A Quantitative Evaluation of Lagrangian Coherent Structure Detection Methods Based on Computational and Experimental Limitations

Aleksandar Badza

May 24, 2022

Thesis submitted for the degree of
Doctor of Philosophy
in
Applied Mathematics
at The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences
School of Mathematical Sciences

Contents

List of Tables

List of Figures

Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Signed: .... Date: 24/05/2022

Acknowledgements

This work was supported by the MATLAB programming software developed by MathWorks, along with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide. I wish to thank Associate Professor Sanjeeva Balasuriya and Dr Trent Mattner for supervising this research and for their support throughout this endeavour. I also wish to extend a hand of gratitude to Prof. Oliver Junge at the Technical University of Munich for his assistance regarding the development of MATLAB code, along with Prof. Gary Froyland of the University of New South Wales for suggesting a time parameterisation for one of the flow systems considered in this research.

Dedication

I wish to dedicate this Thesis to my family, who have supported me at every step of my candidature. I especially wish to pay tribute to my father Milan, who sadly passed away during my candidature and was one of the best mentors and influences a man could ask for.

Abstract

Lagrangian coherent structures are used in fluid mechanics and the analysis of dynamic systems to visualise the most influential flow structures present within a velocity system over a finite period of time. Over the last two decades, a wide variety of methods have been conceptualised for the numerical detection of various forms of these structures within different flows. These include continuous curves of maximal particle repulsion which act as flow barriers, two dimensional objects such as jets or eddies formed from more robust flow behaviour, or larger partitions which remain separated from the rest of the domain over an entire flow interval. While some studies which focus on comparing the basic functionality of groups of these methods have been undertaken, the impact of certain computational factors such as the uncertainty of velocity data or the available resolution of said data on the resultant structures generated from these methods has seldom been investigated. In this Thesis, we address both of these issues by performing a systematic analysis of eight of these Lagrangian coherent structure detection methods using a variety of velocity systems including analytically defined flows (such as the Double Gyre, a non-autonomous Stuart vortex system and the Bickley jet), computational fluid dynamics velocity data (corresponding to flows which each contain two layers of Kelvin-Helmholtz instability) and an oceanographic velocity data set representing the Gulf Stream.

The methods we consider here are the finite time Lyapunov exponent (a measure of the exponential stretching rate of flow trajectories), variational Lagrangian coherent structures (geodesic solutions of variational problems related to flow stretching), Lagrangian averaged vorticity deviation (an objective measure of the vorticity of a flow trajectory against that of the entire domain), stochastic sensitivity (the expected uncertainty of a Lagrangian flow trajectory), the transfer operator (a probabilistic method which seeks density distributions that remain coherent), the dynamic Laplace operator (an extension of the transfer operator method which explicitly includes diffusivity), fuzzy c-means clustering (grouping together collections of flow trajectories based on their consistent proximity) and coherent structure colouring (identifying coherent flow objects from how similarly groups of flow trajectories evolve as a flow advances). We compare the types of Lagrangian coherent structure each method is able to produce, and test how these methods react to the addition of stochastic noise to the velocity data which represents a flow. From our results, methods which detect two-dimensional coherent flow structures rather than the boundaries which separate them, such as coherent structure colouring, Lagrangian averaged vorticity deviation, stochastic sensitivity, the transfer operator and dynamic Laplace operator; are less sensitive to velocity uncertainty and give a more thorough picture of the most influential flow behaviour observable. We also perform a detailed analysis on the impact of spatial resolution in comparison to the size of coherent structures for each of the methods, both qualitatively by visually comparing the coherent structures produced and quantitatively using the absolute errors of various LCS quantities against a "reference case" produced from the best velocity data resolution available.

Chapter 1

Introduction

The study of coherence in general, or the identification of consistent particle patterns and movement within a velocity system, is one of the key foci of fluid dynamics and mechanics. Analysis of dynamical systems provides understanding into the dominant movement of particles, pollutants, or even people, within these systems. Elementary studies of coherence primarily involved observing a static velocity field at a particular instance in time, and studying the system for stable/unstable stagnation points and manifolds (curves in a two–dimensional flow) which flow trajectories within these systems attract towards or repel from, thereby giving some illustration of how we can expect flow trajectories to evolve as they move through a flow (Balasuriya et al. 2018). One can also look for instantaneous coherent surfaces or other objects of dimension higher than one (i.e. other than a point or a curve) using the instantaneous entropy or vorticity of the flow velocity at a particular moment in time (Haller & Yuan 2000).

If we aim to seek coherent flow patterns within an autonomous velocity field, then this is a perfectly appropriate method for identifying dominant flow characteristics. However, the vast majority of flow systems, particularly those in real world settings, are non–autonomous systems within which the velocity field changes as time proceeds. Any analysis performed on an individual Eulerian velocity field at one instance in time is only relevant for that instance of time, and does not provide a practical or reliable illustration of transport.

For a long time, this Eulerian analysis of velocity field snapshots was the only way to obtain any sort of qualitative description of the most important flow behaviour. That all changed when Haller & Yuan (2000) first coined the term Lagrangian coherent structures, which are flow objects that are coherent under Lagrangian motion (that is, following the flow). This takes into account the evolution of a trajectory over a set time interval, and the consistent movement of collections of these trajectories over this time is used to provide a qualitative or quantitative measure of flow coherence for this interval. Portions of the flow domain corresponding to similar patterns of flow behaviour, such as curves along which particle flow is more extensive and unpredictable or regions corresponding to more robust flow patterns; are identified as Lagrangian coherent structures which illustrate how parcels of flow particles evolve together as a flow progresses.

In their paper, Haller & Yuan (2000) defined the first method for Lagrangian coherent structure detection, involving Lagrangian measures of stretching and compression rates of flow trajectories, and curves associated with minimal or maximal stretching were identified as Lagrangian tracers. The method was defined for flows in two–dimensional domains, as are the majority of Lagrangian coherent structure detection methods. There is emerging research into Lagrangian coherent structures in higher dimensions (Blazevski & Haller 2014, Froyland et al. 2010, Froyland & Junge 2018, Haller 2001).

Following up from this development, a range of computational methods used to detect these structures were conceptualised, each one with a different definition of what constitutes a Lagrangian coherent structure. The first, most well known and possibly the easiest to use Lagrangian coherent structure detection method was developed - the finite time Lyapunov exponent method (Haller 2001, 2002, Shadden et al. 2005). The finite time Lyapunov exponent is a measure of the exponential stretching capability of a disk of fluid centred at each initial location over a forward time flow interval (in backward time, it is a measure of compression capability instead). By viewing this measure over all initial positions, this defines a field. Maximal ridges of this field indicate coherent curves that are claimed to correspond to flow barriers which particles repel from (or attract to) at a maximal rate. Despite developing and initially promoting this method for its use in the detection of Lagrangian coherent structures, Haller (2011) later identified issues in using the finite time Lyapunov exponent to find these structures based on degenerate results obtained from applying this method to (mostly) simplistic toy flow models such as autonomous saddle flows. In light of these developments, a new detection method was created and used to detect material hyperbolic flow barriers similar in nature to the maximal ridges of the finite time Lyapunov exponent using variational problems relating to measures of flow stretching (Haller 2011, Farazmand & Haller 2012, Onu et al. 2015); which was later expanded in order to detect closed elliptic structures of high vorticity (Haller & Beron-Vera 2012, 2013) and parabolic current or jet core type structures (Farazmand et al. 2014).

Over the next decade or so, more and more detection methods were conceptualised for finding different types of Lagrangian coherent structures, and not just to find coherent codimension–1 curves. Joseph & Legras (2002) and d'Ovidio et al. (2004) used the finite *size* Lyapunov exponent to detect coherent structures in similar fashion to the finite *time* Lyapunov exponent except the time taken for two flow particles to be stretched apart by a particular, pre–determined length is the quantity that needs to be evaluated. Mezić et al. (2010) developed mesochronic analysis which is used to categorise portions of a flow domain based on the eigenvalues of the gradient (Jacobian) of a Lagrangian flow map, while Mancho et al. (2013) found a way to measure flow coherence by calculating the average velocity for flow trajectories over the flow time interval. The previous two methods come with a considerable drawback in that they are not objective, that is, the results of these methods depend on the flow system's frame of reference. Haller et al. (2016) sought to define "coherence" in terms of a frame–independent vorticity, and developed the Lagrangian averaged vorticity deviation; which is an objective measure of the vorticity of a flow trajectory relative to the mean of this quantity over the entire domain; used to find elliptic coherent structures of strong rotational coherence.

Other detection methods have been conceptualised that are not "geometric", that is, they do not rely on computing characteristics of the flow map. As examples of these kinds of methods, Froyland (2005) developed the transfer operator method, and later the dynamic Laplace operator method (Froyland 2015, Froyland & Junge 2018), which are both used to determine density distributions which change only marginally by the flow. As a first step in seeking "coherent sets", these fields were used to divide a flow domain into two coherent regions. Other methods which relied on the transitional matrix used to execute the transfer operator method were later conceived, such as finite time entropy (Froyland & Padberg-Gehle 2012), which is used to estimate how much stretching flow particles will undergo based on the probabilistic flow information obtained from the transfer operator matrix, or defining hierarchical coherent sets by repeating the transfer operator method on the coherent sets produced (Ma & Bollt 2013).

Alternative definitions of "coherence" were conceived through a series of sparse detection methods, ideal for use in a flow system defined by a collection of Lagrangian flow trajectories instead of Eulerian velocity data. Allshouse & Thiffeault (2012) used the technique of braiding to define coherence by how often trajectories mix or intersect with each other as a flow progresses. Froyland & Padberg-Gehle (2015) proposed dividing a flow domain into a pre–determined number of coherent clusters based on the spatial proximity of flow trajectories over their flow interval using fuzzy clustering techniques. As an alternative to this, Schlueter-Kuck & Dabiri (2017) developed an alternative clustering technique which involves colouring portions of a flow domain based on the kinematic consistency of flow trajectories (i.e. how similarly these trajectories evolve and mix together as a flow proceeds) within these portions.

As a final example, Balasuriya (2020*a*) developed the concept of stochastic sensitivity, which is a measure of the expected uncertainty of flow trajectories. This can be used to obtain flow barriers as well as robust flow regions. There are many more Lagrangian coherent structure detection methods which have been conceptualised in addition to the aforementioned, which have been discussed in a wide array of review papers into this subject (Allshouse & Peacock 2015, Hadjighasem et al. 2017, Balasuriya et al. 2018).

As more Lagrangian coherent structure detection methods were conceptualised and more researchers developed an understanding, and even an appreciation, of the concept; these detection methods were applied to a wide variety of real world flow systems. The vast majority of Lagrangian coherent structure applications are concerned with detecting coherent flow objects within oceanic flow systems, and the finite time Lyapunov exponent method became one of the most commonly used methods; owing primarily to its relative computational simplicity against other methods. Researchers have used patterns emerging from the scalar field of the finite time Lyapunov exponent to identify coherent flow patterns and predict the transport of debris or pollutants within bodies of water. These applications range in scale from identifying flow separating barriers within a smaller body of water such as the Moreton Bay (Suara et al. 2020) or the Pearl River Estuary (Wei et al. 2018); to visualising larger coherent flow objects such as currents within the Indian Ocean (Garcia Llamas et al. 2020). Shadden et al. (2006) have even used the FTLE method to model the formation of vortical coherent structures within bodies of water caused by the propagation of jellyfish.

