Detecting Lagrangian Coherent Structures in Realistic Experimental Fluid-Flow Data Sets

by

Flavio A. C. Martins

A thesis submitted to the
Department of Materials and Mechanical Engineering
in conformity with the requirements for
the degree of Master of Science

Queen’s University
Kingston, Ontario, Canada
-January 2021

Copyright © Flavio A. C. Martins, 2021
Abstract

Detection of coherence in sparse Lagrangian particle tracking (LPT) data is a crucial step towards increasing our understanding in a vast number of biological, engineering and geophysical flows. This thesis explores two Lagrangian-invariant approaches for coherent-structures detection in fluid flows using LPT data. The first tested technique is the Coherent-Structure Colouring (CSC) algorithm (Schlueter-Kuck and Dabiri, 2017a). The performance of this Lagrangian approach is assessed by comparing CSC-coloured tracks with baseline vorticity fields of tracks past an Ahmed reference body and in a swirling jet. The effects of two normalized parameters on the identification of vortical structures were defined and studied: the mean track length; and the mean inter-particle distance. It was found that the CSC algorithm yielded accurate detection of coherent structures when mean inter-particle distances were smaller than 15% of the characteristic length scale of the flow, whereas results quickly deteriorate for sparser Lagrangian data. In the second study, a robust coherent-structure detection framework based on Voronoi tessellation and techniques from spectral graph clustering was developed. In this novel approach, the neighbouring times of two particles, defined as the period of time Voronoi cells of two particles remain connected by a Voronoi edge, is adopted as a metric for coherence. Particles for which Voronoi cells share a common Voronoi edge for longer periods of time present a smaller distance in
the higher-dimensional eigenspace, hence presenting coherent motion. The proposed approach was shown to be successful at identifying coherence with realistic, sparse LPT data from large wind tunnel experiments. Specifically, coherent structures are identified for the first time with inter-particle distances on the order of 100% of the characteristic length scale of the flow.
Co-Authorship

Chapter 3 consists of a completed manuscript that has been accepted for publication. Chapter 4 consists of a manuscript that will be submitted for peer review after completion of this thesis. The citation for each of these publications is as follows:


The copyright of this article is assigned to Springer-Verlag GmbH Germany, part of Springer Nature. The authors retain the right “...to reproduce the Article in whole or in part in any printed volume (book or thesis) written by the Author(s).” Dr. Rival and Dr. Sciacchitano supervised the work, revised and edited the text. Dr. Sciacchitano’s students collected the experimental data. I ran the simulations and wrote the body text.


As this manuscript is not yet published, the copyright is currently retained by the
Authors. Dr. Rival supervised the work. Dr. Rival contributed with helpful discussions, and in the revision of the text. I developed the concept, algorithm and code. I also ran the benchmarks and wrote the body text.
Acknowledgements

This M.A.Sc. thesis is the output of the effort and support of several people to whom I am extremely grateful. First and foremost, I would like to express my sincere gratitude to my supervisor Dr. David Rival for the continuous support of my M.A.Sc. study and research. Your patience and knowledge have greatly contributed to my thesis. Above all, I would like to thank you for the opportunity to study fluid mechanics in depth at one of the most prestigious universities in Canada. I would also like to thank Prof. Piomelli for the best lectures I ever had in fluid mechanics, which greatly contributed to the writing of my thesis.

I am also profoundly grateful for the hard work of my co-authors and for their contribution to my studies. Many thanks to Dr. Linkai, for your kindness and many useful discussions in the OTTER lab. I furthermore am much indebted to Dr. Sciacchitano for all the knowledge you shared and for the many hours you devoted to reviewing my work. I am delighted to have worked with a scientist that puts so much effort into truly understanding experimental fluid mechanics. Your work ethic is truly inspirational, and it is my honest belief that I will one day be able to say that I worked with one of the best experimentalists of my time. Drs. Linkai and Sciacchitano proved to me that scientists can hold great knowledge, while being kind, human, and ethical. You raise my expectations for the future of science.
I was also involved in activities outside the MME Department, where I met amazing people while showcasing Brazilian Culture in Kingston. Thank you so much to all my friends who have supported me for months prior to obtaining this degree through good and bad times. In particular, I would like to thank you Andrew and Alba Sandor Ramos, for all the guidance, the emotional support and for being my family in Canada. I honestly think I would not be able to do it without you. To Esther Greutink, Myron Menezes, Emma Bouillard, Emma Peute, Alex He, Giovanni Pais, Amir Ghorbani, Fernanda Rizzi, Setareh Rahimi, Meli Munoz, Ryo Sakai, John Casnig and Krista Stares, thank you for your camaraderie, encouragement and for all the necessary, but rather inappropriate, jokes. I am confident our coping mechanism for all the hours of intense work we put into our research could not be more efficient. You, my friends, proved me that happiness, kindness and intelligence do not compete for space inside our cognitive system, and with no doubt will contribute to a world where knowledge and arrogance are not strongly correlated. Thank you, gracias, děkuji, bedankt, grazie, merci, arigato, dev bore kurun and obrigado!

I have also had the greatest privilege of meeting Elian Schure. You have inspired me in so many ways that words cannot express it. Ik houd van jou! Finishing my thesis in the middle of a pandemic would be a much harder task if it were not for you. You cherished me with every great moment and supported me whenever I needed it during the most difficult times when writing this thesis. You are a model of a human to me: you are capable of discussing human-defying complex topics while being humble, respectful and cheerful at all times. Your academic talents are truly an inspiration to me. I also extend my thanks to Frank, Nathalie, Stach and Maike Schure, and to Riet and Hub Visschers-Swelsen. You are inspirations of kindness and
I finish with Brazil, where my most basic source of my vital energy resides: my family. I have a truly inspirational and unique family. Their support has been unconditional all these years; they have given up many things for me to be at Queen’s University. Flavio Martins, my father, has proved me that love is a love of giving. He made loans, sold a car, and sacrificed many hours of his life so that I could be here. My M.A.Sc. was a big leap for all of us, and you made it possible. Rachel Avila, my mother, words can not express how much pride I have of being your son. Now, I can hopefully make you proud of having your first born a M.A.Sc. candidate in Mechanical Engineering in Canada. Somehow, we all managed to take me from a poor, public school in Brazil to where I am standing now. To my grandparents: Benedita Avila, Faustino Martins, and Marinete Martins (in memoriam), your existence and trust in me gave me hope and strength. You live in my heart. Eu amo muito vocês! Lastly, I would like to express my gratitude for having you in my life, Camila Martins. I have the luxury of being called your brother, and to witness you tracing a brilliant life. Your sense of justice and your absolute kindness towards all are truly inspirational. I love you all with all my forces, and I could not be more proud of the family I have.

Eu consegui! Eu amo vocês.
List of Acronyms and Symbols

Acronyms

CFD  Computational fluid dynamics
CSC  Coherent-Structure Colouring
ECS  Eulerian coherent structure
FTLE Finite-time Lyapunov exponent
LCS  Lagrangian coherent structure
LPT  Lagrangian particle tracking
PIV  Particle image velocimetry
PTV  Particle tracking velocimetry

Symbols

\( \chi \)  Eigenvector associated with the maximum eigenvalue \( \xi_{max} \)  [–]
\( \delta \)  Dirac’s delta function  [–]
\( \Gamma \)  Fluid domain  [m²]
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Λ</td>
<td>Mean inter-particle distance</td>
<td>[-]</td>
</tr>
<tr>
<td>λ</td>
<td>Eigenvalues</td>
<td>[-]</td>
</tr>
<tr>
<td>A</td>
<td>Adjacency matrix</td>
<td>[-]</td>
</tr>
<tr>
<td>C</td>
<td>Cauchy–Green strain tensor</td>
<td>[s⁻¹]</td>
</tr>
<tr>
<td>D</td>
<td>Degree matrix</td>
<td>[-]</td>
</tr>
<tr>
<td>G</td>
<td>Velocity-gradient tensor</td>
<td>[s⁻¹]</td>
</tr>
<tr>
<td>L</td>
<td>Laplacian matrix</td>
<td>[-]</td>
</tr>
<tr>
<td>M</td>
<td>Material surface</td>
<td>[-]</td>
</tr>
<tr>
<td>S</td>
<td>Symmetric component of the Velocity-gradient tensor</td>
<td>[s⁻¹]</td>
</tr>
<tr>
<td>Ω</td>
<td>Anti-symmetric component of the Velocity-gradient tensor</td>
<td>[s⁻¹]</td>
</tr>
<tr>
<td>ω</td>
<td>Vorticity field</td>
<td>[s⁻¹]</td>
</tr>
<tr>
<td>rᵢⱼ</td>
<td>Mean distance between the i-th and j-th tracks</td>
<td>[m]</td>
</tr>
<tr>
<td>U</td>
<td>Mean particle velocity</td>
<td>[m/s]</td>
</tr>
<tr>
<td>φ</td>
<td>Function of formalized parameters</td>
<td>[-]</td>
</tr>
<tr>
<td>Ψ</td>
<td>Eulerian field</td>
<td>[-]</td>
</tr>
<tr>
<td>ψ</td>
<td>Profile of Eulerian field Ψ</td>
<td>[-]</td>
</tr>
<tr>
<td>τ</td>
<td>Mean characteristic period of a coherent structure, $\tau = l/U_T$</td>
<td>[s]</td>
</tr>
<tr>
<td>xᵢ</td>
<td>Particle’s location vector (xᵢ₁, xᵢ₂, xᵢ₃)</td>
<td>[m]</td>
</tr>
</tbody>
</table>
$\xi$ Eigenvalues of adjacency matrix $\mathcal{A}$ $[-]

C $ Particle concentration, $C = N/D^3, [-]$

$D$ Characteristic length of the flow $[m]$

d $ Euclidean distance $ [m]$

$DT(p_i)$ Delaunay triangulation of particle distribution $p_i$ $[-]

G(V) $ Graph of group $V$ $[-]

l $ Average length of coherent-structures structures $[m]$

$L_s$ Track length $[m]$

$L_t$ Temporal track-length $[-]$

$N$ Total number of tracers $[-]$

$n_{Tij}$ Neighbouring time of the $i$-th and $j$-th particles $[-]$

$p_i$ Particle’s identity $[-]$

$r_{(i,j)}$ Instantaneous distance between the $i$-th and $j$-th tracks $[m]$

$R_{(x,y)}$ Cross-correlation between $x$ and $y$ $[-]$

$Re$ Reynolds number $[m]$

$T$ Characteristic period of the flow $[s]$

$t $ flow time $[s]$

$U_\infty$ Reference velocity $[m/s]$
\[ u_i \quad \text{Velocity vector} \quad [\text{m/s}] \]

\[ U_T \quad \text{Mean tangential velocity of large-scale structures} \quad [\text{m/s}] \]

\[ V(p_i) \quad \text{Voronoi tessellation of particle distribution} \, p_i \quad [-] \]

\[ V_R(p_i) \quad \text{Voronoi cell associated with the particle} \, p_i \quad [-] \]

\[ x_i \quad \text{Position vector}, \, x = x_i \quad [\text{m}] \]
Contents

Abstract i

Co-Authorship iii

Acknowledgements v

List of Acronyms and Symbols viii

Contents xii

List of Tables xv

List of Figures xvi

Chapter 1: Introduction to Theory and Motivation 1

1.1 Fluid motion from a Lagrangian point-of-view 2

1.2 Defining and identifying coherent structures 5

1.3 Overview of engineering and biological applications 6

1.4 Objectives of the thesis 10

1.5 Organization of the Thesis 11

Chapter 2: The Mathematics 13
2.1 Converting track-data into an Eulerian field ........................................... 13
  2.1.1 Discretization of the velocity field based on PIC .............................. 14
  2.1.2 Extracting ECS from velocity fields .............................................. 16
  2.2 Classical LCS theory: FTLE technique ............................................. 18
  2.3 Voronoi tessellation for coherent-structures visualization ..................... 22

Chapter 3: Detection of vortical structures in sparse Lagrangian data using coherent-structure colouring

  3.1 Abstract .................................................................................................. 28
  3.2 Introduction ............................................................................................ 29
  3.3 Methods .................................................................................................. 35
    3.3.1 Mathematical formulation ................................................................. 35
    3.3.2 Normalized parameters ..................................................................... 37
  3.4 Experimental set-up and test cases ......................................................... 38
  3.5 Results and discussion ........................................................................... 42
    3.5.1 Effects of the normalized parameters ............................................... 43
    3.5.2 Correlation between the CSC and Eulerian diagnostics ................... 49
  3.6 Conclusions ............................................................................................ 54

Chapter 4: A robust Voronoi-tessellation-based approach for detection of Lagrangian coherent structures in sparsely-seeded flows

  4.1 Abstract .................................................................................................. 55
  4.2 Introduction ............................................................................................ 56
  4.3 Methods .................................................................................................. 61
List of Tables

2.1 ECS (vortex core) diagnostics based on the $Q$- and $\lambda_2$-criteria (Holmén, 2012). .................................................. 18

4.1 Summary of test cases for the double-gyre flow, where $N$ is the number of tracer particles and $\Lambda$ is the mean inter-particle distance. .................. 73
List of Figures

1.1 Schematics of Lagrangian and Eulerian descriptions for a flow past a circular cylinder. One particle (Lagrangian) trajectory is illustrated in blue, with start and end points at \((x_0, t_0)\) and \((x, t)\), respectively. Particle velocities in the Eulerian frame are set at fixed points in space, spaced by the dashed lines. A square control volume coloured in gray highlights the condition of \(x(t) = X(t)\).  

1.2 (top) Eulerian and (bottom) Lagrangian descriptions of the flow past a square cylinder. \(Q\)-criterion fields show vortical structures defined as convex regions of \(Q > 0\) (in red). In the top row, results obtained with a fine grid are shown. In the bottom row, the same results are extracted with a coarser grid spacing. With a coarser grid spacing, key flow features are not detected using the Eulerian description. In the Lagrangian description, on the other hand, the quality of the results is not necessarily constrained by the number of data points.  

1.3 Map of all the Southern Ocean observational Lagrangian surface drifters in the NOAA GDP Data Set (Lumpkin and Pazos, 2007). Drifters are individually colour-coded, with each data-point being collected every 6h Van Sebille et al. (2018).
2.1 Particle-in-cell based discretization scheme. The velocity field in the Eulerian frame $u_k$ is computed based on statistics of particles contained within the cell. Cells (bins) that do not have particles within their boundaries are treated as empty cells.

2.2 Schematic of the material surface $\mathcal{M}(t)$ evolution as introduced by Haller (2015). Considering a two-dimensional flow $x \in \mathbb{R}^2$, $\mathcal{M}(t)$ evolves under Eq. 2.13. A pathline is shown in the extended state-space $(x, t)$ (line with arrows). Unstable disturbances $y(t_0)$ propagate perpendicularly to $x$.

2.3 Visualizing coherent structures with the FTLE technique. The Locally most attracting material surfaces $\mathcal{M}(t_0)$ (in red) are detected through eigenvalues of $C(x_0) = (\nabla F(t_0))^\top \cdot (\nabla F(t_0))$.

2.4 Decomposition of the space around starts into convex sited as proposed by Descartes (1994). Figure adapted from (Aurenhammer et al., 2013).

2.5 Schematics of a Voronoi diagram. Intersection of planes (---) $D(p_i, p_j)$ define points that are closer to $p_i$ than to $p_j$, forming convex region, $V_R(p_i)$, named Voronoi cells. A set of Voronoi cells define a Voronoi diagram, (—) $V(p_i)$. Neighbouring are naturally identified as particles connected by a single edge of the Delaunay triangulation, $DT(p_i)$.
3.1 (left) Eulerian and (right) Lagrangian descriptions of the flow behind a square cylinder. $Q$-criterion fields show vortical structures defined as convex topologies of $Q > 0$ (in red). In the top row, results obtained with a fine grid are shown. In the bottom row, the same results are extracted with a coarser grid spacing. In the latter, key flow features are not detected using the Eulerian description. In the Lagrangian description, the quality of the results is not necessarily constrained by the number of data points, allowing for the extraction of key flow features even with very sparse data.

3.2 Schematic of fluid domain with key normalized parameters. Fluid travelling at a mean velocity $\overline{U}$ interacts with an obstacle of characteristic length $D$, resulting in a circulation region with structures of radius $l$ and a potential (surrounding) flow. Tracer particles in different flow domains will move with dissimilar kinematic responses, thus being colour-coded accordingly. All tracks possess a temporal track length of $L_t$, but tracks braided in between the large coherent eddies inherit a mean tangential velocity $U_T$ and characteristic period $\tau$. Track lengths $L_s$ are related to the characteristic length of the flow and to the local flow velocity, such that $L_s \sim l \sim D$. 

xviii
3.3 (left) Time-averaged velocity distribution for the flow past the Ahmed body. Iso-contours of the $Q$-criterion ($Q = 4000 \text{ s}^{-2}$) coloured by the stream-wise vorticity, $\omega_x$, highlight the two C-pillar vortices. (center) Normalized stream-wise velocity field, $u_3/U_\infty$, interpolated onto the cross-flow plane $0.5D$ downstream of the rear wall, where $D$ is the height of the Ahmed body. (right) Normalized cross-flow vorticity field, $\omega_3 D/U_\infty$, at the same cross-section position. Reference fields calculated with $N \sim 6.2 \times 10^6$ tracks.

3.4 (left) Time-averaged velocity distribution of the swirling-jet flow. Jet-cores are highlighted by the iso-contours of the absolute-velocity field at $u_i/|u_i| = 0.7$. (center) Normalized stream-wise velocity field, $u_i/U_\infty$, interpolated onto the cross-flow plane $1D$ downstream of the nozzle-exits; and (right) Normalized cross-flow vorticity field, $\omega_1 D/U_\infty$. Reference fields calculated with $N \sim 1 \times 10^6$ tracks.

3.5 (a) Perspective view of the flow past the Ahmed body for $\Lambda/D \approx 0.06$ and $L_s/D \approx 0.8$, and (b) the swirling jet flow, with $\Lambda/D \approx 0.05$ and $L_s/D \approx 0.14$. CSC values range from 1 (dark red) to -1 (dark blue) in the plots.
3.6 CSC-coloured tracks for the flow past the Ahmed body projected onto the cross-flow plane for $x/D \in [0.5, 1.0]$. Rows, top to bottom, show results for $L_s/D \approx 0.4$, 1.0 and 1.6, respectively. Columns, left to right, show results for $\Lambda/D \approx 0.04$, 0.09 and 0.13, respectively. CSC values range in $[-1, 1]$, where -1 corresponds to blue and 1, to red, in the selected colour map. Results indicated as cases (d), (e), ..., (k) in Figs. 3.8a and 3.9a. Last row includes only the two right-most columns because selected $L_s/D$ value would not allow achieving target density magnitude for the third column to be plot. Figures (a) to (c) show vorticity fields obtained with the corresponding $\Lambda/D$ values.

3.7 CSC-coloured tracks for the swirling-jet flow projected onto the cross-flow plane for $x/D \in [-1.0, -0.5]$. Rows, top to bottom, show results for constant $L_s/D$ values of $\approx 0.015$, 0.08 and 0.015, respectively. CSC values range in $[-1, 1]$, where -1 corresponds to blue and 1, to red, in the selected colour map. Columns, left to right, show results for $\Lambda/D$ values of $\approx 0.09$, 0.05 and 0.02, respectively. Results obtained in cases (d), (e), ..., (l) are indicated in Figs. 3.8b and 3.9b. Figures (a) to (c) show vorticity fields obtained with the corresponding $\Lambda/D$ values.

3.8 Cross-correlation function versus normalized inter-particle distance for (a) the flow past the Ahmed body; and the (b) swirling jet flow. Reference cases for the respective benchmarks are highlighted by the dashed lines and refer to results shown in Figs. 3.6 and 3.7. Thick dashed lines indicate the trend of the data points.
3.9 Cross-correlation function versus normalized inter-particle distance for (a) the flow past the Ahmed body; and the (b) swirling jet flow. Reference cases for the respective benchmarks are highlighted by the dashed lines and refer to results shown in Figs. 3.6 and 3.7. Thick dashed lines indicate the trend of the data points.

4.1 Schematic of LCS detection based on Voronoi tessellation and graph clustering. Voronoi diagram is built from the Delaunay triangulation of the seeding particles. Neighbouring Voronoi cells are then identified through the Delaunay triangulation of the tracer distribution. Neighbouring time of Voronoi cells acts as a higher-dimensional criterion for particle clustering, and LCSs are identified as the clusters of Voronoi cells that persist for longer periods of time as neighbouring cells.

4.2 Schematics of coherent-structure detection based on Voronoi tessellation. Voronoi and Delaunay diagrams of the particle distribution $p_i$, $V(p_i)$, $DT(p_i)$, are represented in the graph space, $G(V)$, such that lengths of edges of $DT(p_i)$, $r_{ij}$, are converted into distances in the higher-dimensional eigenspace $G(V)$: $r_{ij} \rightarrow A_{ij} \rightarrow \chi(L(G))$. A Voronoi cell, $V_R(p_i)$, is highlighted in blue in the top-left panel. The top-left panel also highlights in green the triangulation of three particles intersecting a circumcircle with no particles within its interior.