There are several bodies of water that are regularly studied through the analysis of Lagrangian coherent structures, with one of these systems being the Gulf of Mexico. After the Deepwater Horizon oil spill in 2010, several oceanographic researchers used Lagrangian coherent structure detection methods to model how oil within the Gulf of Mexico would flow through the system and over how much of the Gulf it would spread. Modelling how the oil would be advected through the system was accomplished using several detection methods, including the hyperbolic variational Lagrangian coherent structure method (Olascoaga & Haller 2012), mesochronic analysis (Mezić et al. 2010), the hierarchical transfer operator (Ma & Bollt 2013), the finite time Lyapunov exponent and the transfer operator (Bollt et al. 2012). Analysis of Lagrangian coherent structures is also used to illustrate coherent flow behaviour within the Agulhas current, a cold current in the Indian Ocean off the southern coast of South Africa known for its large eddies which flow to the Southern Atlantic Ocean transporting warmth (Byrne et al. 1995) and acidity (Orselli et al. 2019) to this body of water. Researchers have identified the shapes, sizes and locations of these Agulhas eddies or rings and have traced their movement over particular flow periods; using the elliptic variational Lagrangian coherent structure method (Haller & Beron-Vera 2013, Karrasch et al. 2014), the transfer operator method (Froyland et al. 2015) and the dynamic Laplace operator method (Froyland & Junge 2018). Lagrangian coherent structure analysis has also been implemented on the Gulf Stream, a warm current in the North Atlantic Ocean known for its impact on climate in North America and Western Europe (Liu et al. 2018). Researchers have attempted to predict coherent fluid transport from flow separating curves obtained from the finite time Lyapunov exponent (Liu et al. 2018) and the finite size Lyapunov exponent (Haza et al. 2012). In addition to this, the dynamic Laplace operator has been used to identify the locations of warm coherent eddies produced by the Gulf Stream (Froyland et al. 2019). Lagrangian coherent structure analysis has also been performed on the entire global ocean, with the transfer operator (Froyland et al. 2014) and fuzzy c–means clustering methods (Froyland & Padberg-Gehle 2015) used to divide the global ocean into smaller coherent clusters, with bodies of water such as the North/South Atlantic Ocean, the North/South Pacific Ocean and the Indian Ocean identified as these coherent clusters. These results were primarily used to model how patches of garbage and other pollutants move through the ocean.

Applications for Lagrangian coherent structure detection methods are not just restricted to oceanic flow systems, with these methods having been used to find coherent flow behaviour within atmospheric flows. An example of this involves investigating the atmospheric flow around the well known Antarctic Polar Vortex, using the transfer operator (Froyland et al. 2010) or finite time entropy (Froyland & Padberg-Gehle 2012). The finite time Lyapunov exponent has also been used to illustrate the split of the polar vortex witnessed in 2002 (Lekien & Ross 2010). The wind patterns over the east coast of the United States are a popular choice for Lagrangian coherent structure analysis among several researchers, in an effort to predict weather patterns and measure the dispersement of pollutants through the atmosphere. The finite time Lyapunov exponent method has been used on several occasions to identify flow separating barriers within this atmosphere (Garaboa-Paz et al. 2017, Mancho et al. 2013, Nolan et al. 2020, Senatore & Ross 2011). Lagrangian coherent structure analysis has also been used to identify the prominent Red Spot within the atmosphere of Jupiter using parabolic and elliptic variational Lagrangian coherent structures (Hadjighasem & Haller 2016). Hadjighasem et al. (2017) have also used a range of Lagrangian coherent structure detection methods to identify Jupiter's Red Spot, done for the sake of comparative analysis of these methods. Applications of Lagrangian coherent structure detection methods also extend to healthcare, with Teeraratkul et al. (2021) using the finite time Lyapunov exponent to model coherent flow patterns within blood flow to monitor the spread of disease within the bloodstream, and Benaichouche et al. (2013) and Verma et al. (2016) using fuzzy c–means clustering to divide an MRI scan of a brain into segments for further medical analysis.

With the present day saturation of Lagrangian coherent structure detection methods, comparative studies on collections of these methods have been undertaken by a number of researchers. However, these comparisons are mostly limited to qualitatively assessing the types of Lagrangian coherent structures produced by these methods (Allshouse & Peacock 2015, Balasuriya et al. 2018), along with some minor comparisons of the algorithmic complexity of each method taking into consideration factors such as the number of computational parameters that are required to be defined for each of these methods (Hadjighasem et al. 2017). Most of these comparisons between coherent structure detection methods are ill–defined, as every Lagrangian coherent structure detection method has its own definition of what constitutes a coherent structure. There has been little, if any, research and analysis undertaken into the *self–consistency* of these methods and the *reliability* of the results. Factors affecting this include the spatial resolution of velocity and how well this data is refined over a flow domain; and the accuracy of this velocity data due to experimental, observational or modelling errors. As such issues are unavoidable when trying to analyse a real world flow system in this manner, it is imperative that these are taken into consideration when choosing a Lagrangian coherent structure detection method to work with. Developing a better understanding of this concept can allow for Lagrangian coherent structures to be used effectively in more mainstream fluid mechanics applications.

Hence, the aim of this research is to take eight Lagrangian coherent structure detection methods and compare the self–consistency and reliability of these methods; particularly in more *realistic* data situations, in order to determine which of these methods are the most suitable (or the "first choice(s)") for use in Lagrangian coherent structure detection. This will be achieved by writing our own numerical algorithms in MATLAB to execute these methods, implementing these on a variety of analytically and numerically defined flow systems to compare and contrast the types of structures each method is able to produce and, through careful manipulation of the velocity data representing these systems, quantify how each of these methods respond to numerical uncertainty or spatial resolution of the data. The methods we will consider in this study include "geometric" methods which compute coherent structures based on characteristics of the Lagrangian flow maps of particles (such as the finite time Lyapunov exponent, variational Lagrangian coherent structures, Lagrangian averaged vorticity deviation and stochastic sensitivity), probabilistic methods used to divide a flow domain into coherent sets (the transfer operator and dynamic Laplace operator), and sparse methods which can detect coherent flow objects from flow trajectory data with no necessary reliance on the underlying velocity data defining a system (fuzzy c–means clustering and coherent structure colouring).

The structure of this Thesis is as follows. In Chapter 2, we detail the methodology pertaining to our eight methods of choice, discuss the coherent structures we expect to find from these methods based on previous research, and briefly outline the computational procedures undertaken to implement each of these methods numerically. In Chapter 3, we then test all of these methods on a range of flow systems and compare the kinds of coherent structures each method is able to produce. In this Chapter, we focus on six flow systems. The first three systems, the Double Gyre, an unsteady version of the Stuart vortex system and the Bickley jet; are analytically defined toy models with velocity data defined by continuous equations. The other three systems, two simulated flows with Kelvin–Helmholtz vortex layers and an observational data set representing the Gulf Stream; are numerically defined flows available at finite resolution levels. We have implemented our methods on more than just these six systems, however for the sake of brevity we will not document all of these results within this Thesis. We show some of the results obtained for these other systems in Appendix A.

In Chapter 4, we assess how the functionality of each method is affected by the existence of uncertainty within velocity data. We achieve this by transforming each flow system into a stochastic dynamical system by applying normally distributed noise to our Eulerian velocity. We then quantitatively assess this impact by visualising the mean and standard deviation of quantities and attributes relevant to each method after producing a large enough number of stochastic flow map realisations. The noise we apply to our velocity changes with each time step, but is equal over the whole spatial domain in the spirit of random dynamical systems.

In Chapter 5, we perform a similar analysis of our methods this time focusing on how the spatial velocity resolution affects the accuracy of the results. This is another important consideration, for if a velocity data set has poor or limited spatial resolution, this will reduce the accuracy of any Lagrangian flow maps produced and any subsequent Lagrangian coherent structures detected. To assess this, we sub–sample a grid of finite velocity data to reduce the resolution of this data, and then implement the same Lagrangian coherent structure detection methods on these data sets. If we are working with an analytically defined velocity system, we produce an interpolant for the velocity of this system defined over a uniform grid of points. We then (where possible) quantify the differences in relevant Lagrangian coherent structure quantities by analysing the statistics of the error distribution between each resolution and the "reference case". The "reference case" quantities are computed either from analytic velocity equations or the highest resolution possible for a numerically defined data set.

Finally, in our concluding Chapter (Chapter 6), we assess the strengths and weaknesses of all of the methods considered in this research based on the results obtained and give recommendations for the more self–consistent and reliable Lagrangian coherent structure detection methods of those considered in this study. There may be another method that gives a more complete picture of a flow system in a more efficient manner less susceptible to poor velocity data quality and resolution, however we give recommendations from these methods as it is impossible to consider all Lagrangian coherent structure detection methods in a study of this calibre. Bear in mind that we shall not rank the methods considered from worst to best (for instance), instead we will endeavour to recommend the most ideal methods to consider first for accurate, reliable and efficient Lagrangian coherent structure detection within any type of flow system.

Chapter 2

Lagrangian Coherent Structure Detection Methods

In this Chapter, we discuss in detail the eight Lagrangian coherent structure (LCS) detection methods to feature in this Thesis, demonstrate what the structures produced from each method represent, and discuss in full the algorithms and steps involved in detecting these structures numerically using MATLAB. All of these methods are concerned with finding consistent flow patterns within a velocity system defined by the non–autonomous system of ordinary differential equations

$$\dot{\mathbf{x}} = \mathbf{u}\left(\mathbf{x}, t\right), \tag{2.1}$$

where $\mathbf{x} = (x, y)$, defined on a flow domain $\Omega \in \mathbb{R}^2$, $\mathbf{u}$ represents Eulerian velocity, and we seek coherence in a Lagrangian sense, that is, by following the flow of (2.1) over a pre–determined time interval. We define the flow map, or a Lagrangian flow trajectory, $\mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right)$, as the solution generated from advecting an initial point $\mathbf{x}_0 = (x_0, y_0)$ through the flow system (2.1) over the flow interval ranging from t_0 to $t_0 + T$. Various metrics of the flow maps $\mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right)$ defined at different initial points $\mathbf{x}_0$, such as the gradient (or Jacobian) of the flow map $\nabla\mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right)$, help determine the shapes and locations of LCSs within the flow system defined by (2.1). Throughout this study, the flow map of (2.1) is computed numerically in MATLAB for a range of initial points $\mathbf{x}_0$ using `ode45`, a function which solves systems of ordinary differential equations using a Runge–Kutta solver of order (4,5). To estimate the velocity for (2.1) as we generate these flow maps, we either use analytically defined velocity equations, or we interpolate numerically defined velocity data in space and time using MATLAB's `griddedInterpolant` function, with a cubic spline method of interpolation chosen for more accurate velocity estimation. How each of the initial points are seeded is dependent upon the computational requirements of each individual method.

For conciseness through this Chapter (and this Thesis), we will introduce symbols and abbreviations for the most important concepts featured throughout this Thesis, in

Table 2.1: Acronyms, abbreviations and symbols first defined in Chapter 2.

Symbol	Represents	Equation(s)	Defined in/at
LCS	Lagrangian Coherent Structure	N/A	Start of Chapter
FTLE	Finite Time Lyapunov Exponent	(2.3)	Section 2.1
VLCS	Variational LCS	N/A	Section 2.2
VLCS–A	Variational LCS Algorithm A	N/A	Section 2.2.2
VLCS–B	Variational LCS Algorithm B	N/A	Section 2.2.3
VLCS–C	Variational LCS Algorithm C	N/A	Section 2.2.4
LAVD	Lagrangian Averaged Vorticity Deviation	(2.6)	Section 2.3
S^2	Stochastic Sensitivity	(2.9)	Section 2.4
Scaled S^2	S^2 Scaled by Velocity Uncertainty	(2.14)	Section 2.4
$R(L, v_r, h)$	Robust Set Derived Using S^2	(2.15)	Section 2.4
$\mathbf{P}$	Transfer Operator (Finite Approximation)	(2.16),(2.17)	Section 2.5
$\widetilde{\mathbf{P}}$	$\mathbf{P}$ Scaled by Box Measures	(2.19)	Section 2.5
$\boldsymbol{\Delta}^D$	Dynamic Laplace Operator	(2.23),(2.25)	Section 2.6
FCM	Fuzzy C–Means Clusters	(2.27),(2.28)	Section 2.7
CSC	Coherent Structure Colouring	(2.30)	Section 2.8

particular the LCS detection methods used and the flow systems considered. Table 2.1 shows a notation table for concepts first defined in this Chapter.

2.1 The Finite Time Lyapunov Exponent

The finite time Lyapunov exponent (FTLE) is a measure of the largest possible rate of exponential stretching a flow particle is likely to undergo based on the quantifiable change of a flow trajectory when an initial point $\mathbf{x}_0 \in \Omega$ is perturbed over a disk–shaped fluid particle of small radius. Suppose we apply a slight alteration to an initial point $\mathbf{x}_0$ in the form

$$\widehat{\mathbf{x}}_0 = \mathbf{x}_0 + \mathbf{x}_1,$$

where $\|\mathbf{x}_1\|$ is infinitesimally small. Now,

$$\begin{aligned} \mathbf{F}_{t_0}^{t_0+T}\left(\widehat{\mathbf{x}}_0\right) &= \mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0 + \mathbf{x}_1\right) \\ &= \mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right) + \nabla \mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right)\mathbf{x}_1 + \mathcal{O}\left(\|\mathbf{x}_1\|^2\right), \end{aligned}$$

from properties defined in Shadden et al. (2005) and Kent (2008). We use the above expression to define the FTLE as

$$\begin{aligned}\Phi_{t_0}^{t_0+T}(\mathbf{x}_0) &= \frac{1}{|T|}\sup_{\mathbf{x}_1}\ln\frac{\left\|\mathbf{F}_{t_0}^{t_0+T}(\widehat{\mathbf{x}}_0)-\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)\right\|}{\|\mathbf{x}_1\|}\\ &= \frac{1}{|T|}\ln\left\|\nabla\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)\right\|.\end{aligned}$$

After some derivation (Kent 2008), the norm of the flow map gradient

$$\left\|\nabla\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)\right\| = \sqrt{\lambda_{\max}(\mathbf{x}_0)},$$

where $\lambda_{\max}(\mathbf{x}_0)$ is the largest eigenvalue of the Cauchy–Green strain tensor

$$\mathbf{C}_{t_0}^{t_0+T}(\mathbf{x}_0) = \left(\nabla\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)\right)^{\top}\nabla\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0) \tag{2.2}$$

at the point $\mathbf{x}_0$. This provides us with a more computationally friendly definition of the FTLE

$$\Phi_{t_0}^{t_0+T}(\mathbf{x}_0) = \frac{1}{|T|}\ln\sqrt{\lambda_{\max}(\mathbf{x}_0)}. \tag{2.3}$$

which we can use to compute the maximal stretching rate of a flow particle beginning at the point $\mathbf{x}_0$ which moves through the flow system defined by (2.1) (Shadden et al. 2005, Kent 2008). Upon inspecting the scalar field of (2.3), regions of the domain corresponding to low values of (2.3) exhibit more robust flow behaviour, while continuous ridges which correspond to the largest values of (2.3) are defined as codimension–1 LCSs of maximal particle repulsion. This definition of a maximal FTLE ridge is based off simple, analytically defined flow models with an infinite flow time interval, and providing a similar definition for a finite time flow is a little more ambiguous (Balasuriya et al. 2016). Flow trajectories repel these ridges at a maximal rate producing a hypothetical barrier which divides the flow domain Ω into more coherent sub-regions. This is the case for $T > 0$, however if $T < 0$ the maximal ridges of (2.3) correspond to LCSs of maximal particle attraction, around which clusters of flow trajectories compress until they gradually form the shape of these ridges (Shadden et al. 2005).

As (2.3) is relatively easy to evaluate numerically, the FTLE has become a popular method for detecting LCSs within a wide variety of realistic flow systems. The method has primarily been used to model fluid, particle and debris transport in oceanic flow systems (Leclair et al. 2020, Suara et al. 2020), atmospheric flow systems (Garaboa-Paz et al. 2017) or even through arterial blood flow (Teeraratkul et al. 2021). The FTLE has also been used to visualise coherent flow structures within simulated velocity systems modelled to contain flow behaviour not unlike that found in real world systems, either using computational fluid dynamics (Rockwood et al. 2018) or laboratory experiments (Lin et al. 2020). On the other hand, some researchers have expressed concern over the reliability of this method, most notably Haller (2011) who demonstrated a few examples of the FTLE producing misleading results within certain flow systems. Most of these flow systems were analytically defined toy models, including an autonomous saddle flow where the forward time FTLE was evaluated to be constant over the whole domain and the backward time FTLE identified the x axis of the domain (an unstable manifold) as a trough of low particle stretching rather than as a ridge of high particle stretching. Balasuriya et al. (2016) were able to justify why these degenerate results are obtained for this flow through the definition of hyperbolic neighbourhoods, or sub–regions of a domain within which greater particle stretching is observable. This analysis, along with the relatively frequent use of this method by other researchers, suggest that the FTLE otherwise remains a competent and reliable method for use in detecting LCSs within a wide range of flow systems.

To compute (2.3) for flow systems in this study, we define a uniform grid of initial points $\mathbf{x}_0$ within our flow domain Ω and compute the flow map $\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)$ for each of these points. We then calculate $\nabla\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)$ at each grid point using finite differencing techniques, generate the Cauchy–Green strain tensor $\mathbf{C}_{t_0}^{t_0+T}(\mathbf{x}_0)$ and use the largest eigenvalue of this tensor $\lambda_{\max}$ to produce a scalar field for (2.3). In this study, we only consider forward time flows, and are therefore only concerned with detecting LCSs corresponding to ridges of maximal particle stretching rather than maximal particle compression.

2.2 Variational Lagrangian Coherent Structures

The aforementioned degenerate results for the FTLE method prompted Haller (2011) to conceptualise an alternative method for detecting LCSs which correspond to the locally largest levels of particle stretching or compression. This method is referred to as the variational LCS (VLCS) detection method, or sometimes as the geodesic LCS detection method, and is referred to as such because this method defines variational problems for quantities which are used to measure particle stretching or compression capability. The geodesic solutions of these problems are smooth, differentiable curves which represent coherent flow barriers which particles repel from (or attract to) at a locally maximal rate. What constitutes as "local" within the scope of this method will be made clear shortly. This method does carry the additional advantage that the curves produced are "material LCSs", meaning that they can be flowed forwards (or backwards) under the velocity system (2.1) to observe how the shapes of these structures change and evolve away from the initial reference time t_0 (Haller 2011).

VLCSs come in three forms - hyperbolic, parabolic and elliptic. This research will focus on hyperbolic VLCSs only, as this type of structure was the first to be conceptualised and was done so in response to the creation and development of the FTLE method (Haller 2011). In a two–dimensional flow system, the Cauchy–Green strain tensor (2.2) produces two strictly positive eigenvalues λ_1 and λ_2 at each initial point $\mathbf{x}_0 \in \Omega$, where $\lambda_1 \leq \lambda_2$. It

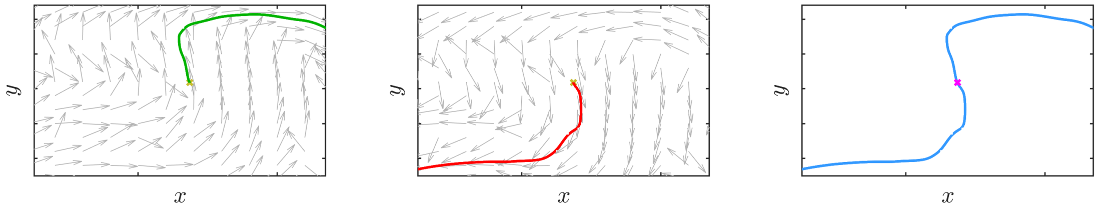

Figure 2.1: Forward (left) and backward (centre) integration of a $\boldsymbol{\xi}_1$ eigenvector field from a point corresponding to a "local maximum" of λ_2 (indicated by a 'x' symbol), which is used to produce a hyperbolic VLCS (right).

is worth noting that λ_2 is the same as $\lambda_{\max}$, the largest eigenvalue of (2.2) used to calculate the FTLE. Two corresponding orthonormal eigenvectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are also obtained, and are referred to as the strain eigenvector and the stretch eigenvector respectively (Farazmand & Haller 2012). The collection of Cauchy–Green eigenvectors at each of these initial points $\mathbf{x}_0$ form eigenvector fields, and solution curves of these fields are used to produce hyperbolic variational LCSs. As illustrated in Figure 2.1, we seed initial integration points for the $\boldsymbol{\xi}_1$ field by obtaining "local maxima" of the flow stretching measure λ_2 and then commence forward and backward integration of the $\boldsymbol{\xi}_1$ field from these points. The combination of the two solution curves produced from this point forms a hyperbolic VLCS of maximal particle repulsion, which should coincide with ridges of λ_2 along which locally maximal flow stretching capability is observed (Haller 2011, Farazmand & Haller 2012). Hyperbolic VLCSs of maximal particle attraction are also attainable, either by producing solution curves of the $\boldsymbol{\xi}_1$ eigenvector field for a backwards time flow commencing from t_0+T and ending at t_0; or solving the stretch eigenvector field $\boldsymbol{\xi}_2$ for the forward time flow from local minima of the other eigenvalue λ_1 (Farazmand & Haller 2012). For the sake of this research, only hyperbolic VLCSs of maximal particle repulsion will be considered.

Even though they will not be considered in this research, it is worth briefly defining the two other types of VLCS which have been conceptualised. Parabolic VLCSs are coherent chains of strain lines (solutions of the $\boldsymbol{\xi}_1$ eigenvector field) and stretch lines (solutions of the $\boldsymbol{\xi}_2$ field) connected by what are known as Cauchy-Green singularity points; which are initial points $\mathbf{x}_0$ at which $\lambda_1 = \lambda_2$. These singularity points come in two forms - trisector type points which break up a collection of strainlines into three distinct zones, and wedge type points which strainlines fold or bend around in a parabolic fashion (Balasuriya & Bollt 2021). To produce parabolic VLCSs, one must solve the $\boldsymbol{\xi}_1$ eigenvector field beginning from trisector type singularity points, solve the $\boldsymbol{\xi}_2$ eigenvector field beginning from wedge type singularity points, and identify chains of curves which connect trisector and wedge type singularity points in an alternating fashion. The resultant parabolic LCSs produced are used to identify centres of coherent streams or jet–type structures which act as more robust flow separators in comparison to their hyperbolic counterparts, as they are typically not positioned within regions of the domain where particle flow is most volatile but rather where greater Lagrangian shear is observable (Farazmand et al. 2014). Elliptic VLCSs are the boundaries of closed, two–dimensional LCSs which correspond to flow objects of strong rotational coherence within a velocity system, such as vortices or oceanic eddies. These structures are generated from vector fields formed from spatially dependent linear combinations of the Cauchy–Green eigenvectors $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ and encapsulate pairs of Cauchy–Green singularity points of the same type (Karrasch et al. 2014, Haller & Beron-Vera 2012, 2013).

The procedure outlined earlier for numerically detecting hyperbolic VLCSs may appear simple to implement, however in reality the steps involved with generating these structures are significantly more complicated due to the dependence of the method on several computational parameters, and the ambiguity involved in deciding which particular eigenvector field solution curves qualify as the LCSs corresponding to the largest "local" particle repulsion capability. This is why several computational algorithms have been outlined in the available literature for producing a complete picture of all of the most relevant hyperbolic VLCSs existent within a flow system. In this research, we explicitly examine two of these published algorithms in Sections 2.2.2 and 2.2.3; and in Section 2.2.4 we propose a third algorithm by using the FTLE as a criterion for selecting initial integration points for the first eigenvector field $\boldsymbol{\xi}_1$. Firstly, we need to demonstrate how we numerically integrate the $\boldsymbol{\xi}_1$ field in each of these algorithms.