4.3 Normalized vorticity fields for the reference double-gyre flow. The plots exhibit contours of normalized vorticity at different time-steps $t$ obtained with a grid of spatial resolution of $\Delta r_i/D = 0.01$. 

xxi
4.4 Raw tracks shown in top left. Reference normalized cross-flow vorticity field obtained from the point-in-cell interpolation of the tracers trajectories is shown in bottom right. Raw reference fields (without smoothing or interpolation) are plotted in specifics planes normal to the streamwise direction at $x_1/D = [-0.5, 0.5, 1.5, 2.2]$. Results obtained with a constant bin size of $\Delta x_i/D = 0.05$. The direction of the reference far-field velocity, $U_\infty$, and the characteristic length, $D$, are also illustrated.

4.5 Arbitrary scalar fields advected on the double-gyre flow evaluated at $t/T = 2$. Plots on top row reveal fields obtained through the evaluation of 64 tracer trajectories, while bottom row shows results obtained with 484 tracer trajectories. Tracer coordinates at end time are indicated by white circles.

4.6 Selected cases for double-gyre flow at $t = 2T$ with mean inter-particle distance values of $\Lambda \approx 0.31, 0.27, 0.20, 0.13, 0.10, 0.09$ (top-to-bottom). Analyses are visualized by $\chi$-coloured Voronoi cells (left), FTLE-fields (center) and CSC-coloured fields overlaid by CSC-coloured tracks (right). Cases (a) to (f) are listed in Table 4.1.

4.7 Selected cases of double-gyre flow at $t = 2T$ with mean inter-particle distance values of $\Lambda \approx 0.31, 0.13, 0.10, 0.09$ (top to bottom) visualized by $\chi$-coloured Voronoi cells (left) and $\omega$-field (right).

4.8 Profiles $\psi$ of the Voronoi diagram at the symmetry section of the double-gyre flow, $\psi = \Psi(x_2/D = 0.5)$ for different mean inter-particle distance values.
4.9 Left and right panels show $\chi$- and vorticity fields, respectively, obtained from the point-in-cell interpolation of the LPT data onto a grid. Results obtained with $\Lambda \approx 0.065$ at $x_1/D = [-0.5,1.25,2.5]$. Panel to the right also shows $\chi$-coloured tracks.  

4.10 Left and right panels show $\chi$- and vorticity fields, respectively, obtained from the point-in-cell interpolation of the LPT data onto a grid. Results obtained with $\Lambda \approx 0.12$ at $x_1/D = [-0.5,1.25,2.5]$. Panel to the right also shows $\chi$-coloured tracks. 

4.11 Left and right panels show $\chi$- and vorticity fields, respectively, obtained from the point-in-cell interpolation of the LPT data onto a grid. Results obtained with $\Lambda \approx 0.20$ at $x_1/D = [-0.5,1.25,2.5]$. Panel to the right also shows $\chi$-coloured tracks. 

A.1 Flowchart of the proposed LCS detection algorithm. 

A.2 Part I of the clustering algorithm. In this sub-part of the code (see Fig. A.1) track data is converted into a Voronoi diagram. 

A.3 (left) Initial Delaunay triangulation consisting on two convex hulls. (center) the initial configuration using the common edge BD is non-Delaunay. (right) By switching the common edge to AC, the two convex hulls match the Delaunay criterion. 

A.4 Part II of the clustering algorithm. In this sub-part of the code (see Fig. A.1) the kinematic dissimilarity of the Voronoi diagram is processed using a spectral-graph clustering problem.
Chapter 1

Introduction to Theory and Motivation

“At present I absolutely want to paint a starry sky. It often seems to me that night is still more richly coloured than the day; having hues of the most intense violets, blues and greens. If only you pay attention to it you will see that certain stars are lemon-yellow, others pink or a green, blue and forget-me-not brilliance. And without my expatiating on this theme it is obvious that putting little white dots on the blue-black is not enough to paint a starry sky.”

-Vincent Van Gogh

Before focusing on the main objective of this thesis - detection of coherent structures from sparse Lagragian data in fluid flows - we begin by presenting the reader to the foundation of the Lagrangian framework for fluid-dynamic analyses, and how it compares to the classical, Eulerian analyses of coherent structures in fluid flows. In section 1.1, Lagrangian and Eulerian descriptions of fluid flows are presented, and their respective advantages and disadvantages for coherent-structures detection are briefly discussed. In section 1.2, we compare Lagrangian and Eulerian frameworks
for their ability to detect coherence. Subsequently, relevant applications of coherent-structures theory that motivated this study are listed in section 1.3. Finally, we present an overview of the objectives, relevance and organization of the thesis in sections 1.4 and 1.5, respectively.

1.1 Fluid motion from a Lagrangian point-of-view

The essence of Lagrangian fluid dynamics is the notion of fluid particle identity acting as an independent variable (Bennett et al., 2006). Such an identity, or label, may for example be the particle’s position or a thermochemical property. In the Lagrangian framework, the time a particle starts to be tracked, i.e., its “labelling time”, and its position, act as dependent variables of the problem. More specifically, fluid particles are typically described by denoting on one side of the flow equation both their location and labelling time, \((t_0, \mathbf{x})\), and on the other side, the scalar or vectorial property of interest:

\[
g = g(t_0, \mathbf{x}).
\] (1.1)

Moreover, in a somewhat equivalent fashion to the conservation of mass or momentum equations, particles in a fluid flow must also conserve their identity \((t_0, \mathbf{x})\), as discussed by Yadigaroglu and Lahey Jr (1976).

Consider a fluid particle at a measurement time \(t\), and a Lagrangian observer at \(\mathbf{x}\), where \(\mathbf{x} = (x_1, x_2, x_3)\), moving with the particle. An Eulerian observer in the same Cartesian system located at \(\mathbf{X}\) detects the particle if, and only if \(\mathbf{X} = \mathbf{x}(t)\). In addition, based on the Labelling theorem, a particle is uniquely labelled in space-time by specifying its coordinates at the labelling time,
1.1. FLUID MOTION FROM A LAGRANGIAN POINT-OF-VIEW

\[ u_0 = u(t = 0) \]
\[ x_0 = x(t = 0) \]

\[ u = u(t) \]
\[ x(t) = X(t) \]

Figure 1.1: Schematics of Lagrangian and Eulerian descriptions for a flow past a circular cylinder. One particle (Lagrangian) trajectory is illustrated in blue, with start and end points at \((x_0, t_0)\) and \((x, t)\), respectively. Particle velocities in the Eulerian frame are set at fixed points in space, spaced by the dashed lines. A square control volume coloured in gray highlights the condition of \(x(t) = X(t)\).

\[ x_0 \equiv x(t = 0), \quad (1.2) \]

where \(x = x(t, x_0)\) (Bennett et al., 2006). Such characteristics reinforce the fact that the state of a Lagrangian fluid particle is a function of time and labelling-state only (see Fig. 1.1). In addition, scalar and vector quantities, \(q\), can be similarly represented in the Lagrangian framework as

\[ q = q(t, x_0). \quad (1.3) \]

In the Eulerian description, on the other hand, quantities \(q\) are a function of both time and a particle’s coordinate \(x\), \(q = q(x, t)\), being this the first important differentiator between both frameworks (Hauke, 2008).

In classical mechanics, laws of motion are directly applied to describe the response
behaviour of finite amounts of matter (Hauke, 2008). Such structure is readily obtained in the Lagrangian framework, since a particle’s accelerations can be determined by simply adding up all forces acting on it. In the Eulerian framework, however, observations are stationary relative to the moving fluid parcels. Thus, if evolution of particle properties are of interest, a mathematical transformation that allows for the recovery of the derivative following the fluid particle becomes necessary (Serra and Haller, 2016). A transformation of this kind naturally takes the form of a derivative and yields the connection between Eulerian and Lagrangian frameworks.

Consider a generic scalar Eulerian field \( c = c(x, t) \). If, instead of arbitrary \( x \)-values, we measure \( c \) while following the trajectory of a fluid parcel, \( r(t) \), it is clear that \( x \equiv r(t) \) and, hence \( c = c(r(t), t) \). Consequently, the total derivative of \( c \) with respect to time, \( Dc/Dt \), can be computed such as

\[
\frac{Dc}{Dt} = \frac{\partial c}{\partial t} + (u \cdot \nabla)c.
\] (1.4)

Equation 1.4, the material derivative, offers the connection between Eulerian and Lagrangian descriptions of a fluid flow. Terms of the material derivative equation are such that the first term on the right-hand side is referred to as the temporal term, while the second term represents the convective effects. The temporal term describes local (instantaneous) changes in the scalar field, whereas the latter represents changes due to convection (Hauke, 2008). A fluid flow in which the temporal term is null is called stationary, such that \( c \) then varies only in space.
1.2 Defining and identifying coherent structures

Coherent structures are reduced order features in a flow field that allow for the decomposition of the originally higher-order systems into distinct sub-domains (Serra and Haller, 2016). Likewise, Serra and Haller (2016) define coherent structures in two-dimensional flows as the “non-autonomous dynamical systems that are instantly the most influential material curves in the field.” Such broad definitions indicate that there is a wide spectrum of approaches one can select from to identify coherent structures. Nonetheless, using a more practical definition, coherent structures can be simply defined as material surfaces that possess temporal coherence, persisting long enough in the fluid flow so that their statistics can still be assessed (Green, 2012).

Figure 1.2 exemplifies typical results for detecting coherent structures in a flow with techniques from Eulerian and Lagrangian frameworks. Eulerian diagnostics for the detection of Eulerian coherent structures (ECS) have long been studied, and the most widely adopted techniques, such as the $Q$, $\lambda_2$- and $\omega$-fields (Wu et al., 2007), are typically based on decompositions of the velocity-gradient tensor, $\mathcal{G} = (\nabla \mathbf{u})^\top$. Lagrangian coherent structures (LCS), on the other hand, can be defined as material surfaces of the flow field, $\mathbf{u}$, that create patterns in passive tracer distributions advected by an unsteady flow (Haller and Yuan, 2000). The term LCS was coined by Haller (2000) to describe the most repelling, attracting, and shearing material surfaces that form in an unsteady flow. The Lagrangian description of coherence is generally aimed at the macro fluid structures formed by the mixing and entrainment of fluid elements, as opposed to the Eulerian point-of-view, in which flow features are extracted from the instantaneous flow field (Haller, 2000).

Overall, both Eulerian and Lagrangian descriptions of coherence each deliver a set
of advantages and disadvantages. While Eulerian methods have been widely implemented, resulting in many Eulerian applications being well catalogued (Green et al., 2007), Eulerian diagnostic fields can also be heavily dependent on user-defined thresholds (Hadjighasem et al., 2017). Moreover, Eulerian methods are also very sensitive to grid sizing, whereas Lagrangian techniques are not necessarily associated to such limitations (Tu et al., 2018). It is also important to highlight the fact that Lagrangian approaches tend to present somewhat stronger dependence on the homogeneity of the seeding particles distribution, rather than on the particle concentration itself. This last characteristic makes Lagrangian approaches promising candidates for the analysis of sparsely-seeded fluid flows (Hadjighasem et al., 2017).

Moreover, Lagrangian methods for coherence detection present an advantageous framework for the analysis of realistic, sparsely-seeded flow fields: calculations of flow gradients, necessary for the characterization of coherence in classical Eulerian approaches (Tu et al., 2018), require data sets to be dense enough so that linearizations remain sufficiently accurate. In the Lagrangian framework, on the other hand, the analysis of the flow gradients can be replaced by evaluations of tracer trajectories. This characteristic allows for the assessment of coherence in much sparser data sets (see Fig. 1.2).

1.3 Overview of engineering and biological applications

In the laboratory, particle tracking velocimetry (PTV), or Lagrangian particle tracking (LPT), data are obtained at relatively high seeding concentrations by following neutrally buoyant particles suspended in the fluid flow (Schanz et al., 2016). With PTV or LPT techniques, individual tracks illuminated by laser sheets on volumes
1.3. OVERVIEW OF ENGINEERING AND BIOLOGICAL APPLICATIONS

- Eulerian description: Q-criterion

High-resolution data

Low-resolution data

Key flow features are missing

- Lagrangian description: tracer particles

High-resolution data

Low-resolution data

Key flow features detected

Figure 1.2: (top) Eulerian and (bottom) Lagrangian descriptions of the flow past a square cylinder. Q-criterion fields show vortical structures defined as convex regions of \( Q > 0 \) (in red). In the top row, results obtained with a fine grid are shown. In the bottom row, the same results are extracted with a coarser grid spacing. With a coarser grid spacing, key flow features are not detected using the Eulerian description. In the Lagrangian description, on the other hand, the quality of the results is not necessarily constrained by the number of data points.
are tracked in either two- or three-dimensional volumes (Dabiri and Pecora, 2019) in applications that range from the analysis of turbulent flows (Dracos, 2013) to medical research, e.g., the work of Sajja et al. (2011), in which magnetic PTV is applied to the sorting of porcine pancreatic islets. Tracks from LPT measurements typically contain valuable spatio-temporal information of the underlying fluid flow that can be better extracted with Lagrangian approaches.

In the atmosphere, wind velocity, air temperature and gases concentration are routinely measured using free-to-move weather balloons. These measurements per usual have high temporal resolution, but relatively low spatial resolution, imposing great barriers for the analyses of wind flows in the Eulerian framework (Bennett et al., 2006). Direct three-dimensional Lagrangian measurements have also been shown to be especially suitable for increasing our understanding of turbulent transport and mixing of pollution in urban and plant canopy flows, allowing for more accurate modelling of the transport mechanisms of air pollution (Britter and Hanna, 2003). In the atmosphere, sparsely-distributed tracks are followed for great periods of time (on the order of weeks), and traditional Eulerian diagnostics can easily fail to provide comprehensive vortex diagnostics under such measurement conditions (Shnapp et al., 2019).

The ocean surface has been extensively investigated using the Lagrangian framework (Gould, 2001). In oceanic sciences, satellite-tracked drifters are typically used to measure ocean currents (Soreide et al., 2001). Radar, radio, or sound-based drift-tracking technologies have also been routinely adopted (Nichols and Williams, 2009).
1.3. OVERVIEW OF ENGINEERING AND BIOLOGICAL APPLICATIONS

Neutrally buoyant drifters such as the SOFAR, RAFOS, PALACE, and ARGO, following tagged water parcels (Haller and Yuan, 2000), are typically released into various regions of the globe, allowing for more accurate measurements of oceanic flows. Important studies concerning the transport of material in the Gulf of Mexico (Miron et al., 2017), estimations of the directions of salt-water flow (Berglund et al., 2017), or the kinetic energy circulation in the ocean (Shenoi et al., 1999) are just some examples of applications of ocean analysis based on Lagrangian measurements. Examples of habitual track data collected by NOAA GDP are illustrated in Fig. 1.3.

Lagrangian-based analysis of the deep ocean has also been the subject of many studies (Gould, 2001). Historically, Swallow (1955) held first trials of deep-ocean Lagrangian measurements over the Iberian Abyssal Plain in 1955. These first tests with buoyant drifters were followed by measurements with buoyant floats of the flow of Arctic water past the Faroes (Crease, 1965) and the deep sea current of Cape St Vincent (Swallow, 1969). Concomitantly, swallow-float measurements of deep currents in the ocean interior, observed off North Carolina, revealed eddies with time and space scales of tens of days and tens of kilometres (Swallow and Worthington, 1961). As highlighted by Gould (2001), such findings are: “arguably the most significant discovery about the nature of the oceans in the twentieth century”.

In the past years, salinity transportation, current temperatures and deep-sea Lagrangian velocity have gained increased attention from the scientific community aimed at environmental and biological issues (Zambianchi et al., 2017). Deep ocean Lagrangian analysis can answer important questions such as: What causes accelerated mass loss of Antarctic ice sheets? (Cook et al., 2016); and What is the real importance of mixing in up-welling of deep water? (Toggweiler and Samuels, 1998). Such
Figure 1.3: Map of all the Southern Ocean observational Lagrangian surface drifters in the NOAA GDP Data Set (Lumpkin and Pazos, 2007). Drifters are individually colour-coded, with each data-point being collected every 6h Van Sebille et al. (2018).

questions motivate the further development of Lagrangian approaches for coherent structure detection. Therefore, the first objective of this thesis is to investigate state-of-art coherent-structures detection techniques and, based upon the deficits found in such approaches, develop a Lagrangian technique for coherence detection for realistic and very sparse flows.

1.4 Objectives of the thesis

Lagrangian (trajectory-based) analysis allows for improved understanding of biological, oceanic and geophysical problems, in which high spatial resolution can be either impractical or impossible (Mathur et al., 2019). Nevertheless, traditional Lagrangian analyses, e.g., finite-time Lyapunov exponents (Haller, 2001), in which calculations of flow gradients are also needed, are not readily appropriate at large spatio-temporal scales. Owing to the limitations of classical approaches for coherent-structures detection based on sparse data sets, new techniques that do not rely on the existence of
dense data have recently been explored and motivate the present thesis.

Among the recently-developed approaches, the most promising technique for large-scale applications is the Coherent-Structure Colouring (CSC) algorithm, originally tested by Schlueter-Kuck and Dabiri (2017a), and investigated in the present study. Sequentially, based upon the limitations found in the CSC approach for coherent-structures detection with realistic, sparse LPT data, a robust technique based on Voronoi tessellation for coherent-structures detection is presented and analysed. The two primary objectives of this thesis are therefore to:

I. Test the CSC technique capabilities of accurately detecting Lagrangian coherent structures using varying LPT data; and

II. Present and validate a robust technique for coherent-structures detection in very sparsely-seeded flows.

1.5 Organization of the Thesis

We begin the thesis with a thorough review of the fundamental concepts of Eulerian approaches and current coherent-structures theory in Chapter 2, addressing the concepts of FTLE and Voronoi tessellation. Classical point-in-cell approaches for mapping velocity from LPT data onto a grid, necessary to produce ECS, is also presented.

In Chapter 3, we build on our efforts to detect coherent structures on sparse Shake-The-Box (STB) data sets using the CSC algorithm. The performance of the approach is assessed by comparing CSC-coloured tracks with the baseline vorticity fields. The ability to extract vortical structures from sparse data is assessed on two Lagrangian particle tracking data sets: the flow past an Ahmed reference body; and
1.5. ORGANIZATION OF THE THESIS

a swirling jet flow. The effects of two normalized parameters on the identification of vortical structures were defined and studied, namely the mean track length, and the mean inter-particle distance.

In Chapter 4, a robust algorithm to detect coherence in LPT data using Voronoi tessellation and spectral clustering is presented and analysed. The technique is tested with two-dimensional, synthetic data from a double-gyre flow and with challenging LPT data past a bluff body from a large wind tunnel. Results obtained with the technique are compared to the CSC-coloured tracks of the CSC approach and to ridges of the FTLE technique.

In Chapter 5, we finish the thesis by summarizing the application of the two investigated coherent-structures detection approaches, followed by a brief discussion on possible future paths for the development of coherent-structures detection techniques. In Appendix A and Appendix B, we detail flow charts and MATLAB routines adopted throughout this study. Lastly, in Appendix C, we present an in depth comparison between FTLE fields and colour-coded Voronoi cells of the technique presented in Chapter 4.
Chapter 2

The Mathematics

This chapter begins with an overview and introduction to the concept of point-in-cell interpolation for LPT data (section 2.1). This overview is then followed by an introduction to the FTLE technique, focusing primarily on the definition of material surfaces, attracting, and repelling surfaces (section 2.2). In section 2.3, we present a detailed introduction to the concepts of Voronoi diagrams and Delaunay triangulations, in the context of coherent-structures detection.

2.1 Converting track-data into an Eulerian field

In this section, we describe the process of mapping velocity from LPT data onto a grid, and based on the resulting velocity field, we show how ECSs are typically identified. It should be noted that in the Eulerian framework, the term ECS is adopted interchangeably with vortex core, or eddies. The ultimate goal of this section is to present the reader to the classical framework for ECS analyses and to develop a foundation for the comparison of classical Eulerian techniques to the investigated Lagrangian methods.
2.1. CONVERTING TRACK-DATA INTO AN EULERIAN FIELD

2.1.1 Discretization of the velocity field based on PIC

Particle-in-cell (PIC) interpolation is a widely adopted particle method to solve complex computational fluid dynamic problems such as multi-material flows, multi-phase flows or flows with spatial discontinuities (Johnson and Beissel, 1996). It was first introduced by the Fluid Dynamics Group at the Los Alamos National Laboratory, LANL (Harlow, 1957). The PIC is a hybrid Eulerian-Lagrangian technique in which the Lagrangian description moves the seeding particles, whereas the Eulerian description interpolates information onto Eulerian nodes. In this study, the PIC is adopted as a means to produce instantaneous velocity fields using velocity from LPT data.