2.2.1 Solving the Eigenvector Fields

Numerically solving the Cauchy–Green eigenvector fields to produce hyperbolic VLCSs should be more or less straightforward, but in reality it is not, primarily due to a lack of smoothness and the presence of orientational discontinuities within these fields; observable when pairs of adjacent eigenvectors face the opposite direction to each other along ridges within these fields. This occurs as the Cauchy–Green eigenvectors are defined up to multiplicative constants, so the direction of these eigenvectors is ambiguous. Because computation of the Cauchy–Green strain tensor is highly sensitive to numerical errors, particularly within regions of a domain where particle flow is more chaotic (which is problematic as the purpose of this method is to detect consistent flow structures within such regions), these computational inconsistencies are generally unavoidable (Farazmand & Haller 2012). However, there are ways in which we can smoothen out these inconsistencies before and while the relevant eigenvector fields are being solved.

Farazmand & Haller (2012) suggest two computational techniques which should both be implemented in order to improve the smoothness of the Cauchy–Green eigenvector fields. The first technique involves defining a "star–grid" (otherwise known as an auxiliary grid) surrounding each original initial point $\mathbf{x}_0$ - consisting of the points $\mathbf{x}_r = (x_0 + \delta_x, y_0)$, $\mathbf{x}_l = (x_0 - \delta_x, y_0)$, $\mathbf{x}_u = (x_0, y_0 + \delta_y)$ and $\mathbf{x}_d = (x_0, y_0 - \delta_y)$; that is, infinitesimally small perturbations of each initial point of size δ_x in the x direction and δ_y in the y direction. We then compute the flow maps of these four new uniform grids of initial points over the

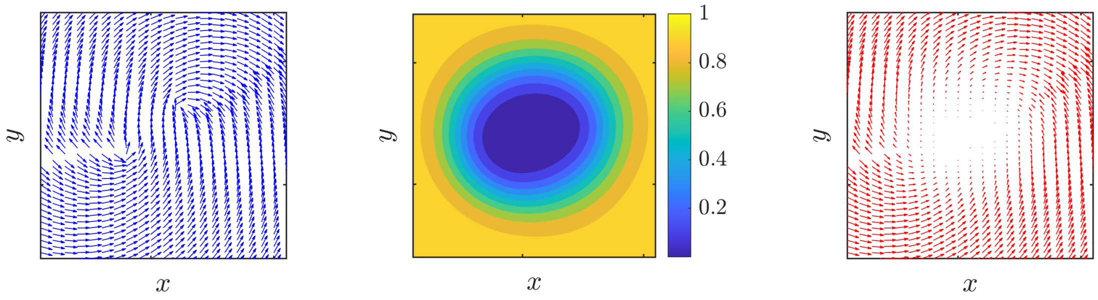

Figure 2.2: An original $\boldsymbol{\xi}_1$ eigenvector field (left), the scalar field for the α parameter (2.4) (centre), and the scaled $\alpha\boldsymbol{\xi}_1$ eigenvector field (right).

same flow interval $[t_0, t_0 + T]$ and use these to generate the flow map gradient using the formula

$$\nabla \mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0) = \begin{bmatrix} \frac{\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_r) - \mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_l)}{2\delta_x} & \frac{\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_u) - \mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_d)}{2\delta_y} \end{bmatrix}.$$

Computing the Cauchy–Green eigenvectors from this definition of the flow map gradient significantly improves the consistency of the orientation of these eigenvectors (Farazmand & Haller 2012). It is advised that the Cauchy–Green eigenvalues should not be computed using the flow map gradient derived from the star grid and instead using the flow map gradient computed from the original uniform grid, because as Lekien & Ross (2010) and Farazmand & Haller (2012) point out, computing the eigenvalues from the star grid gradient definition causes a diffusion in the values for the stretching parameter λ_2, ironing out ridges and key structures of both this quantity and the FTLE and thereby rendering these quantities unreliable for use in locating flow structures corresponding to the largest particle stretching capability.

The second computational technique involves scaling the Cauchy–Green eigenvector fields by the parameter α, defined as

$$\alpha(\mathbf{x}_0) = \left(\frac{\lambda_2(\mathbf{x}_0) - \lambda_1(\mathbf{x}_0)}{\lambda_2(\mathbf{x}_0) + \lambda_1(\mathbf{x}_0)} \right)^2. \tag{2.4}$$

At Cauchy–Green singularity points (where $\lambda_1 = \lambda_2$), the α quantity defined above will equate to 0 thereby eliminating degenerate Cauchy–Green eigenvectors at these points, as we illustrate in Figure 2.2. This is not the only way to deal with this issue, with additional methods defined in Balasuriya & Bollt (2021), and later on in Section 2.2.4 of this Thesis.

These techniques do not remove ridges within the Cauchy–Green eigenvector fields formed by orientational discontinuities between these eigenvectors. This issue can only be resolved by reorientating the eigenvectors as integration proceeds in a manner illustrated in Figure 2.3. As we integrate these eigenvector fields to produce solution curves, we interpolate new eigenvectors along the lengths of these curves. As we do this, the orientation of the interpolated eigenvector must be compared with the orientations of the

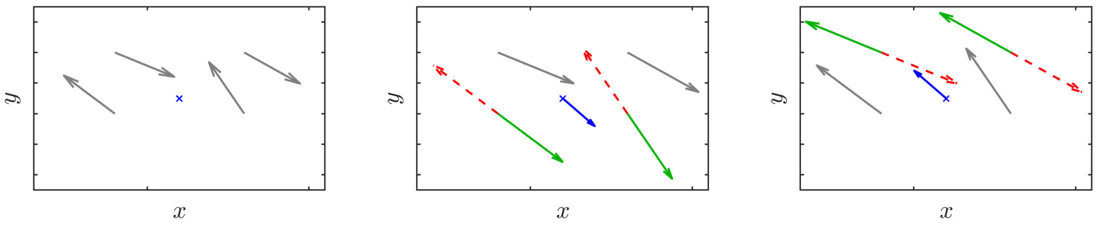

Figure 2.3: A discontinuous grid of four $\boldsymbol{\xi}_1$ eigenvectors along with an interpolation query point indicated by a blue 'x' (left) and reorientation of the discontinuous vectors as integration of the $\boldsymbol{\xi}_1$ field proceeds (centre or right). Red vectors with dashed lines indicate discontinuous original eigenvectors, while the green solid vectors have been reorientated in the same direction in which the hyperbolic strainline is traversing the field to produce the blue interpolated eigenvector.

four surrounding eigenvectors produced from the original grid of initial flow points. This can be done simply by taking the dot product of the interpolated eigenvector with each of the four surrounding grid eigenvectors. If the orientation of the interpolated eigenvector is consistent with all four of the grid eigenvectors then no action is required; however if any of the four grid eigenvectors are pointing in a different direction to the interpolated eigenvector then these grid eigenvectors must be rotated by 180 degrees to match the orientation of the interpolated eigenvector. This will allow the solution curves produced to cross these ridges of orientational discontinuity and for integration of the eigenvector fields to continue to produce smooth solution curves that do not reverse direction or warp in shape (Farazmand & Haller 2012).

Now that we have worked out how to numerically generate these hyperbolic strainlines (candidate VLCSs) from the $\boldsymbol{\xi}_1$ eigenvector field, we must now isolate from the full foliation of these strainlines hyperbolic VLCSs of "locally maximal" particle repulsion capability. We will describe three such algorithms for regulating integration of the eigenvector fields and identifying the most locally repelling strainlines, including an algorithm which identifies hyperbolic LCSs using the theory of VLCSs in Section 2.2.2, a faster algorithm which pre–determines initial integration points from "local maxima" of λ_2 in Section 2.2.3 and, our own algorithm, a combination of this method with the FTLE to identify the most influential hyperbolic VLCSs over the whole domain without overcrowding it with more irrelevant strainlines in Section 2.2.4.

2.2.2 VLCS Algorithm A

One of the first algorithms produced to numerically execute this method was conceptualised by Farazmand & Haller (2012), and was designed to generate hyperbolic VLCSs which are explicitly consistent with the underlying variational theory pertinent to these LCSs. The aim of this algorithm, which we will refer to as VLCS Algorithm A or VLCS–A, is to generate curves representing hyperbolic VLCSs along which the following necessary conditions are met for all points $\mathbf{x}$ along the curve:

1. $\lambda_1(\mathbf{x}) \neq \lambda_2(\mathbf{x})$,
2. $\lambda_2(\mathbf{x}) > 1$,
3. $\langle \boldsymbol{\xi}_2(\mathbf{x}), \mathbf{H}(\lambda_2(\mathbf{x}))\, \boldsymbol{\xi}_2(\mathbf{x}) \rangle \leq 0$, where $\langle \cdot, \cdot \rangle$ represents the dot product and $\mathbf{H}$ symbolises the Hessian, or second derivative matrix of λ_2 (Olascoaga & Haller 2012);
4. $\boldsymbol{\xi}_2(\mathbf{x})$ is always perpendicular to the structure; or equivalently as the Cauchy–Green eigenvectors are orthonormal, $\boldsymbol{\xi}_1(\mathbf{x})$ is always parallel to the structure; and
5. The average value of λ_2 along the structure, $\overline{\lambda_2}$, is maximal with respect to all immediately neighbouring candidate structures.

The first two conditions ensure that any hyperbolic VLCS produced does not traverse through any Cauchy–Green singularity points, at which these conditions are violated, whilst the third condition is an optimisation condition for the particle repulsion parameter λ_2 which ensures that this quantity is maximal within portions of the domain along which these hyperbolic VLCSs are located. The fourth condition is a direction to solve the $\boldsymbol{\xi}_1$ eigenvector field to produce strainlines (which we will refer to by the symbol γ) which act as candidate hyperbolic VLCSs, and the fifth and final condition is used to isolate the structures of locally maximal particle repulsion capability by ensuring that the particle stretching measure λ_2 is largest along a strainline in comparison to its nearest neighbours (Farazmand & Haller 2012).

The steps taken here in generating hyperbolic VLCSs using this algorithm are as follows. We begin by computing the flow map $\mathbf{F}_{t_0}^{t_0+T}$, its gradient and the Cauchy–Green strain tensor (2.2) for a set of initial points $\mathbf{x}_0 \in \Omega$ which have once again been organised within a uniform grid. Next, using the second Cauchy–Green eigenvalues λ_2 and eigenvectors $\boldsymbol{\xi}_2$ corresponding to each of these initial points, we evaluate conditions 1–3 above for each of these initial points. Points which satisfy all three of these conditions within an acceptable enough tolerance form an "acceptable set" of initial points used to integrate the first Cauchy–Green eigenvector field, which will be referred to here as $\mathcal{G}_0$.

The next step would then be to begin integrating the $\boldsymbol{\xi}_1$ eigenvector field from each of the initial points which form the acceptable set $\mathcal{G}_0$. However, seeing as $\mathcal{G}_0$ will often contain a very large amount of points $\mathbf{x}_0$, and generating strainlines from all of these candidate points would be extraordinarily time consuming, we subset a smaller partition of $\mathcal{G}_0$ and solve the $\boldsymbol{\xi}_1$ eigenvector field from these points only. Farazmand & Haller (2012) suggest defining a set of equally spaced horizontal and vertical lines within the domain Ω and solving the $\boldsymbol{\xi}_1$ field from candidate points which lie along these lines. However, to eliminate this extra computational step, we instead randomly seed around 500–2000

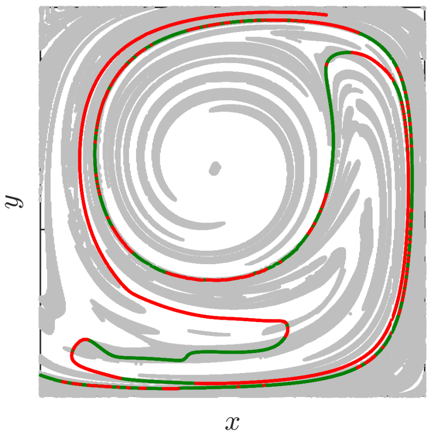

Figure 2.4: The acceptable set $\mathcal{G}_0$ (area shaded grey), and a hyperbolic strainline coloured green for line segments inside $\mathcal{G}_0$ and red for segments outside $\mathcal{G}_0$. The strainline enters and exits $\mathcal{G}_0$ several times as it is being generated, however when the long red line segment of arc length larger than $\mathcal{L}_f$ is produced, integration of the $\boldsymbol{\xi}_1$ field must stop.

candidate points (depending on the system and other factors, e.g. the size of Ω) from $\mathcal{G}_0$. To produce a hyperbolic strainline γ, we integrate the $\boldsymbol{\xi}_1$ field beginning from each candidate point in both forward and backward time to produce a full, cohesive material curve.

We then proceed to solve the $\boldsymbol{\xi}_1$ field from the first candidate point using the field integration techniques detailed in Section 2.2.1. As integration of this field proceeds, checks are performed at each new point along the strainline to ensure that this solution curve remains within the region defined by the acceptable set $\mathcal{G}_0$. Because of numerical errors associated with numerical differentiation and integration; as well as numerical uncertainty expected to be found within numerically defined velocity data, Farazmand & Haller (2012) allow some tolerance when performing these checks. If the conditions of the acceptable set $\mathcal{G}_0$ are found to have been violated, we begin to measure the length of the segment of the strainline which sits outside $\mathcal{G}_0$. If this length exceeds an allowance parameter $\mathcal{L}_f$, we stop integrating the eigenvector field from this candidate point and move on to the next candidate point. If the strainline re-enters $\mathcal{G}_0$ before the parameter $\mathcal{L}_f$ has been exceeded, the length measure may be stricken and integration of the $\boldsymbol{\xi}_1$ field can continue. Figure 2.4 illustrates how integration of the $\boldsymbol{\xi}_1$ field is regulated using $\mathcal{G}_0$.

Completion of this procedure results in the creation of a hyperbolic strainline $\gamma = \gamma(s)$, where s is discretely defined over $N+1$ points along γ, and ranges in value from 0 to what is assumed to be $N\Delta s$, where Δs is the integration time step used to produce γ (Farazmand & Haller 2012). Once a strainline γ has been generated from each seeded candidate point, we single out the hyperbolic VLCSs by first ensuring that each line has an arc length $\mathcal{L}(\gamma)$ of at least $\mathcal{L}_{\min}$, and then by comparing the average value of λ_2, $\overline{\lambda_2}$, of each of these strainlines with their immediate neighbours. The way in which Farazmand & Haller (2012) achieve this involves evaluating $\overline{\lambda_2}$ using the formula

$$\overline{\lambda_2}\left(\gamma\right) = \frac{1}{\mathcal{L}\left(\gamma\right)} \int\limits_{\gamma} \lambda_2\left(\gamma\right) \mathrm{ds},$$

locating the two nearest neighbouring strainlines for each γ (using their corresponding candidate points) and comparing the generated values of $\overline{\lambda_2}$ for this set of lines. If the value of $\overline{\lambda_2}$ for a strainline γ is larger than its two nearest neighbours, this strainline is classified as a hyperbolic VLCS of maximal particle repulsion which satisfies the five necessary conditions listed earlier. The collection of all of these VLCSs forms a foliation of the most repelling barrier–type structures of the flow domain Ω, which clusters of flow trajectories will repel from at a locally maximal rate.

2.2.3 VLCS Algorithm B

While the previous algorithm conforms strongly to the theory of VLCSs, executing the method requires many algorithmic steps and is computationally inefficient to implement. In order to reduce the computational time and effort required to produce these hyperbolic VLCSs, Onu et al. (2015) proposed a new, more simple algorithm for generating these structures. Conceptualised under the name "LCS Tool", the intuition behind this algorithm involves the notion that hyperbolic variational Lagrangian coherent structures are curves which represent flow barriers of maximal particle repulsion capability, but this capability is measured quantitatively by the second Cauchy–Green eigenvalue λ_2. Intuitively, any hyperbolic variational LCSs produced should capture ridges where λ_2 is largest (or at least locally largest) within the domain, so by choosing initial flow points $\mathbf{x}_0$ corresponding to "locally maximal" values of λ_2 as candidate points for solving the $\boldsymbol{\xi}_1$ eigenvector field, hyperbolic VLCSs like those detected from VLCS–A can be detected in a more timely manner. What constitutes a "locally maximum" point of λ_2 as far as Onu et al. (2015) are concerned will be explained shortly.

Unlike the previous algorithm, the LCS Tool algorithm, which we will refer to as VLCS Algorithm B or VLCS–B, relies on just two parameters - a separation radius ρ, and a maximum length for the hyperbolic variational LCSs $\mathcal{L}_{\max}$. The algorithm begins once again by seeding candidate points for integration of the $\boldsymbol{\xi}_1$ field, which is done in this case by identifying initial flow points $\mathbf{x}_0$ which correspond to "locally maximal" values of λ_2. What is meant by "locally maximal" values of λ_2 is as follows: we begin by identifying the initial point $\mathbf{x}_0$ which corresponds to the globally largest value of λ_2 over the whole flow domain Ω. We then draw a circle of radius ρ around this point, and identify the point $\mathbf{x}_0$ corresponding to the global maximum of λ_2 within the remainder of the domain not covered by this circle. We continue in this fashion until no remaining grid points $\mathbf{x}_0$

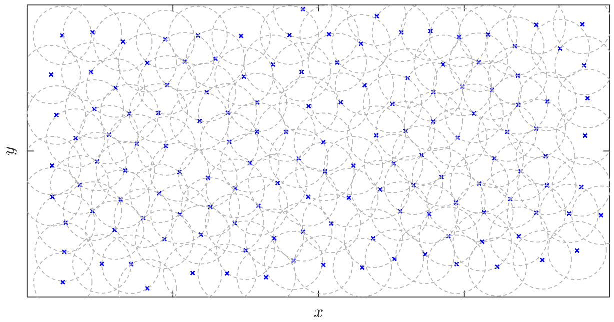

Figure 2.5: Candidate points for integration of the $\boldsymbol{\xi}_1$ field (blue 'x' markers), separated by circular neighbourhoods of radius ρ (indicated by grey dashed lines).

lie beyond a distance of ρ from a previously determined candidate point. This procedure produces an image of candidate points much like the one shown in Figure 2.5.

We then perform forward and backward integration of the $\boldsymbol{\xi}_1$ field from the first point of "locally maximal" λ_2 (that is, the point corresponding to the globally maximal value of λ_2) using the same procedure outlined in Section 2.2.1. This time, we do not regulate this integration using any of the necessary conditions from the previous algorithm or an acceptable set such as $\mathcal{G}_0$, instead integration of the eigenvector field proceeds until a full strainline of length no larger than $\mathcal{L}_{\max}$ has been generated, which is then automatically classified as a hyperbolic VLCS of locally maximal particle repulsion. Integration of the $\boldsymbol{\xi}_1$ field then proceeds from the next candidate point so long as this point does not lie within ρ of the LCS previously detected. Figure 2.6 shows the collection of eligible remaining candidate points from those shown in Figure 2.5 after the first hyperbolic VLCS has been produced. This procedure continues until all candidate points lie within ρ of a hyperbolic VLCS, producing a foliation of these VLCSs which more or less covers the flow domain.

2.2.4 VLCS Algorithm C

The idea of the VLCS–B algorithm is to capture hyperbolic VLCSs which coincide with maximal ridges of λ_2, a measure of the maximal stretching rate of a flow trajectory. However, as noted by Farazmand & Haller (2012), λ_2 is not a smooth measure of particle stretching, with sharp jumps and large variation in this quantity observable throughout the domain. The FTLE (2.3) is a scaled version of λ_2 (the larger of the two Cauchy–Green eigenvalues) used to visualise maximal ridges which represent hyperbolic flow barriers for a dynamic system. This suggests that if the variational LCSs produced using VLCS–B are designed to coincide with maximal ridges of the λ_2 field, these should also coincide with maximal ridges of the FTLE and this quantity could therefore be used to regulate the integration of the $\boldsymbol{\xi}_1$ field instead of λ_2. Hence, we propose a new VLCS detection algorithm which essentially combines the computational metrics of the previous two al-

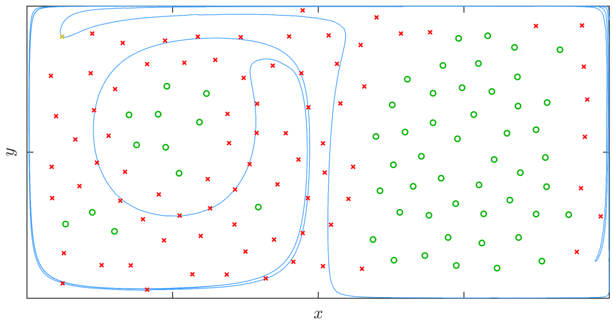

Figure 2.6: The candidate points from Figure 2.5 shown with the hyperbolic VLCS produced from the point labelled with a gold 'x' marker. Points marked with red crosses are within ρ of the VLCS and cannot be selected as the next initial integration point, whereas the points marked with green circles are distanced further than ρ from the VLCS and the point corresponding to the largest value of λ_2 among these remaining points is selected as the next initial integration point.

gorithms with those of the FTLE to find the most influential hyperbolic flow barriers for a dynamic system, essentially creating synergy between a collection of LCS detection methods/algorithms which were all designed to detect the same type of coherent structure.

This algorithm, which we will refer to as VLCS Algorithm C or VLCS–C, follows a similar procedure to VLCS–B, except the FTLE is used to regulate the selection of candidate points for integrating the $\boldsymbol{\xi}_1$ eigenvector field. In addition, the α parameter used to scale the $\boldsymbol{\xi}_1$ field to remove degenerate eigenvectors (see Section 2.2.1) is replaced with a similar quantity,

$$\beta\left(\mathbf{x}_0\right)=\frac{1}{\sigma\left(\mathbf{x}_0\right)}\Phi_{t_0}^{t_0+T}\left(\mathbf{x}_0\right), \tag{2.5}$$

where

$$\sigma\left(\mathbf{x}_0\right)=\max_{\mathbf{x}_0}\Phi_{t_0}^{t_0+T}\left(\mathbf{x}_0\right);$$

that is, the maximum of the FTLE defined discretely over all initial points $\mathbf{x}_0$. We find this maximum in a similar fashion to how Onu et al. (2015) isolate "local maxima" of λ_2, as the FTLE is not defined continuously over the entire flow domain Ω. The scaling parameter $\beta\in[0,1]$, amplifies the $\boldsymbol{\xi}_1$ eigenvectors along ridges where the FTLE is larger, and diminishes these eigenvectors where the FTLE is smaller and a Cauchy–Green singularity point is more likely to exist. Furthermore, in an incompressible flow $\lambda_{\max}=1$ at these singularity points and the FTLE vanishes, causing β to vanish at these points along with the degenerate Cauchy–Green eigenvectors as intended by the original α parameter. An illustration of how this parameter scales the $\boldsymbol{\xi}_1$ eigenvectors can be seen in Figure 2.7. In this algorithm, the β parameter is used to scale the $\boldsymbol{\xi}_1$ field, seed candidate points used to find solution curves for this field, and regulate this process by

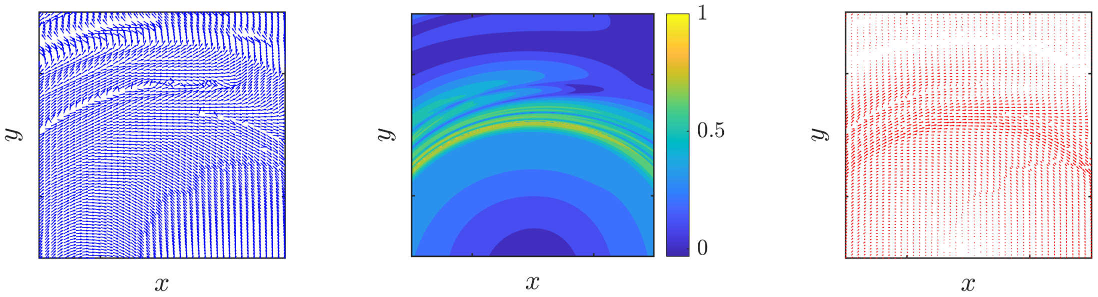

Figure 2.7: An original $\boldsymbol{\xi}_1$ eigenvector field (left), the scalar field for the β parameter (2.5) (centre), and the scaled $\beta\boldsymbol{\xi}_1$ eigenvector field (right).

ensuring that the strainlines produced do not traverse into regions where the FTLE is below a pre–determined threshold β_0.