Consider a two-dimensional fluid-dynamic system that is discrete in space and time, with \( n \) bins labelled \( \Delta x_1, \ldots, \Delta x_n \). We work with a finite number, \( N \), of particles, \( p_i \), in the Euclidean space with Cartesian coordinates \((x_{11}, x_{12}), \ldots, (x_{N1}, x_{N2})\), or location vectors \( \mathbf{x}_i \) (see Fig. 2.1). The \( N \) particles in the Cartesian system are distinct, in the sense that \( \mathbf{x}_i \neq \mathbf{x}_j \), for \( i \neq j \), \( i, j \in I_N = \{1, \ldots, N\} \). Let \( p_i \) be the \( i \)-th tracer particle, with coordinates \((x_{i1}, x_{i2})\), located inside the \( k \)-th bin, \( \Delta x_k \).

Thus, the \( k \)-th component of the velocity field, \( u_k \), can be determined as (Liu and Liu, 2003):

\[
u_k = \frac{1}{N_k} \sum_i U_i \delta(x_i, \Delta x_k), \tag{2.1}\]

where \( N_k \) is an arbitrary number of tracer particles contained within the \( k \)-th bin, \( U_i \) is the velocity vector of the \( i \)-th tracer, and \( \delta \) represents Dirac’s delta function:
2.1. CONVERTING TRACK-DATA INTO AN EULERIAN FIELD

The velocity field is obtained by the weighted summation over the enclosed particles in the cell. Cells without particles are treated as empty cells.

\[ \Delta x_k \]

\[ u_k = \frac{1}{3} \sum_i U_i \]

Figure 2.1: Particle-in-cell based discretization scheme. The velocity field in the Eulerian frame \( u_k \) is computed based on statistics of particles contained within the cell. Cells (bins) that do not have particles within their boundaries are treated as empty cells.

\[ \delta(x_i, \Delta x_k) = \begin{cases} 
1, & \text{if } x_i \in \Delta x_k, \\
0, & \text{otherwise.} 
\end{cases} \quad (2.2) \]

Notice that similar results can be extended to a three-dimensional Cartesian system without loss of generality. Hereafter, grid-indices are omitted for simplicity,
2.1. CONVERTING TRACK-DATA INTO AN EULERIAN FIELD

\( \mathbf{u} \equiv \mathbf{u}_k, \mathbf{x} \equiv \mathbf{x}_k, \) while the subscript notation is adopted to represent tensorial quantities.

2.1.2 Extracting ECS from velocity fields

Subsequent to the definition of \( u_i \equiv u, i \in I_3 = \{1, 2, 3\}, \) traditional Eulerian vortex criteria is specified from the velocity gradient tensor, \( \mathcal{G} = (\nabla \mathbf{u})^T. \) This tensor can be decomposed into its symmetric and anti-symmetric parts, \( \mathcal{G} = S + \Omega, \) such that

\[
S = \frac{1}{2} (\nabla \mathbf{u} + (\nabla \mathbf{u})^T),
\]

and

\[
\Omega = \frac{1}{2} (\nabla \mathbf{u} - (\nabla \mathbf{u})^T),
\]

are referred to as rate-of-strain and vorticity tensors, respectively. Based on these definitions, the eigenvalues, \( \lambda, \) of \( \nabla \mathbf{u} \) satisfy the characteristic equation

\[
\lambda^3 - P\lambda^2 + Q\lambda - R = 0,
\]

where \( P, Q, \) and \( R \) are the three invariants of the velocity gradient tensor (Holmén, 2012). Using the decomposition of the velocity gradient tensor into its symmetric and anti-symmetric parts, these invariants can be expressed as
2.1. CONVERTING TRACK-DATA INTO AN EULERIAN FIELD

\[ P \equiv tr(\mathbf{G}), \quad (2.6) \]
\[ R = det(\mathbf{G}), \quad (2.7) \]
\[ Q \equiv \frac{1}{2} (||\Omega||^2 + ||\mathbf{S}||^2), \quad (2.8) \]

where \( Q \) represents the local balance between shear strain rate and vorticity magnitude. Thus, observing that \( Q \) can be re-written as \( Q = -1/2(\lambda_1 + \lambda_2 + \lambda_3) \) (Jeong and Hussain, 1995), the \( Q \)-criterion defines vortex cores as all elliptic regions in the flow where \( Q(\mathbf{x}) > 0 \). These regions are naturally interconnected by hyperbolic subdomains of \( Q(\mathbf{x}) < 0 \) (JCR et al., 1988).

Another important Eulerian vortex detection approach is the \( \lambda_2 \)-criterion, presented by Jeong and Hussain (1995). Taking the gradient of the Navier-Stokes equations in subscript notation, \( i, j, k \in I_3 = \{1, 2, 3\} \):

\[ a_{ij} = -\frac{1}{\rho} p_{ij} + \nu u_{ijkk}, \quad (2.9) \]

where \( a_{ij} \) is the acceleration gradient, \( \rho \) is the fluid density, and \( p_{ij} \) is the pressure field, and expanding its symmetric part, the following expression can be obtained:

\[ \frac{D\mathbf{S}_{ij}}{Dt} - \nu \delta_{ij} \Omega_{kk} + \Omega_{ik} \Omega_{kj} + \mathbf{S}_{ik} \mathbf{S}_{kj} = -\frac{1}{\rho} p_{ij}. \quad (2.10) \]

The first term on the left-hand side of Eq. 2.10 represents unsteady irrotational strain, whereas the second term, from left to right, represents viscous effects. Considering the third and fourth terms on the left-hand side of Eq. 2.10, \( ||\mathbf{S}||^2 + ||\Omega||^2 \), a vortex core can be defined as all interconnected regions with two negative eigenvalues,
Table 2.1: ECS (vortex core) diagnostics based on the $Q$- and $\lambda_2$-criteria (Holmén, 2012).

<table>
<thead>
<tr>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\lambda_3$</th>
<th>$\sum_i \lambda_i$</th>
<th>Negative $\lambda_2$</th>
<th>Positive $Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>vortex core</td>
<td>vortex core</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>vortex core</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>vortex core</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

$\lambda_i$, of $||S||^2 + ||\Omega||^2$. By ordering the three eigenvalues, $\lambda_1 \leq \lambda_2 \leq \lambda_3$, and considering that the resultant tensor is symmetric (Green, 2012), a vortex core is defined in the $\lambda_2$-criterion framework as all regions in a flow where $\lambda_2 < 0$. Table 2.1 summarizes the relationship between $Q$- and $\lambda_2$-criteria to the eigenvalues $\lambda_i$ for ECS detection.

### 2.2 Classical LCS theory: FTLE technique

The velocity field $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ generates trajectories through the differential equation

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

(2.11)

for $\mathbf{x} = (x_1, x_2, x_3)$, and $t \in [t_0, t_1]$. The solution of Eq. 2.11 can be denoted by $\mathbf{x}(t; t_0, \mathbf{x}_0)$, in which $\mathbf{x}_0$ is the initial position of the trajectory, at $t = t_0$. In fluid mechanics, Lagrangian trajectories $\mathbf{x}(t)$ are referred to as pathlines (Jeong and Hussain, 1995). Typical pathlines take the form of self-intersecting curves in unsteady flows (see Fig. 1.1), becoming more parallel to one another the more steady the flow.

The definition of material surfaces and Lagrangian coherence starts from the concept of the flow map, an operator that is central to the classical Lagrangian description of a flow field, defined as
\[ F^t_{t_0}(x_0) := x(t; t_0, x_0). \] (2.12)

The flow map constitutes an invertible matrix that relates an initial condition at \( t_0 \) to its current state at \( t \), \( F^t_{t_0} : x_0 \rightarrow x(t; t_0, x_0) \). Consider a hyper-surface, i.e., a material surface \( \mathcal{M}(t_0) \), of dimension one less than that of \( u \). The time-\( t \) position and shape of such surface can be determined as

\[ \mathcal{M}(t) := F^t_{t_0}(\mathcal{M}(t_0)), \] (2.13)

and the dynamic stability of the pathlines can then be studied in terms of the stability of \( \mathcal{M}(t) \). Stable and unstable material surfaces create the reciprocal response behaviour in tracer mixing, as is illustrated in Fig. 2.2. In that sense, instabilities of interest, \( y(t) \), present exponential growth and are normal to \( x \), evolving as

\[ y(t) = \nabla F^t_{t_0}(x_0)y(t_0). \] (2.14)

Taking the squared magnitude of \( y(t) \), and replacing \( y \) using Eq. 2.14, the following identity can be obtained

\[
|y(t)|^2 = \langle \nabla F^t_{t_0}(x_0)y(t_0), \nabla F^t_{t_0}(x_0)y(t_0) \rangle \\
= \langle y(t_0), [F^t_{t_0}(x_0)]^\top F^t_{t_0}(x_0)y(t_0) \rangle \\
= \langle y(t_0), C(x_0)y(t_0) \rangle,
\] (2.15)

in which \( C(x_0) = (\nabla F^t_{t_0})^\top \cdot (\nabla F^t_{t_0}) \) is the symmetric and positive-definite right Cauchy-Green strain tensor (Blume, 1989). Computing \( C(x_0) \) requires solving the deformation
2.2. CLASSICAL LCS THEORY: FTLE TECHNIQUE

Figure 2.2: Schematic of the material surface $M(t)$ evolution as introduced by Haller (2015). Considering a two-dimensional flow $x \in \mathbb{R}^2$, $M(t)$ evolves under Eq. 2.13. A pathline is shown in the extended state-space $(x, t)$ (line with arrows). Unstable disturbances $y(t_0)$ propagate perpendicularly to $x$.

gradient $\nabla F_t$, which has prohibitive computational cost in three dimensions and is still a computationally-intensive process in two-dimensions (Haller, 2015). The determination of the evolution of material surfaces based on the deformation gradient is also very sensitive to grid sizing. Moreover, results are also known for becoming noisy depending on initial conditions and adopted discretization scheme (Haller, 2000).

The eigenvalues $\lambda_i(x_0) = \lambda_1(x_0), \ldots, \lambda_m(x_0), x_0 \equiv x_j(t_0) \mid x_j \in \mathbb{R}^m$, of $C(x_0)$ are such that an infinitesimally small sphere centred at $x_0$, carried by the mean flow through $x(t; t_0, x_0)$, deforms into an ellipsoid of which the $i$-th principle axis is proportional to $\sqrt{\lambda_i(x_0)}$. Observing that $\lambda_1(x_0) < \lambda_2(x_0) < 1 < \lambda_m(x_0)$, the average of the largest stretching is directly proportional to $\sqrt{\lambda_m(x_0)}$ and typically represents exponential growth. The exponent of such growth is given by

$$FTLE_{t_0}^{t_1}(x_0) = \frac{1}{|t_1 - t_0|} \ln \left( \sqrt{\lambda_m(x_0)} \right),$$  (2.16)
FTLE technique

Seeding particles

Calculate flow gradients

FTLE field

Figure 2.3: Visualizing coherent structures with the FTLE technique. The Locally most attracting material surfaces $\mathcal{M}(t_0)$ (in red) are detected through eigenvalues of $C(x_0) = (\nabla F_{t_0}^t)^\top \cdot (\nabla F_{t_0}^t)$.

and is known as the Finite-time Lyapunov exponent (FTLE) at $x_0$. Plots of FTLE reveal mixing regions in the flow, presenting the most repelling and attracting coherent-structures. Figure 2.3 shows a schematic diagram on how the FTLE technique identifies coherence based upon the detection of attracting LCS (in red), $\mathcal{M}(x_0)$. If integrated backwards in time, ridges of the FTLE field would reveal repelling coherent-structures.
2.3 Voronoi tessellation for coherent-structures visualization

The coherent structure detection mechanism presented in this thesis is based upon the definition of the so called Voronoi diagram, of which origin dates back to the 17th century (Aurenhammer et al., 2013). René Descartes was one of the first philosophers to claim that our solar system can be sectioned in cells with topologies of vortices (Descartes, 1994), illustrating in his work the decomposition of the space into convex regions revolving around stars (see Fig. 2.4). These regions can be mathematically defined such that for every coordinate $x$, one can specify a site in which a particle $p$ exerts stronger influence on than any other particle.

The concept of Voronoi tessellation has emerged in an independent fashion, with proven usefulness for many research fields. The definition of space, site and influence of a Voronoi diagram can also vary greatly depending upon the application (Voronoi, 1908). As in the summary written by Aurenhammer et al. (2013), Voronoi diagrams can be adopted for numerous and diverse category of applications, e.g., geography and meteorology problems (Shi and Pang, 2000); crystallography (Blatov, 2004), and it plays a central hole in the study of Wigner-Seitz zones, for chemistry and physics (Saye and Sethian, 2011). Voronoi diagrams have also been proved usefull for fluid mechanics, being used as tools for pressure-field extraction on unstructured flow data (Neeteson and Rival, 2015), and as mesh-free alternatives for computing FTLE ridges based on tessellation of unstructured data (Rosi et al., 2015).

Here, I will define the Voronoi diagram and introduce relevant notation and properties to be commonly used in this study. Recalling the notation of Cartesian coordinates in $\mathbb{R}^2$, $x_i = (x_{i1}, x_{i2})$, the Euclidean distance, $d$, of any point $x_i$ to a tracer $p$, with coordinates $(x_1, x_2)$, is given by:
2.3. VORONOI TESSELLATION FOR COHERENT-STRUCTURES VISUALIZATION

Figure 2.4: Decomposition of the space around starts into convex sited as proposed by Descartes (1994). Figure adapted from (Aurenhammer et al., 2013)

\[ d(p, x_i) = \sqrt{(x_1 - x_{i1})^2 + (x_2 - x_{i2})^2}. \]  

(2.17)

Defining the bisector of two distinct points \( p_i \) and \( p_j \) (\( x_i \neq x_j \) if \( i \neq j \)) as \( B(p_i, p_j) \), and taking into account the fact that a bisector is defined as all the points that are at equal distance from both \( p_i \) and \( p_j \), it is evident that \( B(p_i, p_j) \) will lay perpendicular at the middle section of \( \overline{p_i p_j} \), separating the points \( x \) that are closer members to \( p_i \) than to \( p_j \) in the half plane \( D(p_i, p_j) \):

\[ D(p_i, p_j) = \{ x \mid d(p_i, x) \leq d(p_j, x) \}. \]  

(2.18)
The half-plane that consists of points closer to \( p_j \) is denoted by \( D(p_j, p_i) \), and \textit{vice-versa}. Consequently, the intersection of the planes \( D(p_i, p_j) \) surrounding an arbitrary particle \( p_i \), \( V_R(p_i) \), is such that:

\[
V_R(p_i) = \bigcap_{p_i \in \mathbb{R}^2, p_i \neq p_j} D(p_i, p_j). \tag{2.19}
\]

and consists of all points \( x \in \mathbb{R}^2 \) such that \( p_i \) is the nearest tracer. This definition leads to non-overlapping, convex polygons surrounding every tracer \( p_i \) in regions of which all interior points are closer member to than to any other tracer, \( p_j \).

Let \( P = \{p_1, \cdots, p_N\} \) be the collection of the \( N \) distinct tracers in the flow, such that an arbitrary tracer \( p_i \) has coordinates \( x_i \). The polygon defined in Eq. 2.19 is then given by (Ferrero, 2011)

\[
V_R(p_i) = \{x \text{ s.t. } \|x - x_i\| \leq \|x - x_j\| \text{ for } j \neq i, j \in I_N\}, \tag{2.20}
\]

and defines the (ordinary) Voronoi cell associated with the tracer \( p_i \). Axiomatically, the set of all Voronoi cells generated by \( P \), given by

\[
V(p_i) = \{V_R(p_1), \cdots, V_R(p_N)\}, \tag{2.21}
\]

is the Voronoi diagram of the seeding distribution, \( P \). The same definitions can be extended to the \( m \)-dimensional Euclidean space without loss of generality.

Ordinary Voronoi diagrams, \( m \leq 3 \), can be sub-divided into two definitions based upon the number of dimensions of the Euclidean space: in \( \mathbb{R}^2 \), a planar Voronoi diagram are generated by \( P \), whereas in \( \mathbb{R}^3 \), a (higher-order) three-dimensional Voronoi diagram is obtained. Thus, we shall often refer to both planar and three-dimensional
Voronoi diagrams simply as a Voronoi diagram. Furthermore, the edge shared by two Voronoi cells is referred to as a Voronoi edge, whereas the vertices of the Voronoi edges are similarly named as Voronoi vertices. Figure 2.5 shows a schematic representation of all defined entities.

From an intuitive point of view, Voronoi cells can be similarly obtained by expanding a circular surface around each tracer \( p_i \) until their boundaries collapse (Okabe et al., 2009). The resultant interface formed at the intersection of the boundaries is equivalent to those formally defined in Eq. 2.19. Denoting an expanding circle as \( \gamma \), it can take three forms:

i. If \( \gamma \) hits one site \( x_i \), it belong to \( V_R(p_i) \);
ii. If $\gamma$ hits two sites, $x_i$ and $x_j$, it defines the Voronoi edges of $p_i$ and $p_j$;

iii. If $\gamma$ hits three sites, $x_i$, $x_j$, and $x_k$ simultaneously, it is a Voronoi vertex in between $p_i$, $p_j$ and $p_k$. Vertices typically have up to third-degree connections (Okabe et al., 2009).

It is also important to highlight the fact that boundaries of an $m$-dimensional Voronoi cell are $(m-1)$-dimensional faces. Thus, three-dimensional Voronoi polyhedrons are bounded by two-dimensional surfaces, while two dimensional Voronoi cells are bounded by Voronoi edges. In practice, two-dimensional Voronoi cells provide a visual inspection of the flow that is similar in nature to that of classical Eulerian fields without requiring a grid (see Fig. 2.5). Nonetheless, visualization of three-dimensional Voronoi diagrams is challenging and, in the present study, will be avoided in the results.

The next important concept behind the development of the coherent-structures detection technique presented in this thesis is the Delaunay triangulation, $DT(p_i)$. The formal definition of $DT(p_i)$ is such that edges connecting any two points $p_i$ and $p_j$, for which a circle exists that crosses both points, but does not contain any other sites within its boundaries, defines a Delaunay diagram. Thus, Delaunay edges form straight-line triangulations of particles that are closest to one another. Schematic representation of a Delaunay triangulation is shown by dashed lines in Fig. 2.5.

Edges of $DT(p_i)$ connect closest tracer particles. Thus, in this study tracers connected by one Delaunay edge are referred to as Voronoi neighbours. More importantly, if two tracers are connected by a single Delaunay edge, their Voronoi cells axiomatically share a common Voronoi edge. This last property is fundamental for the development of the proposed coherent-structures detection technique. In the proposed
approach, the *neighbouring time* of every two tracers, i.e., the period of time two Voronoi cells remain connected by a single Voronoi edge, is adopted as a metric for coherence. Tracer particles of which neighbouring time remain high throughout the period of analysis are naturally attributed higher degree of coherence between their motion, having higher chances of belonging to the same coherent-structures. Moreover, the process of colour-coding Voronoi cells based on their neighbouring times results in colour patterns that are similar in nature to those observed in the classical FTLE fields. An in depth comparison between results obtained with the FTLE technique and the coherent-structures detection approach proposed in this study is carried out in the Appendix C.
Chapter 3

Detection of vortical structures in sparse Lagrangian data using coherent-structure colouring

3.1 Abstract

In this study, vortical structures are detected on sparse Shake-The-Box data sets using the Coherent-Structure Colouring (CSC) algorithm. The performance of this Lagrangian approach is assessed by comparing the CSC-coloured tracks with the baseline vorticity field. The ability to extract vortical structures from sparse data is accessed on two Lagrangian particle tracking data sets: the flow past an Ahmed body; and a swirling jet flow. The effects of two normalized parameters on the identification of vortical structures were defined and studied: the mean track length; and the mean inter-particle distance. The accuracy of the vortical-structure detection problem through CSC is shown to improve with decreasing inter-particle distance values, whereas little dependence on the mean track length is observed at all. Overall,
the CSC algorithm showed to yield accurate detection of coherent structures for inter-particle distances smaller than 15% of the characteristic dimension of the structure. However, the results quickly deteriorate for sparser Lagrangian data.

3.2 Introduction

LPT is increasingly used due to its ability to quantity fluid-parcel trajectories in three-dimensional volumes. LPT is in principle suitable for the identification of LCSs, i.e. repelling, attracting and shearing material surfaces can be extracted directly (Haller, 2000; Haller and Yuan, 2000). The identification of LCSs allows for a simplified assessment of a flow’s topology, presenting a complete quantitation of material transport within an evolving fluid system (Haller, 2015). In contrast, well-known Eulerian vortex detection methods are based on the decomposition of the velocity gradient tensor (Jeong and Hussain, 1995), relying on the existence of relatively dense experimental data. Lagrangian methods are, however, not necessarily constrained by such requirements (Schlueter-Kuck and Dabiri, 2017b). This last feature drives the development of new post-processing strategies for the detection of coherent structures with data obtained from large-scale applications (Neamtu-Halic et al., 2019; Wei et al., 2019).