This algorithm begins by isolating the initial point $\mathbf{x}_0$ at which $\beta = 1$ (and hence the FTLE is globally maximal) and commencing forward and backward integration of the $\boldsymbol{\xi}_1$ eigenvector field once again using the procedure detailed in Section 2.2.1, with the β parameter used in place of the α parameter. As integration proceeds, we ensure that the value of β along the strainline does not drop below the limiting threshold β_0. Throughout this research we choose $\beta_0 = \overline{\beta}$, the mean of β over the whole domain Ω, which is the simplest possible choice and leaves scope for determining other potential values for this threshold. If a strainline traverses through a sub-region of the domain where $\beta < \beta_0$ for a length greater than $\mathcal{L}_f$ (using a similar technique to that used in VLCS–A), we terminate integration of the strainline. Otherwise, we let the integration of the $\boldsymbol{\xi}_1$ field continue until the length of the strainline exceeds $\mathcal{L}_{\max}$ (in similar fashion to VLCS–B). The next candidate point is then seeded as the initial point $\mathbf{x}_0$ corresponding to the globally largest value of β that does not lie within ρ of any previously generated strainlines. This procedure continues until all initial points $\mathbf{x}_0$ are either within ρ of a hyperbolic strainline or correspond to a value of β which lies below the threshold β_0. The resultant strainlines represent a foliation of maximally repelling hyperbolic VLCSs for the flow system.

This method essentially blends geodesic LCS detection techniques from the previous two algorithms with the FTLE method to produce a collection of the most meaningful hyperbolic VLCSs within a flow system in a computationally efficient manner.

2.3 Lagrangian Averaged Vorticity Deviation

So far, we have considered Lagrangian coherent structure detection methods which seek codimension–1 flow barrier type structures which indicate where flow trajectory stretching is strongest within a flow domain and therefore segregate the domain into more robust sub–regions. There also exist methods used to detect these robust regions themselves, rather than simply detecting the boundaries which separate them from other coherent regions of the domain. One such method uses Lagrangian averaged vorticity deviation (LAVD), a measure of the rotational coherence of pockets of flow trajectories based on the relative vorticity of these trajectories against the average vorticity of the whole domain.

The theory of Lagrangian averaged vorticity deviation was first derived by Haller (2016) and expanded and tested further by Haller et al. (2016). To begin, we will rewrite the velocity system (2.1) in the form

$$\dot{\mathbf{x}} = \mathbf{u}\left(x, y, t\right) = \begin{bmatrix} u\left(x, y, t\right) \\ v\left(x, y, t\right) \end{bmatrix}.$$

Upon computing the flow map $\mathbf{F}_{t_0}^{t_0+T}\left(x_0, y_0\right)$ of an initial point $\left(x_0, y_0\right) \in \Omega$, Haller (2016) shows that the gradient of this flow map can be decomposed in the fashion

$$\nabla\mathbf{F}_{t_0}^{t_0+T}\left(x_0, y_0\right) = \mathbf{H}_{t_0}^{t_0+T}\left(x_0, y_0\right)\ \mathbf{D}_{t_0}^{t_0+T}\left(x_0, y_0\right),$$

where $\mathbf{H}_{t_0}^{t_0+T}\left(x_0, y_0\right)$ represents the gradient of the rotational flow component of (2.1), and $\mathbf{D}_{t_0}^{t_0+T}\left(x_0, y_0\right)$ represents the gradient of the stretching/compression flow component. The rotational flow component gradient can be decomposed further in the fashion

$$\mathbf{H}_{t_0}^{t_0+T}\left(x_0, y_0\right) = \boldsymbol{\Theta}_{t_0}^{t_0+T}\left(x_0, y_0\right)\ \boldsymbol{\Upsilon}_{t_0}^{t_0+T}\left(x_0, y_0\right),$$

where $\boldsymbol{\Theta}_{t_0}^{t_0+T}\left(x_0, y_0\right)$ represents a homogeneous rotation gradient and $\boldsymbol{\Upsilon}_{t_0}^{t_0+T}\left(x_0, y_0\right)$ represents the deviation from this homogeneous rotation under the influence of (2.1). The rotation angle of $\boldsymbol{\Theta}_{t_0}^{t_0+T}\left(x_0, y_0\right)$ is used as an objective, frame independent measure of the rotational coherence of the flow trajectory $\mathbf{F}_{t_0}^{t_0+T}\left(x_0, y_0\right)$, and as shown by Haller et al. (2016) is equal to half of the Lagrangian averaged vorticity deviation of said trajectory,

$$\Theta_{t_0}^{t_0+T}\left(x_0, y_0\right) = \int_{t_0}^{t_0+T} \left|\omega_3\left(\mathbf{F}_{t_0}^{s}\left(x_0, y_0\right), s\right) - \varpi\left(s\right)\right|\ \mathrm{ds}. \tag{2.6}$$

In the above expression, ω_3 represents the third component of the vorticity vector $\boldsymbol{\omega} = \nabla \times \mathbf{u}$, which is the only non-zero component of this vector in a two–dimensional flow system such as (2.1) and carries a value

$$\omega_3 = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y};$$

while ϖ represents the mean vorticity of (2.1) over the entire domain Ω, which is calculated using the formula

$$\varpi\left(t\right) = \frac{1}{A\left(\Omega\right)} \iint_{\Omega} \omega_3\left(x, y, t\right)\ \mathrm{dA},$$

where

$$A\left(\Omega\right) = \iint\limits_{\Omega} 1 \ \mathrm{dA},$$

that is, the area of the domain. In short, (2.6) is a relative measure of the vorticity of a flow trajectory against the mean vorticity of the whole domain, with structures or shapes corresponding to larger values of this quantity exhibiting greater rotational coherence within the flow, in a fashion akin to vortical flow structures such as oceanic eddies Haller et al. (2016).

To calculate the LAVD (2.6), we begin by producing a uniform grid of initial points $\mathbf{x}_0 \in \Omega$ in the same fashion as we have done for the previous two LCS detection methods. We can then proceed by computing the flow map $\mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right)$ for each of these points, retaining flow map information at various time steps $t \in [t_0, t_0 + T]$, and using this information to evaluate the integral (2.6) using a numerical integration method such as the Trapezoidal Formula or Simpson's Formula. However, Haller et al. (2016) note that using this method to calculate (2.6) creates a considerable amount of numerical error and uncertainty for this quantity which can be avoided by solving an extended version of the velocity system (2.1) which takes the form

$$\dot{\mathbf{x}} = \begin{bmatrix} u\left(x,y,t\right) \\ v\left(x,y,t\right) \\ \theta\left(x,y,t\right) \end{bmatrix}, \tag{2.7}$$

where $\theta\left(x,y,t\right) = \left|\omega_3\left(\mathbf{F}_{t_0}^{t}\left(x_0,y_0\right),t\right) - \overline{\omega}\left(t\right)\right|$, that is, the integrand of (2.6). Solving the system (2.7) using a numerical differential equation solver such as ode45 in MATLAB should reduce the estimation error accrued for the LAVD to a similar level as that obtained for the flow maps of the initial points $\mathbf{x}_0$ (Haller et al. 2016).

2.4 Stochastic Sensitivity

The next method we will focus on in this research concerns the use of stochastic sensitivity (S^2) to detect LCSs for a flow system essentially using the variance of Lagrangian flow trajectories. As will be explained in due course, this is one of few LCS detection methods with the flexibility to detect both codimension–1 flow barriers as well as more robust flow regions based on certain characteristics and parameters pertinent to a dynamic velocity system. Given the same deterministic velocity system defined by (2.1), we can convert this system into a stochastic differential equation (SDE) system taking the form

$$\mathrm{d}\mathbf{x}_t = \mathbf{u}\left(\mathbf{x}_t,t\right)\ \mathrm{d}t + \epsilon\ \boldsymbol{\sigma}\left(\mathbf{x}_t,t\right)\ \mathrm{d}\mathbf{W}_t. \tag{2.8}$$

In the above equation, the drift function $\mathbf{u}$ represents the same deterministic velocity system defined by (2.1), ϵ is a parameter representing the known or anticipated level of noise or uncertainty present within our Eulerian velocity representing (2.1), $\boldsymbol{\sigma}$ is a matrix which represents the diffusion of stochastic volatility through the flow map in each spatial direction, and $\mathbf{W}_t \sim \mathrm{N}(0,1)$ represents two–dimensional geometric Brownian motion which has a standard normal distribution. It is important to note that each of these quantities carry relevant units of measurement, particularly when we are considering a dimensional velocity system. In the context of this method, the drift velocity $\mathbf{u}$ has a unit of distance per unit of time, ϵ has a unit of distance per square root of time, the Brownian motion $\mathbf{W}_t$ has a unit of the square root of time and the diffusion matrix $\boldsymbol{\sigma}$ is dimensionless and by default set to the 2×2 identity matrix $\mathbf{I}_{2\times 2}$ (Balasuriya 2020*a*,*b*). If we let $\mathbf{x}_T$ be the solution to (2.8) for the initial point $\mathbf{x}_0$ over a flow interval ranging from t_0 to $t_0 + T$, define the quantity

$$Z_\epsilon(\mathbf{x}_0) = \frac{\mathbf{x}_T - \mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)}{\epsilon},$$

where $\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)$ is the "true" deterministic solution of (2.1). Essentially, $Z_\epsilon(\mathbf{x}_0)$ represents the deviation between the deterministic flow map beginning from $\mathbf{x}_0$ and a stochastic realisation of this flow map produced from (2.8). We then project this deviation over all possible angles $\phi \in [-\pi/2, \pi/2)$ using the formula

$$P_\epsilon(\mathbf{x}_0, \phi) = \begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix}^\top Z_\epsilon(\mathbf{x}_0).$$