Lagrangian (trajectory-based) analysis allows for improved understanding of various processes that take place in the ocean, atmosphere and even in the lab, among other problems, which often suffer from poor spatial resolution (Mathur et al., 2019). In most of these cases, high spatial resolution is either impractical or impossible (Schmale III and Ross, 2015), resulting in incomplete and unreliable diagnostic fields. Resulting low signal-to-noise ratios motivate the development of alternative approaches
3.2. INTRODUCTION

that allow for accurate feature-detection in sparse data. Some of these studies concerning alternative approaches propose an Eulerian framework based on data assimilation (DA) (Ghil and Malanotte-Rizzoli, 1991; Rodell et al., 2004), relying on the mix between numerical models and experimental data to create hybrid fields that generally have reduced uncertainty when compared to the original data sets. DA has been extensively adopted in weather forecasting, and has proved to be successful in detecting flow features in many complex systems such as the ocean–atmosphere coupling model (Schiller et al., 1997) and the Arctic polar-vortex phenomenon (Swinbank and O’Neill, 1994). Other examples where underlying flow dynamics have been revealed include oceanic, geophysical and atmospheric phenomena; see (Bocquet et al., 2010; Cummings, 2005; Rodell et al., 2004). Notwithstanding, due to the necessity of a physical model, DA may not yet be possible for a variety of complex engineering, environmental and biological problems, e.g., the transport of pathogens in atmospheric flows (Schmale III and Ross, 2015) or the advection of micro-organisms (Tallapragada et al., 2011). In such cases, the existence of a physical model is limited and the scales of the flows, up to thousands of kilometres, would make the task of detecting LCSs using DA challenging if not impossible.

Currently, flow structures from LPT data are typically identified using an Eulerian framework, by mapping the sparse LPT data onto a Cartesian grid, either through interpolation (Stüer and Blaser, 2000) or by enforcing compliance with physical laws, e.g., by reconstructing scattered data using B-spline curves (Gesemann et al., 2016) or by interpolating particle track measurements using vortex models (Schneiders et al., 2017). Eulerian coherent structures are generally detected using diagnostic fields
derived from the velocity gradient tensor. In such cases, cores of the vortical structures are defined as the connected regions in the flow where the velocity gradient tensor satisfies the product of the entries on its secondary diagonal being negative and smaller the product of the elements on the main diagonal (Chen et al., 2015). Classical examples of Eulerian methods include the $Q$, $\Lambda_2$- and $\omega$-fields (Zhan et al., 2019), all of those requiring that the velocity field exhibits a swirling pattern (Robinson, 1991). But even though Eulerian techniques have been extensively adopted for the diagnostics and quantification of material surfaces, dense data sets are necessary for accurate estimation of the velocity gradient tensor (Green et al., 2007). Eulerian approaches also suffer from three major drawbacks: they are highly-sensitive to user-defined thresholds, the flow history is not taken into account in individual measurements (Haller, 2015; Rosi et al., 2015), and observations conducted using Eulerian techniques change for non-inertial coordinate transformations (Jeong and Hussain, 1995). Lagrangian descriptions of flow fields, on the other hand, allow for the observation of time-evolving flow features, and diagnostics are independent of the reference frame (Peacock and Dabiri, 2010). Figure 3.1 illustrates typical results obtained from Eulerian and Lagrangian approaches with both densely and sparsely reconstructed flow fields for the canonical flow behind a square cylinder. Although widely adopted, classical Eulerian detection mechanisms are highly sensitive to vector spacing, and can result in large errors due to the larger-scale grid spacing of readily available atmospheric model data or the lack of high-resolution measurements in biological problems (e.g. (Sutton et al., 1994)).

It is also important to highlight the fact that traditional Lagrangian analyses,
3.2. INTRODUCTION

Lagrangian description: tracer particles

High-resolution data

Low-resolution data

Figure 3.1: (left) Eulerian and (right) Lagrangian descriptions of the flow behind a square cylinder. $Q$-criterion fields show vortical structures defined as convex topologies of $Q > 0$ (in red). In the top row, results obtained with a fine grid are shown. In the bottom row, the same results are extracted with a coarser grid spacing. In the latter, key flow features are not detected using the Eulerian description. In the Lagrangian description, the quality of the results is not necessarily constrained by the number of data points, allowing for the extraction of key flow features even with very sparse data.

in which a velocity field is needed, are also not readily appropriate at large spatio-temporal scales. Some examples include the application of the classical finite-time Lyapunov exponent, FTLE (Haller, 2001) and the finite-size Lyapunov exponent, FSLE (Boffetta et al., 2001). FTLE-fields have been extensively investigated in the past years to locate dynamical regions in geophysical flows (Du Toit, 2010; Peacock and Haller, 2013; Pierrehumbert, 1991), and for the detection of transport barriers in canonical flows (Brunton and Rowley, 2010; Kasten et al., 2010). Nonetheless, despite
3.2. INTRODUCTION

the number of successful applications of the FTLE approach, the use of chaos-theory-based techniques for material-coherence detection presents several limitations, such as the fact that a discrete version of the flow field is naturally required. Furthermore, the analysis of the gradient tensor requires tracer trajectories to be sufficiently close so that necessary linearizations of the local velocity field’s gradients remain accurate (Chu et al., 2003; Schlüeter-Kuck and Dabiri, 2017a). Owning to the obstacles of classical approaches for LCS-detection based on sparse data sets, new techniques that do not rely on the existence of dense data have recently been explored. Some of these approaches have been analysed by (Hadjighasem et al., 2017) and motivate the current study.

Among these recently-developed approaches, the most promising technique for large-scale applications is the Coherent-Structure Colouring (CSC) algorithm proposed by (Schlüeter-Kuck and Dabiri, 2017a), and applied in the present study. The CSC belongs to a category of Lagrangian methods that are based on spectral graph theory, originally proposed by Bezdek et al. (1984). The algorithm compares the spatial-evolution of an arbitrary number of tracer particles where tracks are colour-coded based on their dissimilarities. In practice, particle trajectories that present relatively small spatial deviation from each other are coloured with similar CSC values. Additionally, from a mathematical point-of-view, the kinematic dissimilarity between a pair of particles is not based upon the definition of a global coordinate system, making the CSC approach Lagrangian invariant (Haller, 2015). Moreover, clustering techniques are not necessarily constrained by a minimum number of data points, meaning that the technique can still operate with indefinitely sparse data sets (Von Luxburg et al., 2008). Typical results obtained with the CSC approach are
exemplified on the right-hand side of Fig. 3.1.

The convergence of coherent structures detected with CSC for an increasing number of tracer particles is to be expected, and has been verified by Schlueter-Kuck and Dabiri (2017b) in both two-dimensional flows and for synthetic data (see also (Husic et al., 2019)). However, results obtained in these studies do not represent realistic data from large-scale, three-dimensional measurements. The ability to detect LCSs using CSC has only been tested with a relatively high concentration of synthetic tracer particles, precluding the findings of the original papers (Schlueter-Kuck and Dabiri, 2017a,b) to be extrapolated to real or complex atmospheric, oceanic or biological applications. Other questions still remain: for a realistic, three-dimensional flow, what is the minimum recommended particle concentration that would still allow for coherent structure identification in CSC? What is the impact of track length on the overall quality of feature extraction? How does the CSC approach compare to classical Eulerian coherent-structure identification? What are important flow parameters to guarantee consistent results when using the CSC algorithm? In this paper, we address these above questions by using Shake-The-Box data from two benchmark flows: the flow-past the Ahmed body; and a swirling jet flow. Tracks colour-coded using CSC will be compared to the baseline vorticity field, and the cross-correlation between features detected with both diagnostics are calculated and plotted versus flow parameters defined in the subsequent sections.

The current paper is organized as follows: In Section 3.3 we briefly introduce the mathematical background and notation for the implementation of CSC, and describe the adopted key normalized parameters. In Section 3.4, the experimental set-ups for two benchmark flows are presented. In Section 3.5, the correlation analysis between
3.3. METHODS

CSC and the baseline vorticity field are revealed. Lastly, conditions and limitations for the use of the technique with three-dimensional Lagrangian data are established.

3.3 Methods

The following section is organized such that in Section 3.3.1 the mathematical description of the CSC method and the adopted notation are presented succinctly. In Section 3.3.2, normalized parameters for the investigation of the CSC method in realistic flows are developed and discussed.

3.3.1 Mathematical formulation

This section summarizes the CSC algorithm as recently explored by Schlueter-Kuck and Dabiri (2017a). The approach relies on the concept of a graph, a mathematical structure that consists of a set of nodes that are interconnected by edges. Nodes generally represent entities, while the edges describe the relationship between these entities, by both informing rather a pair of nodes is connected or not and what is the weight of that connection. In the CSC framework, the adopted graph is defined such that its nodes represent the measured Lagrangian particles, whilst edges represent the relationship between pairs of particles, being weighted by the kinematic dissimilarity of the respective tracers’ trajectories. Such kinematic dissimilarity is defined as the standard deviation of the distance between two tracer particles during the measurement time, \( t = (L_t - 1)\Delta t \), where \( L_t \) and \( \Delta t \) represent the temporal track length (number of samples within a track) and the measurement’s time step, respectively.

Considering that \( N \) fluid particles with temporal track length of \( L_t \) are known, the weighted adjacency matrix \( \mathbf{A} \) can be defined as:
3.3. METHODS

\[ A_{ij} = \frac{1}{\tau_{ij} L^{1/2}} \left[ \sum_{k=0}^{L-1} (\tau_{ij} - r_{ij}(t_k))^2 \right]^{1/2}, \]  

(3.1)

which contains the weight of the edge that connects the \( i \)-th to the \( j \)-th particle. In Eq. 3.1, \( \tau_{ij} \) is the average distance between the two particles during the measurement time, whereas \( r_{ij}(t_k) \) represents their instantaneous distance at the \( k \)-th time-step. The parameter \( A_{ij} \) is larger for particles that are kinematically dissimilar, whereas it is zero for particles that keep their distance constant over time. For instance, if a pair of particles \((i, j)\) move along straight parallel trajectories at the same velocity, then \( A_{ij} = 0 \). Contrarily, if the \( i \)-th particle moves along a straight trajectory whereas the \( j \)-th moves along a curvilinear one, then \( A_{ij} > 0 \).

Moreover, the degree matrix of a graph, \( D \), defined as:

\[
D_{ij} = \begin{cases} 
0, & \text{if } i \neq j \\
\sum_{k=1}^{N} A_{ik}, & \text{if } i = j,
\end{cases}
\]  

(3.2)

contains in the \( k \)-th diagonal element the sum of the \( k \)-th row of the adjacency matrix \( A \). Defining now the graph Laplacian as \( L = D - A \), the graph colouring problem is such that node pairs with relatively large weights are associated with dissimilar values. This problem is equivalent to maximizing the expression:

\[
z = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - x_j)^2 A_{ij} = X' LX,
\]  

(3.3)

where \( X \) is a row vector containing the value of CSC of each particle, such that kinematically-dissimilar particles will be assigned largely varying \( z \)-values (Schlueter-Kuck and Dabiri, 2017a). The problem of maximizing the expression in Eq. 3.3 is
equivalent to that of finding the maximum eigenvalue $\xi_{\text{max}}$ of the associated eigenvalues/eigenvectors system, $\mathcal{L}X = \xi D X$, such that $CSC \equiv X(\xi_{\text{max}})$. The components of the eigenvector $CSC$ associated with the maximum eigenvalue $\xi_{\text{max}}$ represent the CSC of each particle. Particles with values of $CSC$ similar in magnitude correspond to particles that are kinematically similar, and therefore that belong to the same Lagrangian coherent structure. The CSC is numerically implemented in MATLAB, and a synthetic pseudo-code is given in the B.

### 3.3.2 Normalized parameters

Considering a volume of fluid that scales with the characteristic length of the flow, $D$, and a homogeneous seeding distribution, it has been conjectured that accurate identification of LCSs is dependent on two key parameters: the track length $L_s$ of the tracer particles; and the number $N$ of particles present in the volume. The normalization of these parameters will be presented when considering a generic flow over a bluff body.

It is conjectured that the accuracy of the LCS identification based on the CSC method is dependent on the amount of the Lagrangian information in the flow field, relative to the size of the coherent structures. Hence, two relevant parameters are expected to play a role in the LCS identification: $L_s/D$ and $\Lambda/D$, where $D$ is the characteristic length of the flow, $L_s$ is the track length ($L_s = \bar{U}(L_t - 1)$, being $\bar{U}$ the particle mean velocity along the track), and $\Lambda = [3/(4\pi C)]^{1/3}$ is the average inter-particle distance, being $C$ the tracers concentration.

In general, $L_s$ varies in space and time due to the flow non-uniformity. However, in the following analysis, the average value of $L_s$ is considered. Also, a uniform
seeding concentration is assumed, so that local and average inter-particle distances approximately coincide. From this analysis, the correlation between the results obtained with the CSC approach, and a reference Eulerian field $\Psi$, are a function of the following $\Pi$-groups:

\[
R_{\text{CSC,}\Psi} = \phi (\Pi_1, \Pi_2) = \phi \left( \frac{\Lambda}{D}, \frac{L_s}{D} \right),
\]

(3.4)

where $\Pi_1$ and $\Pi_2$ refer to the normalized inter-particle distance and normalized mean track length, respectively.

The accuracy of the CSC-based LCS identification is assessed by comparing the CSC with a standard Eulerian approach for vortex identification, namely the vorticity field. From preliminary results, it was concluded that the CSC approach holds strong correlation with the absolute-velocity and vorticity fields, but weak to null correlation with the $Q$ or $\Lambda_2$ criteria. The reasons for that will be further discussed in the the Results section of this work.

### 3.4 Experimental set-up and test cases

The first benchmark performed in this study is based on the flow behind a reference Ahmed body. Originally developed to investigate the features of flows around ground vehicles (Ahmed, 1983), this simplified car model comprises a rounded fore-body, a box-like middle section and a slanted surface (slant angle of $25^\circ$) at the rear with rounded edges. The near-wake flow structure comprises two re-circulation regions, and two longitudinal C-pillar vortices. The C-pillar-like structures are the most dominant characteristic material surfaces present in the near-wake region of the flow behind the Ahmed body, and are generated by the pressure difference between the flow coming
3.4. EXPERIMENTAL SET-UP AND TEST CASES

Figure 3.2: Schematic of fluid domain with key normalized parameters. Fluid travelling at a mean velocity $\overline{U}$ interacts with an obstacle of characteristic length $D$, resulting in a circulation region with structures of radius $l$ and a potential (surrounding) flow. Tracer particles in different flow domains will move with dissimilar kinematic responses, thus being colour-coded accordingly. All tracks possess a temporal track length of $L_t$, but tracks braided in between the large coherent eddies inherit a mean tangential velocity $U_T$ and characteristic period $\tau$. Track lengths $L_s$ are related to the characteristic length of the flow and to the local flow velocity, such that $L_s \sim l \sim D$.

from the vertical (lateral) walls of the model and the flow over the slanted surface (Ahmed et al., 1984). The flow behind the Ahmed body is presented in Fig. 3.3.

Data for the flow behind the Ahmed body was collected in the Open Jet Facility (OJF) of the TU Delft Aerodynamics Laboratories. The facility consists of a closed-loop, open test section wind tunnel with octagonal exit-section. The Ahmed body model is a 1 : 2 replica of the reference geometry, with slant angle of $25^\circ$, and rounded edges at the front face with radius of 50 mm. The dimensions of the test model measure $522 \times 194.5 \times 144$ mm. The reference, free-stream velocity was set to 12 m/s, leading to a Reynolds number based on the model height, $H = 144$ mm, of $Re = 1.15 \times 10^5$. The flow is seeded with neutrally buoyant helium-filled soap bubbles, with
3.4. EXPERIMENTAL SET-UP AND TEST CASES

Figure 3.3: (left) Time-averaged velocity distribution for the flow past the Ahmed body. Iso-contours of the $Q$-criterion ($Q = 4000 \, \text{s}^{-2}$) coloured by the stream-wise vorticity, $\omega_x$, highlight the two C-pillar vortices. (center) Normalized stream-wise velocity field, $u_3/U_\infty$, interpolated onto the cross-flow plane $0.5D$ downstream of the rear wall, where $D$ is the height of the Ahmed body. (right) Normalized cross-flow vorticity field, $\omega_3 D/U_\infty$, at the same cross-section position. Reference fields calculated with $N \sim 6.2 \times 10^6$ tracks.

A seeding concentration estimated as 0.3 particles per cm, yielding 0.03 ppp. Data is collected using the LaVision MiniShaker S probe, consisting of four CMOS sensor cameras installed into a single body, with a tomographic aperture of $4.3^\circ$. Lagrangian particle tracking via the Shake-The-Box algorithm is conducted with illumination provided by a Quantronix Darwin Duo Nd:YLF laser located at the center of the probe. Time-resolved data is obtained at 700 Hz, in a measurement volume of up to 130 L, covering from the front of the model to approximately $2.2D$ downstream of the rear face. More detailed information concerning the conducted measurements are presented in (Sciacchitano and Giaquinta, 2019).

The second data set consists of a swirling jet flow. Experiments were performed in the Jet Tomographic Facility at the Aerodynamics Laboratories of TU Delft. The fluid domain consists of an octagonal water tank with diameter of 600 mm, and a
3.4. EXPERIMENTAL SET-UP AND TEST CASES

Figure 3.4: (left) Time-averaged velocity distribution of the swirling-jet flow. Jet-cores are highlighted by the iso-contours of the absolute-velocity field at \( u_i/|u_i| = 0.7 \). (center) Normalized stream-wise velocity field, \( u_i/U_\infty \), interpolated onto to the cross-flow plane \( 1D \) downstream of the nozzle-exits; and (right) Normalized cross-flow vorticity field, \( \omega_1 D/U_\infty \). Reference fields calculated with \( N \sim 1 \times 10^6 \) tracks.

height of 800 mm. Time-resolved data was collected in this volume of fluid at 100 Hz, with a seeding density equal to approximately 300 particles per cm\(^3\). The swirling jets emanate from nozzles of diameter 19.7 mm, and have an average exit velocity of 0.05 m/s, leading to a Reynolds number of \( Re = 1000 \). Results investigated in the present work are based on a flow with swirl number of \( S = 0.2 \). Lagrangian Particle Tracking via the Shake-The-Box algorithm is performed in LaVision DaVis 10. Other characteristics of the experimental measurement system are further discussed by Ianiro et al. (2018). The swirling jet flow was chosen as a benchmark for the CSC approach primarily because of its simple flow topologies and stationary behaviour, containing a long-lasting four-lobed-jet core.
3.5 Results and discussion

In the following, the performance of the CSC scheme is assessed on three-dimensional data based on comparison with the baseline vorticity field, \( \omega \). Tracks coloured using the CSC are interpolated onto a two-dimensional Cartesian grid, and the performance of the approach is calculated on a cell-by-cell fashion. The reference flow fields are shown in Figs. 3.3 and 3.4 for comparison. For each test case, unique set of parameters \( \phi = \phi(\Lambda/D, L_s/D) \) was indirectly selected by filtering from the data sets \( N \) particles with temporal track lengths of \( L_t \). In this study, \( N \) ranged from 50 to 30000 tracks for both benchmarks, while \( L_t \) ranged from 4 to 0.9\( L_{t,max} \) time steps, where \( L_{t,max} \) represents the maximum temporal track length present in the respective data set. For the flow past the Ahmed body, these ranges of \( L_t \) and \( N \) are equivalent to inter-particle distance values of \( \Lambda/D \in [0.02, 0.4] \) and normalized track lengths of \( L_s/D \in [0.1, 1.6] \). For the swirling jet flow, the parameter space consists of \( \Lambda/D \in [0.015, 0.18] \) and \( L_s/D \in [0.02, 0.18] \). It is also important to emphasize that particle tracks in this study were randomly selected, which does not guarantee homogeneous seeding distribution, leading to broaden uncertainty in the estimation of the mean \( \Lambda/D \) value. This issue has been addressed by computing high number of test cases for both benchmark flows.

Solutions of the CSC approach for the two benchmark flows were computed and resultant CSC-coloured tracks are illustrated in Fig. 3.5. Figure 3.5a and 3.5b show perspective views of \( N = 10000 \) colour-coded tracks of the Ahmed body and swirling jet flows, respectively. In each test case, tracks are represented in their entire extension. In these plots, solutions of the CSC approach are re-scaled to the interval \( CSC \in [-1, 1] \), in which -1 is represented in dark blue and 1, in dark red, in the selected colour map. For the flow past the Ahmed body, for which results are shown
in Fig. 3.5a, there are clear evidences of successful detection of the left C-pillar vortex at the center of the tracers’ trajectories. The core of the vortical structure is indicated by the tracks colour-coded in red. The C-pillar vortex is surrounded by green colour-coded tracks, representing the outer shear-layers of the coherent structure.

Lastly, tracks laying inside the vortical material surface are surrounded by blue-coloured tracks, representing the transition to the potential-like flow. The test case shown in Fig. 3.5a contains instantaneous 10000 tracer particles with temporal track lengths of \( L_t = 5 \) time steps (equivalent to \( \Lambda/D \approx 0.06 \) and \( L_s/D \approx 0.8 \)), and only the left-hand side portion of the flow-field is represented, for clarity. In Fig. 3.5b, results for the swirling jet flow are shown for a test case with \( N = 10000 \) particles of temporal track length of \( L_t = 30 \) time steps (equivalent to \( \Lambda/D \approx 0.05 \) and \( L_s/D \approx 0.14 \)). As for the flow past the Ahmed body, the coherent features of the swirling jet flow are also easily identifiable in the CSC-coloured field. The four-lobed jet stream at the center of the fluid domain is colour-coded in red. A narrow transition region, representing the material surface in between the four lobes and the surrounding potential flow, is highlighted in green. The coherent jet stream is surrounded by tracks coloured in blue, identifying the quiescent flow. Moreover, due to the considerably high particle concentration of the selected test case, the rain-drop-like shape of the convex lobes becomes evident in the CSC-coloured tracks.

### 3.5.1 Effects of the normalized parameters

CSC-coloured tracks of the flow past the Ahmed body and the swirling jet flow are shown in Figs. 3.6 and 3.7, respectively. Each figure illustrates the projection onto the cross-flow plane of the instantaneous CSC values, with each particle being represented
by a single point at the instant $t = (L_t - 1) \Delta t/2$. The top row shows the streamwise vorticity component obtained with the $\Lambda/D$ value of the respective column, for comparison. The first benchmark investigated in this work is the flow past the Ahmed body. The two longitudinal coherent eddies are the dominant structures of the flow and their cores travel further downstream symmetrically relatively to the $xz$-plane. For that reason, only the left-hand side portion of the fluid volume is analysed in this study, similarly to what is illustrated in the same figure. The two C-pillar vortexes are also relatively stationary (Sciacchitano and Giaquinta, 2019), being this the reasons for selecting this flow as the first benchmark of the present study. This aspect is crucial for the analysis of this flow field since all tracks in the time interval are analysed simultaneously (i.e., as if the flow was stationary).

Here we examine two-parameters cases, for $L_s/D \approx [0.4, 1.0, 1.5]$ and $\Lambda/D \approx [0.04, 0.09, 0.13]$. Figure 3.6 includes examples of baseline vorticity fields and CSC-coloured tracks obtained in the analysis. Columns, left to right, show results for...
decreasing $\Lambda/D$ values (increasing particle concentration). Rows, top to bottom, show results for increasing $L_s/D$ values. In Figs. 3.6a to 3.6c, vorticity fields obtained with $\Lambda/D \approx 0.04, 0.09, 0.13$ are shown for comparison (note that the reference velocity and vorticity fields are shown in Fig. 3.4).

For increasing $L_s/D$ values, barely any difference is noticed in the colouring pattern of the tracer particles. This result is most likely associated with the fact that the relatively parallel streamlines do not allow for considerable increments of the kinematic dissimilarity as the flow moves further downstream. Based on the mathematical description of the CSC algorithm, it is evident that the method benefits from strong divergence in the tracers’ trajectories, which is not sufficiently dominant in the data from flow past the Ahmed body. Another interesting remark is the fact that vortex detection is observed to work reasonably well independently of seeding gaps: the sparse seeding region centred at $(x_2/D, x_3/D) \approx (1, -0.5)$ did not prevent the algorithm from correctly distinguishing between particles (at the vicinity of the vortex’s boundary) that belong to the vortex core from those at the potential-like flow. Non-uniform particle seeding is a common weakness among classical clustering techniques that rely on absolute particle positions (Bezdek et al., 1984) since tracers from distant regions of the flow would be mistakenly clustered together by the adjacency matrix due to the absence of tracks in between them.

In the second benchmark, the swirling jet flow, the flow features a four-sectors-jet divided by the shear-layers resulting from the interactions along the vane walls. The four jet cores rapidly develop a quasi-homogeneous inner momentum distribution after the exit, featuring separation of their respective cores for increasing azimuthal distance due to the swirl number. In all test cases, the CSC approach was able
Figure 3.6: CSC-coloured tracks for the flow past the Ahmed body projected onto the cross-flow plane for $x/D \in [0.5, 1.0]$. Rows, top to bottom, show results for $L_s/D \approx 0.4, 1.0$ and 1.6, respectively. Columns, left to right, show results for $\Lambda/D \approx 0.04, 0.09$ and 0.13, respectively. CSC values range in $[-1, 1]$, where -1 corresponds to blue and 1, to red, in the selected colour map. Results indicated as cases (d), (e), ..., (k) in Figs. 3.8a and 3.9a. Last row includes only the two right-most columns because selected $L_s/D$ value would not allow achieving target density magnitude for the third column to be plot. Figures (a) to (c) show vorticity fields obtained with the corresponding $\Lambda/D$ values.
Figure 3.7: CSC-coloured tracks for the swirling-jet flow projected onto the cross-flow plane for $x/D \in [-1.0, -0.5]$. Rows, top to bottom, show results for constant $L_s/D$ values of $\approx 0.015$, 0.08 and 0.015, respectively. CSC values range in $[-1, 1]$, where -1 corresponds to blue and 1, to red, in the selected colour map. Columns, left to right, show results for $\Lambda/D$ values of $\approx 0.09$, 0.05 and 0.02, respectively. Results obtained in cases (d), (e), ..., (l) are indicated in Figs. 3.8b and 3.9b. Figures (a) to (c) show vorticity fields obtained with the corresponding $\Lambda/D$ values.
to detect the existence of the four-lobed jet with some degree of accuracy, as it is shown in Fig. 3.7. Better feature-extraction performance occurs for decreasing inter-particle distance values, while the benefits of longer normalized track lengths are hardly noticeable, agreeing with the findings of Schlueter-Kuck and Dabiri (2017b). It is also important to highlight the fact that the CSC algorithm assigns arbitrarily large or small values to the coherent structures. For this reason, colour inversions e.g., between panels 3.7k and 3.7l, occur. Alternatively, since the magnitude of the CSC values have no physical meaning, one could artificially change the resultant colour-mapping by computing $1 - \text{CSC}$ so that LCSs are always coloured with the same side of the colour spectrum.

Results in the left-most column of Fig. 3.7 represent flows with sparse particle seeding, resulting in $\Lambda/D \approx 0.09$. At this condition, the spatial particle density is sufficient to observe the existence of the vortices, which can be distinguished by the region of red-coloured particles, but the outer shear-layers of the jets are still hardly noticeable. At $\Lambda/D \approx 0.05$, at the center column, higher particle concentration values start to allow for the detection of both the existence of the four lobes and the stagnation region in between them. For $\Lambda/D \approx 0.02$, at the right-most column, boundaries of the jet core start to become well defined and the characteristic four-lobed jet is easily identifiable in the Lagrangian data. Moreover, Figs. 3.7d and 3.7a represent a case in which the CSC method outperformed the baseline Eulerian diagnostic field. While the sparse CSC-coloured tracks shown in Fig. 3.7d already exhibit evidences of the existence of four lobes, with a small blue-coloured region separating them, the equivalent vorticity field, shown in Fig. 3.7a, presents only a single macro structure that surrounds the four cores. This effect is directly associated
with the fact that the sparsely interpolated vorticity field filters small discontinuities out from the interpolated velocity field, i.e., it “ignores” the three or four tracks with different velocity magnitude at the center of the jet stream. Such sequelae is not observed in a track colouring approach.

Effects of the track lengths are somewhat more ambiguous in comparison to those of the inter-particle distance. Comparing the results from $L_s/D \approx 0.015$ to $L_s/D \approx 0.15$ (top to bottom rows of Fig. 3.7), no noticeable improvement in the identification of the Lagrangian coherent structures can be appreciated, even if the tracks length increases by one order of magnitude. Hence, it is concluded that the CSC results are mainly dependent on the inter-particle density, and therefore on the tracers concentration, whereas they are rather insensitive to the track length. Overall, for the CSC approach, it is more valuable to have more tracks of shorter length, rather than less longer tracks. This result belongs to the maxima that, as for any other clustering technique, more data points act as to converge the solution toward the “ground truth” (Von Luxburg et al., 2008). Results should also improve based upon the quality of the data, which, in the case of the CSC algorithm, is not directly dictated by the amount of Lagrangian information tracks carry (i.e., tracks lengths), but by how divergence tracks are. In that sense, the CSC approach depends on the number of tracer particles and on the flow characteristics, but not on the track lengths, as proposed in the second section of this work.

3.5.2 Correlation between the CSC and Eulerian diagnostics

The accuracy in detecting Lagrangian structures using the CSC approach was evaluated by computing the cross-correlation coefficient between CSC-field mapped onto
3.5. RESULTS AND DISCUSSION

(a) Results for flow past the Ahmed-body. (b) Results for swirling-jet flow.

Figure 3.8: Cross-correlation function versus normalized inter-particle distance for (a) the flow past the Ahmed body; and the (b) swirling jet flow. Reference cases for the respective benchmarks are highlighted by the dashed lines and refer to results shown in Figs. 3.6 and 3.7. Thick dashed lines indicate the trend of the data points.

A Cartesian grid and the stream-wise vorticity component, $\omega_1$, $R_{CSC, \omega_1}$. Both the velocity and CSC-fields were interpolated into a Cartesian two-dimensional grid, projected onto the cross-flow plane. Flow parameters of each grid cell were computed using the average of the $(u, CSC)$-values of all particles that lay within it. The sparse velocity and CSC values are mapped onto a Cartesian bin by averaging their values within cubic bins of 0.02 cm size, with 75% overlap, following the approach proposed by Agüera et al. (2016). The binning is conducted for varying values of the tracers concentration (viz. $\Lambda$) and of the tracks length ($L_s$) to assess the effect of these two parameters onto the quality of the CSC results.
3.5. RESULTS AND DISCUSSION

Figure 3.9: Cross-correlation function versus normalized inter-particle distance for (a) the flow past the Ahmed body; and the (b) swirling jet flow. Reference cases for the respective benchmarks are highlighted by the dashed lines and refer to results shown in Figs. 3.6 and 3.7. Thick dashed lines indicate the trend of the data points.

Figure 3.8 shows results of the CSC algorithm using increasing inter-particle distance values. For both flows, increasing inter-particle distance values leads to loss in the accuracy of the method, as expected. Schluter-Kuck and Dabiri (2017a) obtained similar results, emphasising also that the method is dependent on prior knowledge of the size of the vortical structures in order to secure appropriate sparsification levels. In their work, Schluter-Kuck and Dabiri concluded that numbers of particles in the order of hundreds, equivalent to $\mathcal{O}(\Lambda/D) > 10^{-1}$ in the investigated flows, would be insufficient to accurately detect LCSs, agreeing with what is shown in Fig. 3.8.

The trend of the results, on the other hand, is not monotonic. For the swirling
jet flow, for example, there is a slight decrease in the overall cross-correlation values for low values of $\Lambda / D$. This particular characteristic is most likely due to the fact that the four-lobed convex jet stream goes from an ellipsoidal to a “raindrop-like” shape for $\Lambda / D < 0.05$ when computed by the CSC (compare Figs. 3.7e and 3.7f, for example). On the other hand, the vorticity or absolute-velocity fields tend to exhibit predominantly rounded-like jet-cores. Another remark observed in the same plots is the fact that the correlation function converges to a single value for decreasing values of $\Lambda / D$ independently of the track length. This characteristic allows for the technique to be considered spatially-convergent, part of the set of basic requirements that have proven necessary for self-consistent LCS results (Haller, 2015).

Contrarily to what was observed for the normalized inter-particle distance, the cross-correlation function does not show any particular trend with varying normalized track lengths, as it is shown again in Fig. 3.9. Short track lengths will naturally lead to a less accurate estimation of the kinematic dissimilarities between track pairs, but the reciprocal is not necessarily true for increasing track lengths. The comparison of cases 3.7e and 3.7k proves that the increment of $L_s / D$ value by 500% did not result into an appreciable improvement of the estimated cross-correlation coefficient. A similar conclusion can be drawn for the flow past the Ahmed body, for which results are shown in Fig. 3.9a. Based on the results presented so far, it is concluded that the results of the CSC algorithm are independent of the tracks length, as long as the latter exceeds a certain minimum. Based on the current results, such minimum is estimated to about 15% of the flow characteristic length.

As a final remark, it is also important to highlight one of the biggest drawbacks of the current description of the CSC: the method works in a binary fashion, by
detecting that a single ensemble of tracer particles behaves differently from all the rest, regardless of the real number of vortical structures shed in the flow. This characteristic can be illustrated by analysing the CSC-coloured tracks of the swirling jet flow: in all test cases, its characteristic four-lobed jet is colour-coded in a single colour (e.g., red), while the remaining particles in the quiescent flow are colour-coded in blue, for example. If all flow features were individually detected by the CSC approach, as hoped-for, one would expect to see the four lobes of the swirling jet coloured with four dissimilar CSC-values. In that sense, the CSC approach distinguishes rotational zones from the potential-like ones, but it is not meant for individual features detection, per se.

This characteristic of the CSC approach is the key reason behind the fact that strong correlation is observed between the CSC-coloured tracks and the absolute-velocity field (up to 95% in the tests with the present benchmarks), but small correlation is obtained between the Lagrangian approach and the $Q$ or $\Lambda_2$ criteria (less than 20% with the present benchmarks), for example. In the swirling jet flow, the four lobes of the jet travel at a similar mean velocity that is considerably higher than that of the surrounding quiescent mean. The CSC approach detects the high congruence of the relative motion of particles inside the jet cores in comparison to the particles at the quiescent mean, and this result is very similar in nature to the definition of the absolute-velocity field. Both the $Q$ and $\Lambda_2$ criteria are, on the other hand, proper individual features detection mechanisms (Jeong and Hussain, 1995), highlighting individual convex helical structures that rotate about the symmetry axis of the jet cores (see (Dulin et al., 2017)) and, therefore, are not strongly correlated to what the CSC-coloured tracks express.
3.6 Conclusions

In this study, three-dimensional coherent flow structures were detected using sparse track data for the first time using the Coherent-Structure Colouring (CSC) algorithm. Three-dimensional PTV data of the canonical flow past the Ahmed reference body and a swirling jet flow were adopted as benchmarks. The accuracy in the detection of coherent eddies was evaluated using the cross-correlation between the CSC-field with the baseline vorticity field, based on the two-dimensional projection onto the cross-flow plane of both diagnostic fields. It was concluded that higher accuracy in the coherent structures identification is obtained for decreasing inter-particle distance. Conversely, the CSC results were independent of the tracks length, as long as the latter exceeded 15% of the flow characteristic length.
Chapter 4

A robust Voronoi-tessellation-based approach for detection of Lagrangian coherent structures in sparsely-seeded flows

4.1 Abstract

A robust algorithm to detect coherent structures with sparse Lagrangian particle tracking data, using Voronoi tessellation and techniques from spectral graph theory, is proposed and tested. Neighbouring tracer particles are naturally identified through the Voronoi tessellation of the tracers’ distribution. The method examines the neighbouring time of tracer trajectories, defined as the total flow time two Voronoi cells share a common Voronoi edge, by converting this information into a Cartesian distance in the graph representation of the Voronoi diagram. Coherence is assigned to groups of Voronoi cells whose neighbouring time remains high throughout the time interval of analysis. The technique is first tested on the two-dimensional synthetic data of a double-gyre flow, and then with challenging, large-scale three-dimensional Lagrangian particle tracking data behind a bluff body at high Reynolds number. The
novel approach proves to be successful at identifying coherence with realistic and very sparse experimental data. Specifically, it is shown that coherent structures are identifiable for the first time for mean inter-particle distances of the order of the largest length scales in the flow.

4.2 Introduction

Lagrangian coherent structures (LCS) are ubiquitous to all fluid flows, and are often observable through collective paths of buoyant drifters (Olascoaga et al., 2013; Pearson et al., 2019; Shadden et al., 2009) or tracer particles (Allshouse and Peacock, 2015; Rosi et al., 2015). The concept of identifying LCSs directly from Lagrangian particle tracking (LPT) data emerges from the clear fact that tracers inherit their motion directly from the carrier fluid, allowing for the detection of the most influential surfaces in the velocity field (Serra and Haller, 2016). LCS-detection can also be understood as both a tool for qualitative flow visualization and as a basis for reduced-order modelling (Balasuriya et al., 2018; Schlueter-Kuck and Dabiri, 2019). Despite these observations, direct measurements of LCSs in flow fields has remained a challenging task despite countless efforts to extract materially-coherent domains from arbitrary complex flows (Farazmand and Haller, 2012; Haller, 2001, 2002; Hussain, 1983; Shadden et al., 2005). LCS-detection is especially challenging for many sparse data sets, characteristic of realistic problems, in which the mean distance between tracer particles can be on the order of the largest length scales present in the flow. Here, we present and discuss an alternative approach for LCS-detection based on Voronoi tessellation (Rosi et al., 2015) and graph clustering that has potential to detect coherent motion in sparse data (see also (Fukami et al., 2021)).
As discussed by Schlueter-Kuck and Dabiri (2017a), most traditional approaches in determining coherent structures from LPT data are ultimately Eulerian in nature, and generally consider single snapshots in time of typical streamline structures so as to define their respective vortex criteria. The Eulerian framework for Eulerian coherent structure (ECS) detection typically consists of mapping sparse LPT data onto a grid, in a points-in-cell fashion (Harlow, 1957), having the velocity field $u_i, i \in I = \{2, 3\}$, indirectly determined from the statistics of points laying within the boundaries of the grid’s cells. In the Eulerian framework, coherent structures are typically identified using vortex criteria such as the $Q$-, $\Lambda_2$- or $\omega$-fields (Zhan et al., 2019); all of those requiring computations of the velocity gradient tensor, $\partial_j u_i$ (Green et al., 2007). This inherent necessity for computing spatial gradients based on pre-interpolated velocity-fields can produce unreliable diagnostic fields for readily available, moderately-sparse seeding concentrations (Hadjighasem et al., 2017). In contrast to classical Eulerian approaches, various processes that take place in the ocean, atmosphere and other large-scale flows, often result in data that suffers from poor temporal and spatial resolution (Mathur et al., 2019; Schmale III and Ross, 2015). Resulting low signal-to-noise ratios from real-life data motivates the current development of Lagrangian approaches that allow for accurate feature-extraction in sparse data (see review by Hadjighasem et al. (2017)).

To avoid the inherent challenges associated with experimentally-acquired LPT data (e.g., sparsity and spatial clustering of tracks), classical coherence detection mechanisms involve artificially seeding of regularly-spaced Eulerian data. Such an approach produces Eulerian data that is subsequently integrated over time, reconstructing pathlines from tracers (Raben et al., 2014). This method naturally inserts bias
and truncation error into acquired data at high computational cost. Moreover, traditional Lagrangian analyses, in which gradients of the velocity field are also needed, are not readily appropriate for feature-extraction from sparse data sets either. Classical examples include the finite-time Lyapunov exponent, FTLE (Haller, 2000), and the finite-size Lyapunov exponent, FSLE (Boffetta et al., 2001). FTLE-fields have been successfully adopted to locate the most influential material-surfaces in geophysical (Du Toit, 2010), atmospheric (Garaboa-Paz et al., 2015) and oceanic flows (Beron-Vera et al., 2008). Nonetheless, chaos-theory-based approaches for coherence detection such as the FTLE are also based upon the analysis of flow gradients (Martins and Zanotello, 2018), requiring tracer trajectories to be sufficiently close with one another so that necessary linearizations of the local velocity field’s gradients remain accurate (Schlueter-Kuck and Dabiri, 2017a).

Owing to the obstacles of classical Eulerian and Lagrangian approaches for coherent-structure detection, new techniques that do not rely on the computations of flow gradients have been proposed (Peacock and Dabiri, 2010). Among these, one of the most promising techniques for real-life applications is the Coherent-Structure Colouring (CSC) approach, originally tested by Schlueter-Kuck and Dabiri (2017a), and critically assessed by Schlueter-Kuck and Dabiri (2017b), Schlueter-Kuck and Dabiri (2019), and Husic et al. (2019). The CSC approach is based on the spectral graph clustering theory for pattern identification of tracer motion, as first introduced by Bezdek et al. (1984). The algorithm compares the kinematic responses of individual tracer particles and colour-codes their tracks based on their kinematic dissimilarities. In practice, particle trajectories that present relatively small spatial deviation from one another are coloured with similar CSC values.
The CSC algorithm colour-codes LPT data based on the kinematic dissimilarity between all pairs of particles. Such definition of coherence is not based upon fixing a global coordinate system, making the approach Lagrangian invariant (Haller and Yuan, 2000). Moreover, the number of clusters (coherent structures) present in the flow does not need to be determined a priori, avoiding the need of biased inputs from the user, such as needed in classical approaches (Saqib et al., 2017; Zhao et al., 2020). Clustering techniques are also not necessarily constrained by a minimum number of data points, meaning that this category of LCS-detection approaches can theoretically operate with indefinitely sparse data sets (Von Luxburg et al., 2008). Nonetheless, the CSC technique presents critical limitations when adopted to data sets from real-life applications: for a flow with largest length scales on the order of $D$, the CSC technique was found to have limited value for mean inter-particle distances of $\sim 15\% D$ or greater (Martins et al., 2021) whereas, for typical experimental data, seeding can be considerably sparser (Lumpkin and Johnson, 2013). Another drawback of the CSC approach is the fact that data must be comprised of tracks of equal length, which is infeasible in most experimental settings. Moreover, since particles in the CSC approach are viewed as scattered, discrete points in space, the vast majority of the physical portion of the fluid domain is left empty, making diagnostics difficult to interpret for sparse seeding-distributions.