As the noise level ϵ tends to 0, one might expect that $Z_\epsilon(\mathbf{x}_0)$ (and therefore also $P_\epsilon(\mathbf{x}_0, \phi)$ regardless of the projection angle ϕ) will tend to 0 and that a stochastic solution of (2.8) will tend towards the "true" deterministic solution of (2.1) (Balasuriya 2020*a*). However, seeing as $\mathbf{W}_t$ is not bounded, individual realisations of $\mathbf{x}_T$ (and hence $Z_\epsilon(\mathbf{x}_0)$) can also be unbounded. We can therefore show, as Balasuriya (2020*a*) has done, that the expectation of $Z_\epsilon(\mathbf{x}_0)$ as ϵ tends to zero, is zero. What we are interested in is the variation of this deviation in the flow map, particularly of the quantity $P_\epsilon(\mathbf{x}_0, \phi)$. The quantity

$$A(\mathbf{x}_0, \phi) = \left[\lim_{\epsilon \to 0} \mathrm{Var}\left(P_\epsilon(\mathbf{x}_0, \phi)\right)\right]^{1/2},$$

which is known as the Anisotropic Uncertainty of a flow trajectory, is the expected uncertainty or variation of a flow trajectory over a particular projection angle ϕ. If we maximise this uncertainty over all possible projection angles, we obtain what we can essentially classify as the variance of a Lagrangian flow trajectory, the stochastic sensitivity

$$S^2(\mathbf{x}_0) = \lim_{\epsilon \to 0} \max_{\phi \in [-\pi/2, \pi/2)} \mathrm{Var}\left(P_\epsilon(\mathbf{x}_0, \phi)\right) \tag{2.9}$$

(Balasuriya 2020*a*,*b*). Larger values of this quantity indicate a greater susceptibility of Lagrangian flow trajectories beginning from a point $\mathbf{x}_0$ to velocity uncertainty, and therefore the existence of more chaotic flow behaviour within this region of the flow domain. In fact, as shown by Balasuriya (2020*a*) and as will be further illustrated in later Chapters, for most flow systems the greatest stochastic sensitivity of Lagrangian flow trajectories is observable along codimension–1 ridge type structures not unlike those observable from the FTLE. On the other hand, smaller values of this quantity indicate greater resistance of Lagrangian flow trajectories to numerical uncertainty, thereby offering some guarantee of more robust and consistent particle flow behaviour within the sub–regions of Ω which correspond to these low values.

Balasuriya (2020*a*) has obtained an analytical expression for S^2 that does not rely on the computation of hundreds or even thousands of stochastic realisations of $\mathbf{x}_T$. To derive this expression, we must begin by defining the matrix

$$\mathbf{\Lambda}\left(\mathbf{w},t\right) = \exp\left(\int_{t_0}^{t_0+T} \left[\nabla\cdot\mathbf{u}\right]\left(\mathbf{F}_{t_0+T}^{\tau}\left(\mathbf{w}\right),t\right)\,\mathrm{d}\tau\right)\boldsymbol{\sigma}\left(\mathbf{F}_{t_0+T}^{t}\left(\mathbf{w}\right),t\right)^{\top}\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\nabla\mathbf{F}_{t_0+T}^{t}\left(\mathbf{w}\right), \tag{2.10}$$

where $\mathbf{w} = \mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0\right)$ for initial points $\mathbf{x}_0 \in \Omega$ and $\nabla\cdot\mathbf{u}$ represents the divergence of the velocity system (2.1). If we choose the default $\boldsymbol{\sigma} = \mathbf{I}_{2\times 2}$, (2.10) simplifies to

$$\mathbf{\Lambda}\left(\mathbf{w},t\right) = \exp\left(\int_{t_0}^{t_0+T} \left[\nabla\cdot\mathbf{u}\right]\left(\mathbf{F}_{t_0+T}^{\tau}\left(\mathbf{w}\right),t\right)\,\mathrm{d}\tau\right)\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\nabla\mathbf{F}_{t_0+T}^{t}\left(\mathbf{w}\right). \tag{2.11}$$

Furthermore, if (2.1) is incompressible, i.e. $\nabla\cdot\mathbf{u} = 0$ over the entire domain, (2.11) simplifies even further to

$$\mathbf{\Lambda}\left(\mathbf{w},t\right) = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\nabla\mathbf{F}_{t_0+T}^{t}\left(\mathbf{w}\right), \tag{2.12}$$

which is just a counterclockwise rotation of the backwards time flow map gradient of $\mathbf{w}$ by an angle of $\pi/2$ (Balasuriya 2020*a*,*b*). We then use the elements of $\mathbf{\Lambda}$ to compute the integrals

$$L\left(\mathbf{w}\right) = \frac{1}{2}\int_{t_0}^{t_0+T}\left(\mathbf{\Lambda}_{12}^2 + \mathbf{\Lambda}_{22}^2 - \mathbf{\Lambda}_{11}^2 - \mathbf{\Lambda}_{21}^2\right)\,\mathrm{d}t,$$

$$M\left(\mathbf{w}\right) = \int_{t_0}^{t_0+T}\left(\mathbf{\Lambda}_{11}\mathbf{\Lambda}_{12} + \mathbf{\Lambda}_{21}\mathbf{\Lambda}_{22}\right)\,\mathrm{d}t,$$

and the subsequent quantity $N\left(\mathbf{w}\right) = \sqrt{L^2\left(\mathbf{w}\right) + M^2\left(\mathbf{w}\right)}$. Generating these quantities is essential as the optimal projection angle in (2.9) turns out to be $-\phi_*/2$, where $\cos\left(\phi_*\right) =$

$L(\mathbf{w})/N(\mathbf{w})$ and $\sin(\phi_*) = M(\mathbf{w})/N(\mathbf{w})$ (Balasuriya 2020*a*,*b*). We then complete the procedure by calculating

$$\widetilde{S}^2(\mathbf{w}) = \frac{1}{2}\int_{t_0}^{t_0+T} \left(\mathbf{\Lambda}_{11}^2 + \mathbf{\Lambda}_{21}^2 + \mathbf{\Lambda}_{12}^2 + \mathbf{\Lambda}_{22}^2\right)\, \mathrm{d}t + N(\mathbf{w}), \tag{2.13}$$

and mapping the values of $\widetilde{S}^2(\mathbf{w})$ to their corresponding initial points $\mathbf{x}_0 \in \Omega$ at time t_0 to obtain the stochastic sensitivity $S^2(\mathbf{x}_0)$ (Balasuriya 2020*a*,*b*).

Numerically evaluating (2.13) for the velocity system defined by (2.1) does involve a rather lengthy computational process. We begin as we have with the previously defined LCS detection methods by seeding a uniform grid of initial points $\mathbf{x}_0 \in \Omega$ and flowing these forward under (2.1) to obtain the flow maps $\mathbf{w} = \mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)$ where $\mathbf{w} = (w, z)$. We then need to produce the matrix $\mathbf{\Lambda}$ for each of these advected flow points $\mathbf{w}$ over a range of time steps $t \in [t_0, t_0 + T]$. These computations rely on generating the backwards time flow map gradient for the advected points $\mathbf{w}$, which is challenging as these points are extremely unlikely to be spaced uniformly after the initial grid points have been advected through the flow system defined by (2.1). There are a few ways of overcoming this issue to allow the gradients to be computed using standard finite differencing, and in this research we shall consider two of these methods.

The first method involves seeding a uniform grid of particles $\mathbf{w}_0$ at time $t_0 + T$ which covers the spread of all of the advected points $\mathbf{w}$, flowing these backwards to time t_0 and computing the gradient of the backwards time flow map at various time steps $t \in [t_0, t_0 + T]$ using the same finite differencing techniques employed in some of the previously discussed LCS detection methods such as the FTLE or VLCS methods. The other method involves generating the flow map gradient using an auxiliary star grid defined for each advected point $\mathbf{w}$ in similar fashion to the technique discussed in Section 2.2.1 for the VLCS method, which involves generating the backwards time flow maps for the points $\mathbf{w}_r = (w + \delta_x, z)$, $\mathbf{w}_l = (w - \delta_x, z)$, $\mathbf{w}_u = (w, z + \delta_y)$ and $\mathbf{w}_d = (w, z - \delta_y)$ which surround each $\mathbf{w}$ by a small distance of δ_x in x and δ_y in y; and generating the flow map gradient by computing

$$\nabla \mathbf{F}_{t_0+T}^{t}(\mathbf{w}) = \begin{bmatrix} \frac{\mathbf{F}_{t_0+T}^{t}(\mathbf{w}_r) - \mathbf{F}_{t_0+T}^{t}(\mathbf{w}_l)}{2\delta_x} & \frac{\mathbf{F}_{t_0+T}^{t}(\mathbf{w}_u) - \mathbf{F}_{t_0+T}^{t}(\mathbf{w}_d)}{2\delta_y} \end{bmatrix}.$$

The latter method is recommended when working with observational data sets, or with a velocity system where the spread of advected particles $\mathbf{w}$ is considerably large (Balasuriya 2020*b*). We then use these gradients to compute the elements of the matrix $\mathbf{\Lambda}$, using (2.12) or (2.11) if the flow system in question is compressible. In the latter case, computation of the divergence integral required to evaluate (2.11) can be done numerically using a numerical integration technique such as the Trapezoidal Formula or Simpson's Formula; the former of which will be used in this study.

We then numerically evaluate $L\left(\mathbf{w}\right)$, $M\left(\mathbf{w}\right)$ and the integral in (2.13) using the Trapezoidal Formula (or Simpson's Formula if desired) and use these quantities along with $N\left(\mathbf{w}\right)$ to compute $\widetilde{S}^2$ for the points $\mathbf{w}$ at time t_0+T. If we have used the star grid method to compute the backwards time flow gradients, we complete the computation of S^2 by simply assigning the values of $\widetilde{S}^2\left(\mathbf{w}\right)$ to the corresponding points $\mathbf{x}_0$ at time t_0. However, if we have instead defined a new uniform initial grid $\mathbf{w}_0$ at time t_0+T to calculate these gradients, we assign the values of $\widetilde{S}^2\left(\mathbf{w}_0\right)$ to the points $\mathbf{x}=\mathbf{F}_{t_0+T}^{t_0}\left(\mathbf{w}_0\right)$ at the initial time t_0, and use scattered interpolation to determine the values of S^2 for our original initial points $\mathbf{x}_0$ (Balasuriya 2020*a*,*b*).

The raw S^2 quantity carries a drawback in that it tends to spike to very large maximal values in not a very smooth or cohesive fashion, resulting in large variation (and large gradients) between the values of S^2 throughout the flow domain and the production of a scalar field that can be almost impossible to interpret for some dynamic systems (Balasuriya 2020*a*). There are various methods of scaling this quantity to make the key patterns observable for this quantity more distinguishable within its scalar field, using parameters relevant to our velocity system such as the resolution of the velocity data or the initial grid of points; or the known or expected uncertainty present within this data. The scaled version of S^2 which we will use in this research takes the form

$$S_v\left(\mathbf{x}_0\right)=\epsilon\sqrt{S^2\left(\mathbf{x}_0\right)}=\sqrt{h v_r S^2\left(\mathbf{x}_0\right)}, \tag{2.14}$$

where h is the resolution of the velocity data and v_r is the anticipated uncertainty present within this data. This scaled quantity essentially provides a measure of the lengthscale over which the flow map of a particle which begins at $\mathbf{x}_0$ will experience considerable variation or uncertainty (Balasuriya 2020*a*,*b*).

As mentioned earlier, the stochastic sensitivity can also be used to detect sub–regions of Ω over which particle flow is less sensitive to velocity uncertainty and is therefore expected to be more robust or consistent. We define the robust set

$$R(L,v_r,h)=\left\{\mathbf{x}_0\in\Omega\ :\ S_v(\mathbf{x}_0)<L\right\}, \tag{2.15}$$

which is comprised of partitions of Ω within which the scaled stochastic sensitivity lengthscale $S_v\left(\mathbf{x}_0\right)$ is below a threshold of choice L, and therefore particle flow is expected to be more consistent under the influence of (2.1) and any associated velocity uncertainty. This method therefore provides a way to visualise both the most chaotic and the most consistent particle flow behaviour within a velocity system, with the added flexibility of these coherent structures being able to be fine tuned using lengthscales and velocity uncertainty levels of choice (Balasuriya 2020*a*,*b*).