In the present study, rather than viewing a fluid parcel as a means for discretization, we employ the simple concept that a seeding particle, with its associated position and momentum, is representative of a volume of the fluid flow (Espanol and Serrano, 2009; Krueger et al., 2019; Padberg-Gehle and Schneide, 2017; Rosi et al., 2015). One of the most natural ways of assigning volume to a tracer is through the
Voronoi tessellation (Neeteson and Rival, 2015). The proposed technique employs Voronoi tessellation to detect neighbouring particles in the flow, using the period of time a pair of Voronoi cells persist as neighbouring cells, referred to here as their \textit{neighbouring time}, as a metric for coherence. The algorithm then employs spectral graph clustering theory to compare the neighbouring times of an arbitrary number of Voronoi cells, colour-coding them based on the period of time they share a common Voronoi edge (as shown in Fig. 4.1). In practice, Voronoi cells that present relatively long neighbouring times are colour-coded with similar colours. Since the neighbouring time is independent of reference frame, the proposed technique is naturally Lagrangian-invariant (Haller, 2015). Moreover, the approach is based on spectral clustering, which does not necessarily constrain the technique to a minimum number of data points, making it a strong candidate for the use with very sparse LPT data.

It therefore becomes our first objective to prove and quantify the ability of the current technique to detect coherent structures in LPT data. Towards this goal, the proposed approach is described and tested with synthetic data on the two-dimensional, mixing double-gyre flow, and then on challenging three-dimensional LPT data behind a bluff body at high Reynolds numbers. Coloured Voronoi cells of the current technique are then compared to CSC-coloured tracks, FTLE-fields, and to diagnostics obtained with the baseline vorticity fields.

The outline of this chapter is as follows. In section 4.3, we review the concepts for defining coherence using topological properties of the Voronoi diagrams. This analysis is followed in section 4.3.1 by the description of the adopted key normalized parameters for the evaluation of the current method. In section 4.4, we describe the adopted benchmark flows. In section 4.5, analysis of the technique is performed with
4.3 Methods

The following section is organized as follows. In section 4.3.1, the mathematical description of the proposed method and adopted notation are presented. Subsequently in section 4.3.2, we explain the normalized parameter space for the analysis of the current technique.
4.3. METHODS

4.3.1 Mathematical description

This section presents the mathematical description of the proposed technique. Key steps of LCS detection within the technique’s framework are also schematically represented in Fig. 4.2. Consider the instantaneous seeding of \( N \) particles, \( p_i \ (i \in I_N = \{1, 2, \cdots, N\}) \), in \( m \)-dimensional Euclidean space, \( m \in I = \{2, 3\} \), with location vectors \( \mathbf{x}_i \). The Voronoi diagram of \( p_i \), \( V(p_i) \), is such that for every point \( \mathbf{x} = x_i \) in space, \( x_i \in \mathbb{R}^m \), a convex polygon, referred to as a Voronoi cell, \( V_R(p_i) \), can be defined as a region in which \( \mathbf{x} \) is closer member of \( p_i \)

\[
V_R(p_i) = \{\mathbf{x} \text{ s.t. } \|\mathbf{x} - \mathbf{x}_i\| \leq \|\mathbf{x} - \mathbf{x}_j\| \text{ for } j \neq i, j \in I_N\}, \tag{4.1}
\]

than to any other particle \( p_j \) \((p_i \neq p_j \text{ if } i \neq j)\) with location vector \( \mathbf{x}_j \). In Eq. 4.1, \( \mathbf{x} \) is any point within the fluid domain, and \( V_R(p_i) \) encloses each particle into an individual Voronoi cell (see highlighted polygon in Fig. 4.2). The boundaries of such a polygon are referred to as Voronoi edges, \( r_{ij} \) (see thick black lines in top-left panel in Fig. 4.2). The collection of all Voronoi cells at the time instant \( t \), \( V(p_i) \), given by

\[
V(p_i) = \{V_R(p_1), \cdots, V_R(p_N)\}, \tag{4.2}
\]

is the Voronoi diagram of the seeding distribution, \( p_i \) (Ferrero, 2011). Based on \( p_i \), a network called a Delaunay triangulation, \( DT(p_i) \), (see dashed gray lines in top-left panel in Fig. 4.2) can be defined, where the vertices of the network correspond to particle positions, \( \mathbf{x}_i \), and a triangle (i.e., a set of three edges) between three adjacent nodes is created if no other particles lay within the circumcircle area defined by the three nodes (see green circle in Fig. 4.2). It is evident from the definition of \( DT(p_i) \)
that Voronoi cells only share a common Voronoi edge when their associated particles are connected by a Delaunay edge (Chew, 1989). It is also axiomatic from the given definitions of $DT(p_i)$ and $V(p_i)$ that edges of the Voronoi diagram are normal to the Delaunay triangle edges.

Hereafter, we define neighbouring tracers, $p_i$ and $p_j$, all pairs of tracer particles that, at the $n$-th time-step, describe the two vertices of the $(i,j)$-th Delaunay edge (see particles highlighted in green in Fig. 4.2). The total neighbouring time, $n_{T_{ij}}$, is subsequently defined as the total number of time-steps, $n_{T_{ij}} \in [0,n]$, for which the $i$-th and $j$-th tracers persisted as neighbouring tracers. Based upon such definitions, the problem of coherent-structure detection is naturally reduced to that of finding groups of Voronoi cells that persist for higher numbers of time-steps as neighbouring cells, since the cells of two particles in a Delaunay edge are neighbouring cells. The present technique solves this last problem by adopting spectral graph theory (Spielman, 2007). Consider a graph $G(V)$ defined such that its $i$-th node represents the $i$-th Voronoi cell, whereas its $(i,j)$-th edge connects the $i$-th cell to its $j$-th neighbouring Voronoi cell. The weight (or length) $A_{ij}$ of the $(i,j)$-th edge is defined as

$$A_{ij} = \left(\frac{1}{2}\right)^{n_{T_{ij}}},$$

where $A$ is referred to as the adjacency matrix of the graph $G(V)$. From the definition in Eq. 4.3, it is clear that the length of the $(i,j)$-edge is smaller for greater neighbouring times. Hence, long-lasting neighbouring tracers will be closer in the graph representation of the Voronoi diagram, $G(V)$, and are naturally clustered together. It is also evident that such graphs share the same topology of $DT(p_i)$ (see Fig. 4.2). Nonetheless, while edges of $DT(p_i)$ yield a physical distance between pairs of tracers,
Figure 4.2: Schematics of coherent-structure detection based on Voronoi tessellation. Voronoi and Delaunay diagrams of the particle distribution $p_i$, $V(p_i)$, $DT(p_i)$, are represented in the graph space, $G(V)$, such that lengths of edges of $DT(p_i)$, $r_{ij}$, are converted into distances in the higher-dimensional eigenspace $G(V)$: $r_{ij} \rightarrow A_{ij} \rightarrow \chi(\mathcal{L}(G))$. A Voronoi cell, $V_R(p_i)$, is highlighted in blue in the top-left panel. The top-left panel also highlights in green the triangulation of three particles intersecting a circumcircle with no particles within its interior.

edges of $G(V)$ represent what we define as the kinematic distance between the same tracer particles. Moreover, since the neighbouring time is a frame-invariant property, coherent structures obtained with the proposed technique are also Lagrangian invariant (Haller, 2000).

Central to the spectral clustering technique are also the definitions of the graph’s degree matrix, $D$, given by

$$D_{ij} = \delta_{ij} \sum_{k=1}^{N} A_{ik}, \quad (4.4)$$

and the graph’s Laplacian, $\mathcal{L}$, defined as
\[ \mathcal{L}_{ij} = D_{ij}^{-1/2}(D_{ij} - A_{ij})D_{ij}^{-1/2}, \quad (4.5) \]

where \( \delta_{ij} \) represents Dirac’s delta function. In Eq. 4.5, \( \mathcal{L} \) is pre-normalized by \( D \).

The corresponding eigenvalue problem that maximizes the differences between the Voronoi cell’s neighbouring time is

\[ \mathcal{L}_{ij}X_j = \xi X_j, \quad (4.6) \]

such that \( \chi_j \equiv X_j(\xi_{\text{max}}), j \in I_N = \{1, \cdots, N\} \), is the eigenvector with associated maximum eigenvalue, \( \xi_{\text{max}} \) (Schlueter-Kuck and Dabiri, 2017b). Voronoi cells, \( V_R(p_j) \), assigned with values \( \chi_j \), present the largest dissimilarities in the graph representation space, hence representing the most influential coherent structures in the Voronoi-representation space. This coherence criterion is such that Voronoi cells that share a common Voronoi edge for longer flow times present a shorter distance in the higher-dimensional eigenspace (lower eigengaps), hence presenting coherent motion. Coherence is then visualized by mapping the solution \( \chi \in [\chi_{\text{min}}, \chi_{\text{max}}] \) onto a colour map, allowing for coherent-structure identification in groups of Voronoi cells colour-coded with similar \( \chi \)-values.

### 4.3.2 Normalized parameters

Considering a volume of fluid that scales with the characteristic length of the flow, \( D \), and a homogeneous seeding distribution, accurate coherent-structure identification is assumed to be dependent on the number of tracers, \( N \), present in the flow. Consider a flow with largest length scales \( D \). Defining the characteristic period of the largest scales as \( T \), coherence shall be visualized in normalized time, \( t/T \). Moreover, the
amount of Lagrangian information in the flow relative to $D$ is assumed to be proportional to the number $N$ of tracer particles tracked per characteristic length $D$, as denoted by $C = N/D^3$. Subsequent to the definition of $C$, here we define the mean inter-particle distance as

$$\Lambda \equiv \left( \frac{3}{4\pi C} \right)^{1/3},$$

and the correlation between two diagnostic fields $\Psi_1$ and $\Psi_2$ within the proposed framework, $R_{\Psi_1,\Psi_2}$, can be approximated as a function of a single variable $\Lambda$:

$$R_{\Psi_1,\Psi_2} = R_{\Psi_1,\Psi_2}(\Lambda).$$

Accuracy of LCS detection with the proposed approach is assessed by comparing results obtained within the technique’s framework with results obtained from the baseline Finite-time Lyapunov exponents (FTLE) approach (Haller and Yuan, 2000) and with the state-of-art Coherent-Structure Colouring (CSC) algorithm (Schlueter-Kuck and Dabiri, 2017a). Both of these approaches will be summarized in the following paragraphs.

The FTLE measures the maximum linearised growth rate between initially adjacent particles. The FTLE approach is founded on the definition of the Cauchy-green tensor, $C$, given by

$$C(x_0) = \left( \nabla F_{t_0}^t \right)^\top \cdot \left( \nabla F_{t_0}^t \right),$$

where $F_{t_0}^t$ represents the flow map, which operates over a fluid parcel at $x_0 \equiv x(t_0)$,
and tracks the parcel’s position from an initial to a final state, $\mathbf{F}_{t_0}^{t} : \mathbf{x}_0 \to \mathbf{x}(\mathbf{x}_0)$. Identification of coherent structures with the FTLE technique relies upon the definition of regions of highest eigenvalue $\lambda_{\text{max}}$:

$$FTLE_{t_0}^t = \frac{1}{|t - t_0|} \ln \left( \sqrt[\lambda_{\text{max}}(C)} \right). \quad (4.10)$$

In the FTLE framework, LCSs are identified as barriers to fluid motion. Extrema values of $FTLE_{t_0}^t$ represent the locally more influential material surfaces. If integrated forward in time, extreme positive values represent the locally most repelling material surfaces, whereas opposite in signal extrema represent the most attracting ridges. Both material surfaces exchange definitions if the FTLE is integrated backward in time.

Moreover, $\chi$-coloured Voronoi cells will also be compared to the CSC-coloured tracks from the CSC approach. Similar to the currently-proposed technique, in the CSC framework the dissimilarities between a pair of tracer trajectories is represented by the adjacency matrix of the associated graph-clustering problem, $\mathcal{A}$, defined as

$$\mathcal{A}_{ij} = \frac{1}{\overline{r}_{ij} L_t^{1/2}} \left[ \sum_{n=0}^{L_t-1} (\overline{r}_{ij} - r_{ij}(t_n))^2 \right]^{1/2},$$

where $\overline{r}_{ij}$ represents the mean distance between the $i$-th and $j$-th particles, and $L_t$ is the temporal track-length. The spectral clustering process based on the CSC approach follows from Eq. 4.6 such that $CSC \equiv \mathcal{X}(\xi_{\text{max}})$. Results obtained with both the FTLE and CSC techniques for varying $\Lambda$-values will be used for comparison.
4.4 Description of benchmarks

4.4.1 Algorithm verification via case study: double-gyre flow

We specify an unsteady velocity field \( u_i, i \in I = \{1, 2\} \), the double-gyre flow, that has had extensive use as a test-bed for LCS detection (Balasuriya, 2016; Balasuriya et al., 2018; Pratt et al., 2015; Senatore and Ross, 2011; Shadden et al., 2005; Williams et al., 2015):

\[
\begin{align*}
    u_1 &= -\pi A \sin(\pi h(x_1, t)) \cos(\pi x_2), \\
    u_2 &= \pi A \cos(\pi h(x_1, t)) \cos(\pi x_2) \frac{\partial h}{\partial x_1}(x_1, t),
\end{align*}
\]

where

\[
h(x_1, t) := \epsilon \sin(\Omega t) x_1^2 + (1 - 2\epsilon \sin(\Omega t)) x_1.
\]

For a domain \( \Gamma = [0, 2] \times [0, 1], x_i \in \Gamma \), it is well-known that the topology of the flow comprises hyperbolic trajectories near \( x_i = (1, 0) \) and \( (1, 1) \), which move position for \( \epsilon > 0 \) (Williams et al., 2015). In the present study, \( A = 1, \epsilon = 0.25 \) and \( \Omega = 2\pi/T = 2\pi/10 \). The characteristic length \( D \) and velocity \( U_\infty \) are such that \( U_\infty/D = T \). Figure 4.3 shows the normalized vorticity field at different characteristic times \( T \) for reference. The reference flow is simulated with spatial resolution of \( \Delta x_i/D = 0.01 \), and was evolved with a constant time-step of \( \Delta t/T = 0.02 \).

4.4.2 Case study: High-reynolds-number bluff-body

The second benchmark conducted in this study consists of LPT data behind a bluff body collected in the 24 m × 9.1 m × 9.1 m test section of a large low-speed wind
Figure 4.3: Normalized vorticity fields for the reference double-gyre flow. The plots exhibit contours of normalized vorticity at different time-steps $t$ obtained with a grid of spatial resolution of $\Delta x_i/D = 0.01$. tunnel at the National Research Council in Ottawa, Canada (Hou et al., 2021). An overview of the raw-tracks data and resultant, interpolated vorticity field are provided in Fig. 4.4. Tracks behind the bluff body are measured using a single-camera set-up (Photron mini-WX 100, AF Micro-Nikkor 60 mm f/2.8, with the aperture set to 11) in a measurement volume of $4D \times 1.5D \times 1.5D$, where $D$ is the characteristic length scale of the flow ($D \approx 1$ m). Large air-filled soap bubbles generated with two commercial bubble generators (Antari B200) are adopted as tracer particles, yielding bubbles-production rates of $\sim 80$ bubbles/s. Soap bubbles of average diameter $\sim 17.5$ mm are illuminated by an array of four pulsed high-power LEDs (LED-Flashlight 300, LaVision GmbH). Characteristic wake-flow behind the bluff body at free-stream velocity of $U_\infty \approx 8$ m/s, yielding $Re \sim 6 \times 10^5$ based on the model’s height, is assessed at a constant sampling frequency of $f_s = 150$ Hz. More details concerning
4.4. DESCRIPTION OF BENCHMARKS

Figure 4.4: Raw tracks shown in top left. Reference normalized cross-flow vorticity field obtained from the point-in-cell interpolation of the tracers trajectories is shown in bottom right. Raw reference fields (without smoothing or interpolation) are plotted in specifics planes normal to the streamwise direction at $x_1/D = [-0.5, 0.5, 1.5, 2.2]$. Results obtained with a constant bin size of $\Delta x_i/D = 0.05$. The direction of the reference far-field velocity, $U_\infty$, and the characteristic length, $D$, are also illustrated.

the measurement process and general flow field are presented by Hou et al. (2021).

A total of 5718 bubble tracks are extracted from this large-scale measurement, whereas in the present study we filter-out from the raw track data trajectories of temporal length $L_{s,\min} \leq 9$ units. The filtering process yields 1940 tracks of temporal length $L_{s,\min} \geq 10$ units. Figure 4.4 represents raw pathlines and a reference normalized vorticity field obtained from point-in-shell interpolation of the 1940 filtered
4.5. RESULTS AND DISCUSSION

tracer trajectories. We also highlight the fact that no smoothing or interpolation schemes are adopted to the reference vorticity field shown in Figure 4.4.

The analysis of the reference vorticity field indicates that tracks undergo a strong swirling motion due to the C-pillar vortex formed in the wake of the bluff body. Coherence detection from this LPT data is challenging due to a lack of tracks in the lower-left corner of the measurement volume, and due to the lack of tracks in the vortex-core region. Low tracer concentration in the vortex core region is associated with bursting of bubbles due to a large pressure and shear-stress gradients near the centerline of the coherent structure. Strong advection outwards from the vortex’s core also yields inhomogeneity in the tracer distribution. The sparsity and non-homogeneous tracer distribution characteristics of the current data set allow for the assessment of a single C-pillar-like vortex structure in the track data. However, the coherent structure is hardly visible in the vorticity field due to the poor extraction of spatial-gradients achieved with this very sparse data set.

4.5 Results and discussion

4.5.1 Double-gyre Flow

In the following section, the performance of the proposed approach first is assessed on simple two-dimensional data from the double-gyre flow. Voronoi cells coloured with \( \chi \)-values are compared to the FTLE- and CSC-coloured fields in a side-by-side fashion. CSC- and FTLE-fields are re-scaled in all cases to the intervals \([-1, 1]\) and \([0, 1]\), respectively, to allow for direct comparison of results. We extract coherent structures using the proposed technique by simply thresholding \( \chi \) to 70% of its maximum value, and re-scaling the resultant field to the interval of \([-1, 1]\) (see (Senatore and Ross,
4.5. RESULTS AND DISCUSSION

The first analysis consists of plotting results obtained with the three approaches for decreasing mean inter-particle distances value, whereas the normalized flow time is fixed at $t/T = 2$. Subsequently, the same analysis is repeated for a comparison of the current technique to the baseline vorticity field. Lastly, $\chi$-coloured Voronoi cells are interpolated onto a two-dimensional Cartesian grid, $V(p_i) \rightarrow \Psi(x_i), \Psi \in \mathbb{R}^2$, and profiles $\psi \equiv \Psi(x_i = a), \psi \in \mathbb{R}$, for constant $a$, of the resultant two-dimensional field $\Psi$ are compared to profiles of baseline vorticity fields. Here, the objective is to perform a systematic comparison of the different techniques, using them on a test-bed flow in which a ground truth can be reasonably established. Nevertheless, due to the absence of a standardized method of quantifying the quality of coherent-structure detection techniques that are so different in nature, diagnostic tools are compared in this study purely through visual inspection. Even so, analysed techniques must meet a minimum requirement: they must consistently outperform visual inference from randomly advected scalar fields (Hadjighasem et al., 2017; Peacock and Dabiri, 2010), such as those shown in Fig. 4.5.

In the present analysis, the number of tracers, $N$, ranges considerably ($N = 16, 25, 64, 225, 484$ and $625$), corresponding to mean inter-particle distances of $\Lambda \approx 0.31, 0.27, 0.20, 0.13, 0.1$ and $0.09$. In a recent publication, it was concluded that particle concentrations of $N/D^3 \leq \mathcal{O}(10^2)$ ($\Lambda \geq 0.13$) are insufficient for coherent-structure detection in boundary-layer or free-jet flows based on LPT data (Schneiders and Scarano, 2016). Hence, here we consider particle distributions of $\Lambda \geq 0.13$ as very sparse data sets. Tracers are homogeneously-seeded using the following procedure: in every test case, the domain $\Gamma_i, i \in I = \{1, 2\}$, is split into bins of size

2011)).
4.5. RESULTS AND DISCUSSION

Figure 4.5: Arbitrary scalar fields advected on the double-gyre flow evaluated at $t/T = 2$. Plots on top row reveal fields obtained through the evaluation of 64 tracer trajectories, while bottom row shows results obtained with 484 tracer trajectories. Tracer coordinates at end time are indicated by white circles.

Table 4.1: Summary of test cases for the double-gyre flow, where $N$ is the number of tracer particles and $\Lambda$ is the mean inter-particle distance.

<table>
<thead>
<tr>
<th>Case</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>16</td>
<td>25</td>
<td>64</td>
<td>225</td>
<td>484</td>
<td>625</td>
</tr>
<tr>
<td>$\Lambda$</td>
<td>0.31</td>
<td>0.27</td>
<td>0.20</td>
<td>0.13</td>
<td>0.10</td>
<td>0.09</td>
</tr>
</tbody>
</table>

$\Delta x_1 \times \Delta x_2 = \Gamma_1/N^{1/2} \times \Gamma_2/N^{1/2}$, followed by the seeding of a single particle into a random position inside one of the $N$ uniformly-distributed bins. After the homogeneous seeding of synthetic tracers, tracer trajectories are integrated using a second-order central scheme with time-steps of $\Delta t/T = 0.2$ in $t/T \in [0, 10]$. Test cases are summarized in Table 4.1.

Figure 4.6 presents the double-gyre flow for, from left-to-right, $\chi$-coloured Voronoi
cells, FTLE- and CSC-fields. CSC-fields are obtained through the mapping of the instantaneous seeding distribution onto two-dimensional Cartesian grids. The resultant field is then subjected to a robust spline smoothing algorithm developed by Garcia (2010), which adopts a type-2 discrete cosine transform to assign missing data to values that are estimated using the entire dataset. Moreover, we use this blue-green-yellow-red colour-scheme consistently for all diagnostics fields for ease of comparison.

From these first results, it is observed that without requiring specification of the number of gyres, the proposed technique was able to reveal physically-interpretable coherent regions for the sparsest test case, at $\Lambda \approx 0.31$ (see case (a) in Fig. 4.6). Despite the very low particle concentration at $\Lambda \approx 0.31$, the two gyres can be identified in the flow as two regions in the Voronoi diagram coloured in red and blue, corresponding to the two extrema of $\chi$. At this same particle concentration, both the CSC and FTLE techniques were unable to provide physically-interpretable results.

Further decreasing the inter-particle distance to $\Lambda \approx 0.2$, for case (c), the boundary between the two distinguished regions in red and blue in the Voronoi diagram becomes well-defined, whereas no discernible pattern has emerged from the FTLE or CSC approaches. Moreover, further decreasing the mean inter-particle distance to $\Lambda \approx 0.13$, for case (e), a clear pattern of red-coloured Voronoi cells near $x_1 = 1$ becomes visible in the Voronoi diagram. The comparison of such a pattern with ridges of the FTLE-field for the same flow, shown in the center column of Fig. 4.6, reveals that the pattern created by the Voronoi cells in this region of the flow represents an attracting material surface, i.e., an unstable manifold that represents a barrier to fluid motion, classically observed in the FTLE-field. Also for case (e), two regions start to form in the CSC-fields, which are evident from the two clusters of tracers somewhat
Figure 4.6: Selected cases for double-gyre flow at $t = 2T$ with mean inter-particle distance values of $\Lambda \approx 0.31, 0.27, 0.20, 0.13, 0.10, 0.09$ (top-to-bottom). Analyses are visualized by $\chi$-coloured Voronoi cells (left), FTLE-fields (center) and CSC-coloured fields overlaid by CSC-coloured tracks (right). Cases (a) to (f) are listed in Table 4.1.
reproducing the expected topologies of the two gyres. The fact that material coherence can be observed in CSC-fields with the particle concentration of case (e) agrees with observations of Husic et al. (2019), who concluded that the CSC is expected to accurately reconstruct the two gyres of the double-gyre flow with a minimum number of tracks of $N \sim 300$, or $\mathcal{O}(\Lambda) \leq 0.1$.

The current technique presents an advantage when compared to the two classical approaches since in the current technique’s framework coherence can still be assessed at scales $\sim \mathcal{O}(D)$, whereas both the FTLE and CSC techniques are known to breakdown (Green et al., 2007; Husic et al., 2019; Rosi et al., 2015; Schlueter-Kuck and Dabiri, 2017a; Serra and Haller, 2016). Furthermore, results obtained with the current technique, for cases (a-f), are such that physically-distinguishable coherence is always obtained, but coherence is observed as features of different length scales depending upon $\Lambda$. We observe that for $\Lambda \sim \mathcal{O}(D)$ (cases (a-c)), the two gyres are visible in the Voronoi diagram, whereas for $\Lambda \sim \mathcal{O}(l)$ (cases (d-f)) coherence is manifested as a sharp boundary between the two disparate rotational regions. In the latter, $l$ represents the length scale of the attracting surface at the symmetry section of the double-gyre flow. We also observe that for $\Lambda \approx 0.2$ (case c in Fig. 4.6), the green-like region is the manifestation of the separation between the two vortices and corresponds to the dark-red central regions seen in case (f). These conclusions can be re-stated such that at $\mathcal{O}(\Lambda/D) \sim 1$, $\chi$-coloured coherence is manifested as the regions of coherent motion, whereas at $\mathcal{O}(\Lambda/D) \ll 1$, $\chi$-coloured coherence is observed as regions of locally maximum and minimum stability separating the regions of coherent motion.

We repeat the comparative analysis using an Eulerian vortex criterion. Due to the simple topology of the flow, the vorticity field, $\omega = \nabla \times u$, is sufficient for this
4.5. RESULTS AND DISCUSSION

Figure 4.7: Selected cases of double-gyre flow at $t = 2T$ with mean inter-particle distance values of $\Lambda \approx 0.31, 0.13, 0.10, 0.09$ (top to bottom) visualized by $\chi$-coloured Voronoi cells (left) and $\omega$-field (right).

study. The $\omega$-fields are obtained as follows. The domain $\Gamma_i$ is divided into bins of size $\Delta x_i/D = 0.05$. Velocity components of the $(i,j)$-th bin are computed using the average velocity of the tracers within its domain, in a point-in-cell fashion (Garth and Joy, 2010), whereas empty cells are assigned not-a-number (NaN). After mapping the reconstructed Lagrangian velocity to Eulerian space, the curl of the interpolated...
velocity field is computed such that partial derivatives are calculated using a second-
order central difference scheme for interior data points, whereas first-order, single-
sided (forward) difference schemes are adopted for points along the edges (Moukalled
et al., 2016). ω-fields are also subjected to a partial reconstruction process using
a linear interpolation algorithm (Henn et al., 2013) that replaces NaNs in interior
points by the arithmetic mean of the respective field variable in the adjacent cells.
The same colour scheme implemented for other diagnostics is adopted for the ω-field
for ease of comparison.

For case (a) using the Eulerian approach, shown in the top-right panel of Fig. 4.7,
the very sparse Lagrangian data (16 particles) does not allow for the velocity field to
be reconstructed, as expected, whereas the two gyres can be identified as two clusters
of Voronoi cells when using the current approach. For Λ ≈ 0.13, corresponding to
case (d), a noisy vorticity-field can be reconstructed from the Lagrangian particle
tracking data. Concomitantly, the unstable flow-region in between the two gyres is
physically distinguishable in the red-coloured Voronoi cells. Finally, for Λ ≈ 0.09,
the strongest ridge is observed in the Voronoi diagram in the bottom panels of Fig.
4.7. At this particle concentration level, the two counter-rotating gyres are also easily
distinguishable in the vorticity field.

Recalling the definition of Ψ(x_i) as the projection of a coherent-structure diag-
nostic onto a two-dimensional Cartesian grid, and the definition ψ as ψ ≡ Ψ(x_i = a),
for constant a, here we compare the profiles ψ, ψ ∈ ℝ, of the four tested coherent-
structure diagnostics in a grid of constant bin sizes of Δx_i/D = 0.05. The majority
of the Voronoi diagrams in Fig. 4.6 indicate the presence of the two gyres. Out of
those offering more accurate coherent structure boundaries, here we analyse results
4.5. RESULTS AND DISCUSSION

Figure 4.8: Profiles $\psi$ of the Voronoi diagram at the symmetry section of the double-gyre flow, $\psi = \Psi(x_2/D = 0.5)$ for different mean inter-particle distance values.

for cases (a), (c) and (f), corresponding to mean inter-particle distances of $\Lambda \approx 0.31$, 0.20 and 0.09, respectively. Profiles of the Voronoi diagrams at the symmetry section of the double-gyre flow, $\psi = \Psi(x_2/D = 0.5)$, are compared to a reference baseline vorticity-field evolved in the same grid directly from the flow’s stream-function (see Eq. 4.11).

Figure 4.8 shows regions where shape coherent sets may exist. For $\Lambda \approx 0.31$, these regions are revealed by the positive and negative peaks of $\psi(\chi)$ symmetrically spaced from $x_1/D \approx 1$. The same symmetry is revealed for the $\Lambda \approx 0.20$ case, and it is such that $\chi$ encompasses vortex boundaries of the two gyres. In both test cases, strong correlation with the vorticity-field profile is observed. However, the analysis of the profile for $\Lambda \approx 0.09$ reveals that the two peaks of $\psi$ no longer exist in the most densely-seeded case, but rather we observe a negative peak near $x_1/D = 1$, in the middle of the two plateaus, corresponding to large FTLE-values (see case (f) in
4.5. RESULTS AND DISCUSSION

Fig. 4.6), and two plateaus at $x_1/D \sim 0.5$ and $\sim 1.5$, which are indicative of the two vortex cores.

Recalling that the profile $\psi$ is naturally also a representation of the eigenvector $\chi$, extrema of $\psi$ reveal the most dominant features of the spectral clustering problem, i.e., partitions of data of maximum dissimilarities. Thus, from the point of view of the graph clustering problem, the dominant features of the flow field at $O(\Lambda/D) \sim 1$ are the two gyres, as predicted from classical ECS theory (Jeong and Hussain, 1995), whereas at $O(\Lambda/D) \ll 1$, the dominant features are the locally-most stable and unstable material surfaces of the velocity field, as suggested by Haller and Yuan (2000).

4.5.2 High-Reynolds-number bluff-body data

The second benchmark conducted in this study consists of three-dimensional LPT data collected behind a bluff body at a free-stream velocity of $U_\infty \approx 8$ m/s ($Re \sim 6 \times 10^5$). In the current study, the total number of tracks in each test case is selected by imposing thresholded temporal track lengths of $L_{s,min}$, such that each analysed case consists of tracks of $L_s \geq L_{s,min}$ only. The analysis is conducted with $N = 1940, 289$ and $70$ particles, equivalent to inter-particle distances of $\Lambda \approx 0.065, 0.12$ and $0.20$, respectively. We concluded from a preliminary analysis that both the CSC and FTLE techniques fail to provide physically-distinguishable flow features with such challenging LPT data and hence in the current benchmark, results obtained with the proposed technique are compared only to baseline vorticity fields.

In contrast with the two-dimensional Voronoi tessellation, three-dimensional Voronoi
diagrams present complex topologies and make the visualization of coherent structures in LPT data quite challenging. We thus opted to omit the Voronoi diagram from the current section as well. Instead, the three-dimensional track data is colour-coded directly with the corresponding \( \chi \) values. Subsequently, a point-in-cell interpolation is adopted to map the \( \chi \)-coloured tracks onto three-dimensional \( \chi \)-fields. The results of the proposed approach are then compared to normalized cross-flow vorticity fields, \( \omega_1 D/U_\infty \). Both \( \chi \)- and vorticity-fields are analysed at three planes parallel to the cross-flow direction, at \( x_1/D = [-0.5, 1.25, 2.5] \). A constant bin size of \( \Delta x_i/D = 0.05 \) is adopted for the interpolation of the velocity- and \( \chi \)-fields, which are subsequently subjected to a partial reconstruction process using the robust smoothing algorithm developed by Garcia (2010).

Results for the challenging three-dimensional bluff-body data with \( \Lambda \approx 0.065 \) are
Figure 4.10: Left and right panels show $\chi$- and vorticity fields, respectively, obtained from the point-in-cell interpolation of the LPT data onto a grid. Results obtained with $\Lambda \approx 0.12$ at $x_1/D = [-0.5, 1.25, 2.5]$. Panel to the right also shows $\chi$-coloured tracks.

Figure 4.11: Left and right panels show $\chi$- and vorticity fields, respectively, obtained from the point-in-cell interpolation of the LPT data onto a grid. Results obtained with $\Lambda \approx 0.20$ at $x_1/D = [-0.5, 1.25, 2.5]$. Panel to the right also shows $\chi$-coloured tracks.
shown in Fig. 4.9, in which the left panel shows $\chi$-fields overlaid by $\chi$-coloured tracks of the current technique, and the right panel shows the vorticity field resulting from the mapping of LPT data onto a grid. As guaranteed by the analysis conducted with the double-gyre flow, the vortex boundaries of the C-pillar vortex are well-defined in the $\chi$-coloured tracks with the selected inter-particle distance value. In contrast, the vortex core, which is located near the center of these planes, can only be partially visualized by the vorticity field.

Results for $\Lambda \approx 0.12$ are shown in Fig. 4.10. With a mean inter-particle distance of about one-tenth of the characteristic length scale of the flow, $\Lambda/D \sim \mathcal{O}(10^{-1})$, the Eulerian approach (right panel) lacks a nested sequence of smooth closed contours that allows for coherent-structure detection. Nevertheless, an intuitive visual inspection of the $\chi$-field in the left panel still picks up the convex, C-pillar vortex core at the center of the $\chi$-coloured track data. Lastly, results for $\Lambda \approx 0.2$ are shown in Fig. 4.11. Interpolation of the velocity field with a large $\Lambda$ value is naturally unreliable, indicated by the poorly-reconstructed vorticity field in the right panel of the figure. On the other hand, as shown in the left panel, boundaries of a convex structure representing the C-pillar vortex can still be partially visualized by the $\chi$-fields, indicating that coherent structures can be identified with the current technique despite the severe sparsity of the investigated experimental data set. Notwithstanding, due to the evident sparsity and non-homogeneous tracer particles distribution of the current data set, we are unable to go much beyond a general visual assessment of the results. Nonetheless, the current technique, in view of the results presented in this study, proves to be a strong candidate for the detection of coherent structures in realistic, very sparse LPT data.
It is important to highlight the fact that the analysed flow is a very challenging benchmark due to data gaps near the core of the C-pillar vortex. Data gaps in the interior of a vortex structure are a consequence of the intense drifting away of tracers from the vortex core due to centrifugal forces (Marshall, 2005). In the analysed data set, strong shear is also responsible for the bursting of tracers (helium-filled soap bubbles) near the center of the C-pillar vortex (Faleiros et al., 2019). The combination of both effects causes radial-dispersion of tracers and data-gaps along the vortex core, omitting or strongly dampening the coherent structure in diagnostic fields of gradient-based approaches, despite the average particle concentration of the data set (see vorticity fields in Fig. 4.9, for example). Unlike the neighbouring times of particles advected by the mean flow, which are typically high, the complex advection processes of tracers in regions of high shear forces yield low neighbouring times. Unlike in the classical approaches, such contrast between the behaviour of particles inside and outside of vortical regions allows for the visualization of coherent structures using the proposed technique even if the particle concentration in the vortex core region is considerably low (see results in Fig. 4.11).

Moreover, in some cases, strong rotational motion of coherent structures may also induce the formation of a thin sheet of high particle concentration trapped between the columnar vortex core and the external vortical structure (Marshall, 2005), as shown in the results for high particle concentrations for the double-gyre flow (see bottom panels of Fig. 4.7). In that case, the coherent structures visualized through the proposed technique would assume the shape of the ridges forming along the exterior shear layers of the coherent structures, i.e., the attracting surfaces of the flow (see Voronoi cells colour-coded in red in Fig. 4.7). Although the visual interpretation of these coherent
structures is straightforward, time-resolved data may be necessary in order to indicate to the user the direction of the coherent motion. Nonetheless, the manifestation of coherence as attracting ridges, as observed here in the double-gyre flow, is not always achievable in realistic data sets. The visualization of attracting ridges separating vortical structures from the mean flow requires inter-particle distance values to fit to characteristic length-scales of the very thin shear layers formed in the outer layers of a vortex core. In such cases, coherence obtained with the present technique would only be manifested as the dissimilarity between tracers inside and outside a vortical structure.

4.6 Conclusions

A robust algorithm to detect coherence on very sparse Lagrangian particle tracking data, using Voronoi tessellation and techniques from spectral graph theory, is proposed and tested. Neighbouring time of tracer trajectories, defined as the total flow time two Voronoi cells remain connected by a Voronoi edge, was adopted as a metric for coherence. The novel approach was tested with LPT data from a simple double-gyre flow and then with challenging, realistic LPT data past a bluff body. The method was compared to CSC-coloured tracks of the CSC technique, FTLE ridges and to baselines vorticity field. In this study, capabilities of visualization was evaluated using as a metric the inter-particle distance. It was concluded that, in general, higher accuracy in coherent-structure identification is obtained for decreasing inter-particle distances. For low inter-particle distance values, colour-coded Voronoi cells of the current technique were found to exhibit patterns that are strongly correlated to the attracting surfaces (unstable manifolds) of the FTLE technique. For very
high inter-particle distance values, on the order of the characteristic length scale of the flow, coherence-detection is still achievable with the proposed technique, whereas classical approaches broke down.
Chapter 5

Conclusions

As presented in this thesis, Lagrangian methods for coherent-structure detection present a pivotal step for the assessment of the underlying fluid dynamics in important biological, engineering and geophysical flows. It was therefore the main goal of this thesis to advance theories in coherent-structure detection, as well as to develop a robust trajectory-based detection technique that allowed us to observe flow features in a realistic, very sparse data set. In this chapter, we recapitulate the major findings from this study, and comment on the contributions made to knowledge, and the limitations of the studies conducted in Chapters 3 and 4. Lastly, we end this thesis with considerations for future studies.

5.1 Summary

As noted in a recent review conducted by Hadjighasem et al. (2017), methods for coherence detection in sparse data sets have been sorely lacking in the published literature to date, motivating the current study. While most techniques presented are not readily appropriated for the detection of coherent structures within sparse data sets, few prospective approaches present alternative mechanisms that would
a priori grant them the potential to identify coherence in realistic fluid problems. Among these recently proposed approaches, the Coherent-Structure Colouring (CSC) technique, first tested by Schlueter-Kuck and Dabiri (2017b) and assessed in this study, presented an attractive framework for the observation of coherence in fluid flows.

The replacement of calculations of velocity-field gradients by the track’s kinematic dissimilarity, as done in the CSC framework, proved useful to visualize coherent structures in densely-seeded synthetic flows, while maintaining results Lagrangian-invariant. However, as noted in Chapter 3, coherence-detection with the CSC approach is inaccurate for typical tracer concentrations of realistic flows, i.e., $\Lambda \ll 1$. In fact, we proved, based on Shake-The-Box data past an Ahmed reference body and of a swirling jet, that flow-features extraction with the CSC approach is unreliable for inter-particle distances, $\Lambda$, that exceeded 15% of the characteristic length scale of the flow. Thus, the CSC technique was found to be insufficient for the detection of coherence in realistic, sparsely-seeded flows.

The process of clustering tracer trajectories based on their kinematic dissimilarities can also be flawed, e.g., trajectories that belong to distant regions of the potential-flow region move relatively parallel to each other and, therefore, would be mistakenly clustered together as belonging to a coherent structure; Moreover, in flows with strong mixing, a pair of tracks of low kinematic dissimilarity can be instantly located at distant regions of the flow. Thus, a visual assessment of instantaneous CSC-coloured tracks will not always reveal closed (convex) sets of particles composing a coherent structure.

Nonetheless, we believe that the use of spectral graph theory, as done in the
CSC framework, is a promising mechanism for very sparse data sets since it sets the minimum required number of tracer trajectories in the flow to, axiomatically, as few as three particles only (Bezdek et al., 1984). Instead, we argue that the adopted definition of coherence of the CSC technique, i.e., low kinematic dissimilarity, is the cause of its limited potential for sparse data sets. Based on all these conclusions, we propose and test in Chapter 4 a robust technique for coherence detection that replaces the kinematic dissimilarity with the neighbouring times of tracer distributions as a metric for coherence, while still adopting CSC-technique’s spectral clustering approach.

In the proposed framework, a spectral clustering problem was fed with Voronoi cell’s neighbouring times, which allowed us to assess coherent structures with very sparse data sets (a minimum of 16 tracers only, for a double-gyre flow, and 70 particles, for a challenging LPT data set). Moreover, as illustrated with a simple double-gyre flow, the proposed technique offered convergence towards the attracting surfaces (unstable manifolds) of the FTLE technique for more densely-seeded flows, while still being able to pick up coherent structures within very sparse data sets. We believe therefore that the proposed combination of Voronoi tessellation and techniques in spectral graph theory presents a promising framework for coherence detection in realistic flows.

5.2 Perspectives

The long-term purpose of this study is to introduce and test a robust approach for coherent-structure detection in very sparse LPT data, so as to generate Lagrangian-invariant alternatives to flow analysis in situations of challenging application. While
we consider this thesis successful at providing such diagnostics, one limitation of our study is that the application of the aforementioned approach has generally been limited to two test cases only, and there is unarguable gain in extending the conducted analyses to a variety of other applications.

Extension of the presented approach to ridge detection in geophysical models and to biological applications, e.g., the transport of pathogens in atmospheric flow (Schmale III and Ross, 2015), would be valuable to the grounding of the technique presented here. Extending track-based approaches to oceanic floats’ data (Puckett et al., 2012) has been a challenge due to the characteristic sparsity of these data sets, and the authors believe that the present technique has potential value in such applications. In the range of biological applications, swarming and colony phenomena (Allshouse, 2013; Zhang et al., 2010) also represent challenging, sparse problems, with great potential value for the analysis using the current technique.

Aside from the extension of the range of applications, it is also reasonable to assume that there is a time-dependent variation in the results of which the effects could not be accounted for in the presented test study. In Chapter 4, analyses of the proposed coherence-detection technique were carried out in investigation-time windows on the order of the characteristic periods of the flows. However, we believe that the exponential decay proposed for the Adjacency matrix’s components,

\[ A_{ij} = \left( \frac{1}{2} \right)^{n_{Tij}}, \]

will naturally yield singularities in the spectral clustering problem for sufficiently high neighbouring time values, \( n_{Tij} \). Notice that the neighbouring time is an inverse function of the sampling frequency, hence such singularities could happen earlier.
5.2. PERSPECTIVES

in the results of high-sampling frequency data sets. The authors also believe that future studies should include the characteristic period of the flow normalized by the sampling period (inverse of sampling frequency) as a normalised parameter of future analyses. Very low values of this parameter would yield high neighbouring times for all, quasi-stationary, Voronoi cells and the method would eventually break down.
Bibliography


Pep Espanol and Mar Serrano. Voronoi fluid particles & tessellation fluid dynamics.


Sebastian Gesemann, Florian Huhn, Daniel Schanz, and Andreas Schröder. From noisy particle tracks to velocity, acceleration and pressure fields using b-splines and penalties. In 18th international symposium on applications of laser and imaging techniques to fluid mechanics, Lisbon, Portugal, pages 4–7, 2016.


Paul S Krueger, Michael Hahsler, Eli V Olinick, Sheila H Williams, and Moham-madreza Zharfa. Quantitative classification of vortical flows based on topological


Appendices
Appendix A

Flowcharts

Here we present in details the working principle of the proposed coherent-structure detection technique, which can be sub-divided in two parts: the first, concerning steps for the tessellation of the Voronoi diagram based on the instantaneous seeding distribution; and the second, in which a graph-clustering problem based on the Voronoi cell’s neighbouring times is built. Both steps are represented as Part I and Part II, respectively, in the schematic diagram shown in Fig. A.1.

In Part I, typical LPT data is converted into structured data, such that every particle is assigned an unique particle ID $\in I = \{1, N\}$ (see Fig. A.2), where $N$ represents the total number of tracked particles, $p_i, i \in I_N \in \{1, \cdots, N\}$. In this step, the Voronoi tessellation and Delaunay triangulation of the tracer distribution, $p_i, V(p_i)$ and $DT(p_i)$, are initialised. Sequentially, in Part II, spectral clustering of the graph $G(V)$ is done via the graph’s Laplacian, $\mathcal{L}(G)$, and it is such that Voronoi cells connected by a single edge of $DT$ for longer periods of times, $t$, are assigned smaller distances in the graph-representation space of $DT(p_i), A_{ij}$, through Eq. 4.3. The problem is then solved iteratively until $t_{\text{max}}$ is reached, where $t_{\text{max}}$ represents either the maximum temporal track length, $L_t$, or a user-defined threshold-period.
The algorithm starts in Part I by reading typical LPT data containing the absolute coordinates, \( x_i \), time, \( t \), and the ID of each track. The respective two- or three-dimensional tracer distribution, \( p_i \), is then assigned a Voronoi diagram \( V(p_i) \) and Delaunay triangulation, \( DT(p_i) \). Notice also that a provisional solution \( \chi^*(t) \) is created before the final solution \( \chi(t) \). This step allows for the correction of the orientation of the eigenvector \( \chi(t) \) with respect to \( \chi(t - 1) \): in case the former represents an inflection of the latter relatively to \( \chi = 0 \), an inversion of signals

\[
\chi(t) = \begin{cases} 
\chi^*(t), & \text{if } R(\chi^*(t), \chi(t - 1)) > 0, \\
-\chi^*(t), & \text{otherwise,}
\end{cases}
\]

(A.1)

is adopted so as to keep the colour-coding process consistent in time, i.e., to avoid colours switching when colour-coding the Voronoi cells.

In the present work, the Voronoi diagram is computed using the algorithm developed by Barber et al. (1996), namely the *QuickHull*. Their algorithm represents convex hulls as a set of facets and a set of adjacency lists numbering the Voronoi edges of each facet. The QuickHull starts by computing the Delaunay triangulation in \( \mathbb{R}^d \), using convex hulls of dimension \( \mathbb{R}^{d+1} \), \( d \in I = \{2, 3\} \). To determine the Delaunay triangulation of a given particle distribution, extrema points are lifted to a paraboloid and their convex hull is computed. Next, a line segment is traced between the pair of particles that are furthest apart. Sequentially, a perpendicular line is traced between the furthest point from the first line segment to the current line segment, both intersecting perpendicularly to each other. Three points, defined as the ends and crossing of the two line segment, build the initial hull, which consists of a two-dimensional triangle. The process runs in an incremental fashion, adding
Read LPT data: 
\( x_1, x_2, x_3, t, \text{ID} \)

Determine: \( p_i, G(V), \mathcal{L}(G), \) and \( D(p_i). \)

Solve spectral clustering problem

Calculate solution at \( t, \chi^*(t) \)

If \( R(\chi^*(t), \chi(t-1)) < 0, \chi(t) = -\chi^*(t); \) c.c., \( \chi(t) = \chi^*(t) \)

\( t = t + 1 \)

no

\( \chi(t - 1) = \chi(t) \)

yes

\( t < t_{\text{max}}? \)

end

Figure A.1: Flowchart of the proposed LCS detection algorithm.

points and triangles to the convex hull of previously processed points.

The process of building the Delaunay triangulation from convex hulls can be visualized as follows: considering two triangles ABC and BCD, such as in Fig. A.3, if the sum of angles \( \alpha + \gamma \) is such that \( \alpha + \gamma \leq 180^\circ \), they meet the Delaunay condition. If they do not match the Delaunay condition, then switch the common edge from BD to AC, and test the criterion. Repeat until condition \( \alpha + \gamma \leq 180^\circ \) is matched. The
Figure A.2: Part I of the clustering algorithm. In this sub-part of the code (see Fig. A.1) track data is converted into a Voronoi diagram.

The complete algorithm, which takes place in I.1 in Fig. A.2, is described in details by Barber et al. (1996).

Starting from the Delaunay triangulation of the seeding distribution, the Voronoi diagram must be processed. This stage of the algorithm is indicated as I.2 in Fig. A.2. Edges of the Voronoi diagram are segments of the perpendicular bisector of the edges of the Delaunay triangulation. Computing such segments is straight-forward
Figure A.3: (left) Initial Delaunay triangulation consisting on two convex hulls. (center) the initial configuration using the common edge BD is non-Delaunay. (right) By switching the common edge to AC, the two convex hulls match the Delaunay criterion.

Subsequently, coordinates of the Voronoi edges must be determined and the Voronoi neighbouring matrix, $VN$, must be computed. $VN$ is a $N \times N$ matrix that is updated every time step, such that its $(i, j)$-th element is 1 if the $i$-th Voronoi cell share a common Voronoi edge with the $j$-th Voronoi cell at the $n$-th time-step, or 0, otherwise. Equivalently, it is 1 if both cells correspond to the extrema of the $m$-th, $m \in I = \{2, N - 1\}$, Delaunay edge:

$$VN_{ij} = \begin{cases} 1, & \text{if } (i, j) \text{ are extrema of the same Delaunay edge,} \\ 0, & \text{otherwise.} \end{cases}$$  \hspace{1cm} (A.2)

For instance, the $VN$ matrix of the Delaunay triangulation shown in the right-most panel of Fig. A.3 is given by
Calculate Adjacency matrix, $A$

Compute degree matrix, $D$

Compute graph’s Laplacian, $\mathcal{L}$

Solution at time $t$, $\chi^*(t)$

If $R(\chi^*(t), \chi(t-1)) < 0$, $\chi(t) = -\chi^*(t)$; c.c., $\chi(t) = \chi^*(t)$

$t < t_{\text{max}}$?

$t = t + 1$

no, $t = t + 1$

yes

Figure A.4: Part II of the clustering algorithm. In this sub-part of the code (see Fig. A.1) the kinematic dissimilarity of the Voronoi diagram is processed using a spectral-graph clustering problem.

$$VN = \begin{pmatrix}
1 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1
\end{pmatrix}, \quad (A.3)$$

for nodes-numbering such as $(A, B, C, D) \rightarrow (1, 2, 3, 4)$.

The second segment of the algorithm consists of the spectral-clustering step. The flowchart of this segment of the code is shown in Fig. A.4. The Adjacency matrix, $A_{ij}$,
is defined in Eq. 4.3. It is clear from Eq 4.3 that if the \( i \)-th and \( j \)-th Voronoi cells share a common boundary indefinitely, \( A_{ij} \to 0 \) as \( n \to \infty \). Following to the computation of the Adjacency matrix, the degree matrix, \( D \), and the graph’s Laplacian, \( L \), must be determined. The spectral-clustering problem,

\[
L X = \xi X,
\]

is then solved by finding the maximum eigenvalue \( \xi_{max} \) of the second-order tensor \( L \), such that \( \chi \equiv X(\xi_{max}) \). The components of the eigenvector \( \chi \) associated with the maximum eigenvalue \( \xi_{max} \) are then mapped onto a colour-map, such that the \( i \)-th Voronoi cell is colour-coded a \( \chi_i \)-colour.
Appendix B

MATLAB routines as pseudo code

Here we share the pseudo MATLAB routines, as pseudo code, implemented in this work for the computation of the CSC-coloured tracks, $\chi$-coloured Voronoi cells and for the point-in-cell interpolation of velocity fields based on input LPT data.

Algorithm B.1: Proposed coherent-structure detection technique in pseudo code as MATLAB a routine.

```matlab
% Pseudo code of MATLAB routine for the Voronoi-based LCSs
% detection algorithm
% Author: F. A. C. Martins. Date: November 2020
% Queen's University. Kingston, ON - Canada

% # INPUTS: The code reads LPT data in the format
% xp = input(points,1); % x-coord of particles at t
% yp = input(points,2); % y-coord of particles at t
% ID = input(points,3); % Particles' ID at t
% up = input(points,5); % x-coord of particles at t
% vp = input(points,6); % y-coord of particles at t
% npt = length(xp); % number of particles at t
% #OUTPUT: chi-values

points = find(input(:,4)==t); % all data points at t=t
xp = input(points,1); % x-coord of particles at t
yp = input(points,2); % y-coord of particles at t
```
ID = input(points,3); % Particles' ID at t
up = input(points,5); % x-coord of particles at t
vp = input(points,6); % y-coord of particles at t
npt = length(xp); % number of particles at t

% VE = voronoi edge's coordinates
% CL = voronoi cells
[VE,CL] = voronoin([xp, yp]); % calculate voronoi diagram
nc = size(CL,1); % number of cells at t

%==========================================================================%
% Compute the coordinates of the cell's edges:
%==========================================================================%
COORD = cell(nc,1);
for p=1:nc
  COORD{p} = [VE(CL{p}',1), VE(CL{p}',2)];
end

%==========================================================================%
% Returns: Which edges are connect to the i-th particle, i.e., the Neighbours cells (VN).
%==========================================================================%
VN = zeros(np,np);
for i = 1:npt
  for j= i+1:npt
    s = size(intersect(CL{i}, CL{j}));
    if (1<s(2))
      VN(ID(i), ID(j)) = 1;
      VN(ID(j), ID(i)) = 1;
    end
  end
end

%==========================================================================%
% Solve graph-clustering problem
%==========================================================================%
A = ones(np,np);
A(find(VN==0)) = A(find(VN==0))/10;
A(1:1+size(A,1):end) = 0; % sets main diagonal to zero

% Compute degree matrix:
D = zeros(np); % sums of adjacency matrix
for i=1:np
    D(i,i) = sum(A(i,:));
end

% Compute Laplacian matrix:
L = D - A; % Compute the graph Laplacian:
L = (D^-0.5)*L*(D^-0.5); % normalize L

%==========================================================================%
% Solve graph-clustering problem
%==========================================================================%
[eigVec,eigVal] = eig(L);

if mode == 0
    % Selects best eigenvector
    [~,index] = max(diag(eigVal));
    solution = real(eigVec(:,index)); % colors to plot
else
    % Selects another eigenvector that is expected to work better
    % for the data-set. In the future, an optimization function
    % should be able to identify the optimum mode-value.
    solution = real(eigVec(:,mode)); % colors to plot
end

%==========================================================================%
% Make sure that colouring scheme
% doesn't invert:
%==========================================================================%
R = corrcoef(solution,chi);
if R(2,1) < 0
    solution = -1*solution;
end

% Colour-code Voronoi cells with chi-values:
chi = solution;
Algorithm B.2: CSC technique in pseudo code as a MATLAB routine.

```matlab
% Pseudo code of MATLAB routine for the CSC algorithm
% Author: F. A. C. Martins. Date: November 2020
% Queen's University. Kingston, ON - Canada

% # INPUTS: The code reads LPT data in the format
% - x[Nt,Np], y[Nt,Np], z[Nt,Np],
% - Fs represents the sampling frequency [1/s]
% - Nt represents the (fixed) number of time-steps
% - Np is the number of seeding particles.
% #OUTPUT: CSC-values for the Np tracks

rsum = zeros(Np,Np);  % SUM(R(t)) for all t
for t=1:Nt  % iterate over time
    for p1=1:Np  % iterate over particles
        xp1=x(t,p1); yp1=y(t,p1); zp1=z(t,p1);
        for p2=p1:Np
            xp2=x(t,p2); yp2=y(t,p2); zp2=z(t,p2);
            r(p1,p2) = sqrt((xp1-xp2)^2 + (yp1-yp2)^2 + (zp1-zp2)^2);
        end
    end
    R(:,:,t) = r+r';  % Adjacency matrix at the time t
    rsum = rsum + R(:,:,t);  % SUM(R(t)) for all t
end

Tf = (1/Fs)*(Nt-1);  % last physical time-step
s = zeros(Np,Np);  % standard-deviation matrix
for t=1:Nt  % iterate over time
    s = s + (rsum/Nt-R(:,:,t)).^2;
end
A = (1/sqrt(Tf))*(sqrt(s)./rsum/Nt);  % apply definition of A

D(i,i) = sum(A(i,:));

L = (D^-0.5)*(D - A)*(D^-0.5);  % normalized Laplacian matrix
```
% Calculate CSC solution  
% --------------------------------------
[eigVec,eigVal] = eig(L);  % Solve eig.val./eig.vec. problem
[-,Index] = max(diag(eigVal));  % find max. eig. val.
CSC = real(eigVec(:,Index));  % Solution = colors to plot
Algorithm B.3: Routine for binning LPT data onto three-dimensional Cartesian grids as a MATLAB routine.

```matlab
points = find(input(:,5)==time); % all data points at time=t
xp = input(points,1);       % x-coord of particles at t
yp = input(points,2);       % y-coord of particles at t
zp = input(points,3);       % z-coord of particles at t
up = input(points,5);       % x-coord of particles at t
vp = input(points,6);       % y-coord of particles at t
wp = input(points,7);       % v-coord of particles at t
ID = input(points,4);       % Particles' ID at t
npt = length(xp);          % number of particles at t

% Define grid:
%=====================================================================
bin = 0.1;                  % bin size
x = Axis(1):binsize:Axis(2);
y = Axis(3):binsize:Axis(4);
z = Axis(5):binsize:Axis(6);
x = length(x);
y = length(y);
z = length(z);
[X,Y,Z] = meshgrid(x,y,z);

% Define velocity fields:
%=====================================================================
U = NaN(ny,nx,nz);
V = NaN(ny,nx,nz);
```
$W = \text{NaN}(ny,nx,nz)$;

%===============================================================================
% Bin data:
%===============================================================================

for i=1:nx-1
    binx = find(xp>=x(i) & xp<=x(i+1));
    for j=1:ny-1
        biny = find(yp>=y(j) & yp<=y(j+1));
        binxy = intersect(binx,biny);
        for k=1:nz-1
            binz = find(zp>=z(k) & zp<=z(k+1));
            binxyz = intersect(binz,binxy);
            if ~isempty(binxyz)
                U(j,i,k) = mean(up(binxyz));
                V(j,i,k) = mean(vp(binxyz));
                W(j,i,k) = mean(wp(binxyz));
            end
        end
    end
end

%===============================================================================
% Smooth velocity-field:
%===============================================================================

U = fillmissing(U,'linear','EndValues','none');
V = fillmissing(V,'linear','EndValues','none');
W = fillmissing(W,'linear','EndValues','none');
Appendix C

Metrics for self-consistent results and correlation to FTLE technique

Standards imposed by the scientific community to new candidate Lagrangian techniques require that relevant new coherent-structure detection methods are self-consistent. In a recent publication, Haller (2015) proposed a set of basic requirements for self-consistent coherent-structures results that serves as a guideline for future development of Lagrangian techniques. In his work, Haller (2015) imposes the following four criteria:

i. Objectivity: a candidate method must provide diagnostics that are invariant under Euclidean coordinate changes. Instantaneous Eulerian techniques are therefore non-objective (Haller, 2005);

ii. Finite time: A real flow’s patterns evolve and are advected during finite time. Within finite times, asymptotic methods such as chaos-theory-based and stability analysis become mathematically undefined (Martins and Zanotello, 2018);

iii. Lagrangian invariance: requires that detected structures are uniquely defined along the assimilation-time window, i.e., coherent structures evolve into each other;

iv. Spatial convergence: In order to ensure that the results uniquely identifies a material point as being part of a LCS, the candidate vortical detection technique needs to converge to an unique solution under data refinement (Stone et al., 1991).

The majority of proposed coherence-detection techniques to date do not match these criteria (Hadjighasem et al., 2017). The current approach, on the other hand, matches the four criteria for self-consistent features-extraction. As follows:
i. The proposed, Voronoi-tessellation based technique adopts neighbouring times, $n_{T_{ij}}$, of tracer particles, $p_i, p_j$, as a criterion for coherence. Such metric for coherence is independent of reference frame, making it objective;

ii. The proposed method also proved capable of feature-extraction in periods of time, $t$, on the order of the characteristic period, $T$, of the flow, $t \sim T$. Thus, the technique is also finite-time;

iii. Coherence in the proposed framework is uniquely obtained based upon the definition of a single graph $G(V)$ that shares common topology with the Delaunay triangulation of the particles distribution, $p_i, DT(p_i)$. Since any Delaunay triangulation (axiomatically) uniquely assigns IDs to tracer particles (Chew, 1989), so does $G(V)$. Therefore, solutions of the proposed technique are also Lagrangian invariant;

iv. Lastly, the technique proved to converge coherent structures under data refinement. Thus, the method is also spatially-convergent.

In Chapter 4, the authors argued that results obtained with the current technique have strong correlation to ridges of the FTLE technique for densely-seeded flows. In this section of the appendix, we adopt theories in dynamical system analysis (Haller and Yuan, 2000; Tu et al., 2018) to give mathematical foundation to such claim.

Let the instability time, $T_u(x_0, t_0)$, in the time interval $[t_0, t_1]$ be defined such as

$$T_u(x_0, t_0) = \frac{1}{t_1 - t_0} \int_{M(x_0)} dt.$$  \hspace{1cm} (C.1)

In the FLTE framework, a material surface $M(x_0)$ along which $T_u$ is locally minimal acts as an unstable manifold (attracting surface) of the velocity field (Haller and Yuan, 2000), i.e., it forms a material barrier between two vortex cores. Nonetheless, the current technique identifies coherent regions as groups of Voronoi cells of high neighbouring time or, equivalently, in groups of Voronoi cell that present relatively low instability times, $T_u$. Based on these observations, spatially-converged results of the proposed technique present an alternative assessment of $T_u^{-1}$, which, on its turn, is also an alternative description of the attracting LCSs of the FTLE technique (Haller, 2000).