2.5 The Transfer Operator

All of the LCS detection methods considered so far have been "geometric" in nature, that is, the computations relevant to each method rely on characteristics of flow trajectories such as the flow map gradient and the Cauchy-Green strain tensor; which are usually generated from initial points that seem to pertain to a uniform grid structure even though this does not necessarily happen to be the case (Allshouse & Peacock 2015, Hadjighasem et al. 2017). There are ways of generating LCSs without having to rely on quantities like these. An example of this involves the consideration of probabilistic coherence, which involves seeking an initial density of fluid particles over a flow domain which undergoes the least change under the influence of a dynamic system. One of these methods involves dividing a flow domain Ω into coherent sets using a flow particle transition matrix which acts as a finite dimensional approximation of the transfer operator.

The transfer operator (or the Perron–Frobenius operator) operates on an infinite-dimensional space of density functions in order to model how these functions are mapped from an initial time t_0 to the final time t_0+T under a dynamic velocity system. However, in order to approximate this operator over a flow domain Ω, we use the transition matrix $\mathbf{P}$, which is a standard finite-dimensional discretisation of the operator defined over a finite number of boxes covering Ω at time t_0 as well as the advected flow domain $\widetilde{\Omega} = \mathbf{F}_{t_0}^{t_0+T}(\Omega)$ at time t_0+T (Froyland et al. 2010).

Suppose we divide the domain Ω into a finite collection of boxes $B_j, j = 1, 2, \ldots, M$; and the domain $\widetilde{\Omega}$ into a similar (or sometimes identical) collection of boxes $C_i, i = 1, 2, \ldots, N$. The transition matrix $\mathbf{P}$ is theoretically defined by the column–stochastic matrix

$$\mathbf{P}_{ij} = \frac{\mu\left(B_j \cap \mathbf{F}_{t_0+T}^{t_0}\left(C_i\right)\right)}{\mu\left(B_j\right)}, \tag{2.16}$$

where μ represents Lebesgue measure (area in two dimensions) (Froyland 2005, Allshouse & Peacock 2015). Essentially, if we take a box C_i at time t_0+T, flow it backwards under the system (2.1), measure the overlap between the resultant, deformed box and the box B_j defined at time t_0, and divide this by the area of B_j, we obtain a probability of a fluid particle beginning in box B_j at time t_0 and ending up in box C_i at time t_0+T under the influence of the velocity system (2.1) (Tallapragada & Ross 2013). We define the transition matrix $\mathbf{P}$ as column–stochastic, in contrast to the majority of the literature where it has been defined as a row–stochastic operator (Froyland 2005, Froyland et al. 2010, Froyland 2013, Tallapragada & Ross 2013, Allshouse & Peacock 2015). To estimate (2.16) numerically, we use a discrete version of this matrix

$$\mathbf{P}_{ij} = \frac{\left|\mathbf{x}_k : \mathbf{x}_k \in B_j \wedge \mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_k\right) \in C_i\right|}{Q}, \tag{2.17}$$

for Q initial points $\mathbf{x}_k = (x_k, y_k)$ within the box B_j seeded randomly under a uniform distribution (Froyland et al. 2010, Allshouse & Peacock 2015). Figure 2.8 gives an illustration

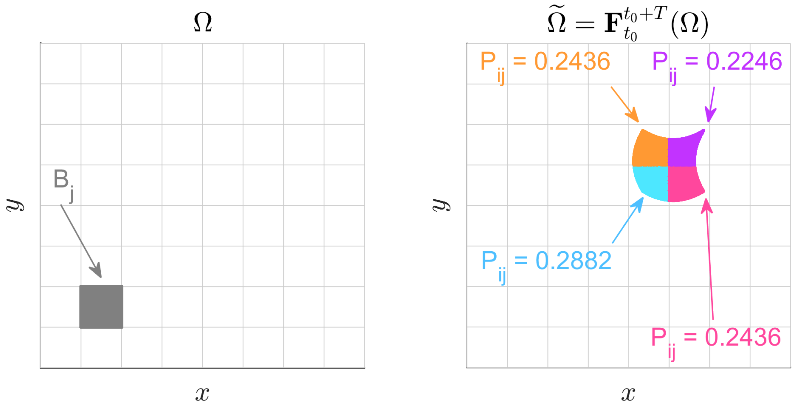

Figure 2.8: A set of initial points $\mathbf{x}_0$ within the box B_j at time t_0 (left), and the transfer operator probabilities $\mathbf{P}_{ij}$ for a range of boxes C_i at time $t_0 + T$ following advection of the initial points through a dynamical system (right).

of how we estimate the elements of $\mathbf{P}$ using this discrete technique.

Ideally, $\mathbf{P}$ should define how the distribution of a concentration of fluid particles at time t_0 is advected through a flow system under the influence of (2.1) over the time interval $[t_0, t_0 + T]$. If we let the vector $\mathbf{v}_0$ define this concentration over all boxes B_j at time t_0, the vector

$$\mathbf{v}_T = \mathbf{P}\mathbf{v}_0$$

defines the spread of these particles over the boxes C_i at time $t_0 + T$ following their advection through the system. The pullback operator, $\mathbf{P}^\top$, sends flow particles at the final flow time $t_0 + T$ back to the initial time t_0. Applying this to the above equation gives

$$\mathbf{P}^\top \mathbf{v}_T = \mathbf{P}^\top \mathbf{P}\mathbf{v}_0.$$

The aim of the transfer operator method is to find an initial particle concentration $\mathbf{v}_0$ such that $\mathbf{P}^\top \mathbf{v}_T \approx \mathbf{v}_0$, thereby preserving $\mathbf{v}_0$ after flowing these distributions forward by $\mathbf{P}$ and flowing them back again using $\mathbf{P}^\top$; that is,

$$\mathbf{v}_0 \approx \mathbf{P}^\top \mathbf{P}\mathbf{v}_0.$$

This goal yields the eigenvalue problem

$$\mathbf{P}^\top \mathbf{P}\mathbf{v}_0 = \lambda \mathbf{v}_0. \tag{2.18}$$

The matrix $\mathbf{P}^\top \mathbf{P}$ is symmetric and positive definite, so all eigenvalues will be non–negative; and in the case of this matrix, the largest eigenvalue is $\lambda_1 = 1$. The corresponding eigenvector $\mathbf{v}_1$ is the indicator function over the entire flow domain. The next largest eigenvalue λ_2 has a value slightly less than 1, and its corresponding eigenvector $\mathbf{v}_2$

produces a "strongly coherent" initial distribution over all of the boxes B_j at time t_0. By partitioning these distributions over a zero contour level, or an optimal non–zero level, we (in a rather ad hoc fashion) divide Ω into two coherent sets based on the flow behaviour of particles under the influence of (2.1) over the time interval $[t_0, t_0 + T]$ (Froyland et al. 2010, Froyland 2013). Subsequent eigenvectors such as $\mathbf{v}_3$ can also be used to find other pairs of coherent sets provided that there is a sufficient gap (referred to as a spectral gap) between the corresponding eigenvalue and the subsequent eigenvalue (Tallapragada & Ross 2013); or a combination of the first k eigenvectors can be used to partition the flow domain Ω into k coherent flow objects using techniques such as k–means clustering (Froyland & Junge 2018) or the SEBA algorithm (Froyland et al. 2019). If we define a large number of boxes B_j at time t_0 and C_i at time $t_0 + T$, the matrix $\mathbf{P}$ will be considerably large and calculating the matrix product $\mathbf{P}^\top\mathbf{P}$ will take up a large amount of memory in programming software such as MATLAB. To overcome this issue, Froyland et al. (2010) use a memory–conserving calculation technique which involves obtaining the same coherent sets for Ω from singular values and vectors of $\mathbf{P}$ instead. We define $\mathbf{v}_0$ as a vector of the Lebesgue measures (areas) of all initial boxes B_j at time t_0, and define the spread of these measures $\mathbf{v}_T$ in the manner done above. The elements of these vectors $\mathbf{v}_0$ and $\mathbf{v}_T$ define the non–zero entries of the diagonal matrices $\mathbf{\Pi}_0$ and $\mathbf{\Pi}_T$ respectively, which are used to define the matrix

$$\widetilde{\mathbf{P}} = \mathbf{\Pi}_T^{-\frac{1}{2}} \mathbf{P} \mathbf{\Pi}_0^{\frac{1}{2}}. \tag{2.19}$$

We then generate the largest singular values (not the largest eigenvalues) of $\widetilde{\mathbf{P}}$, the first of which, σ_1, should equal 1. The next largest singular value, σ_2, is the square root of λ_2 which is the second largest eigenvalue of $\mathbf{P}^\top\mathbf{P}$, and corresponds to a right singular vector $\mathbf{y}_r$ and a left singular vector $\mathbf{y}_l$. Again, singular vectors corresponding to a subsequent singular value such as σ_3 could be used instead. We use these right and left singular vectors to define "optimal vectors" which are used to obtain coherent sets for Ω at time t_0 and $\widetilde{\Omega}$ at time $t_0 + T$, respectively. The optimal vector

$$\mathbf{v}_0^* = \mathbf{\Pi}_0^{-\frac{1}{2}} \mathbf{y}_r$$

is used to detect two coherent sets at time t_0, while the advection of these coherent sets under (2.1) to the time $t_0 + T$ can be detected using the other optimal vector

$$\mathbf{v}_T^* = \mathbf{\Pi}_T^{-\frac{1}{2}} \mathbf{y}_l$$

(Froyland et al. 2010). The primary focus of this research will be on coherent sets relevant to the initial flow time t_0 obtained using $\mathbf{v}_0^*$.

To carry out this method numerically, we divide the flow domain Ω at time t_0 into $M = M_x \times M_y$ rectangular boxes of equal size, and the anticipated advected domain $\widetilde{\Omega}$ at time $t_0 + T$ into $N = N_x \times N_y$ rectangular boxes, initialising the transition matrix $\mathbf{P}$

as a $N_xN_y \times M_xM_y$ matrix. Within each of the boxes B_j at time t_0, we seed Q initial flow points $\mathbf{x}_0$ under a uniform distribution. As far as the literature is concerned, only a small value for Q ranging between 400 (Froyland et al. 2010, Froyland 2013) and 1600 (Froyland 2015) points is sufficient. Over each box B_j, we flow these particles forward for T units of time under (2.1), identify which boxes C_i at time $t_0 + T$ these particles have "landed" into, and use these to create vectors of sample particle advection probabilities which form columns of $\mathbf{P}$. The procedure continues for all remaining boxes B_j defined at t_0. We then obtain singular values for $\mathbf{P}$ and their corresponding singular vectors using the technique described above as outlined by Froyland et al. (2010). We will primarily consider the second right singular vector of $\mathbf{P}$ to obtain the optimal vector $\mathbf{v}_0^*$ for all of our flow systems, though we will also use subsequent singular vectors if the coherent sets produced by these vectors give a meaningful visualisation of the flow system in question.

2.6 The Dynamic Laplace Operator

The dynamic Laplace operator is very similar to the transfer operator defined in the last section in that it is used to split a flow domain Ω into two (or potentially more) coherent sets. The key difference is that this operator uses a Laplacian matrix (which will be defined shortly) to explicitly incorporate diffusion into the flow defined by the transfer operator matrix. Froyland (2015) first used this technique to divide a flow domain Ω into two coherent sets using $\mathbf{\Delta}$, the static Laplace operator matrix relevant to a velocity system where Ω does not change shape. The primary motivation of this method is to find an optimal boundary ζ that divides Ω into two coherent sets by minimising the ratio between the length of the boundary and the area of the coherent sets subsequently produced; otherwise known as the Cheeger constant which in this case is defined as

$$\chi = \min_{\zeta} \frac{\mathcal{L}(\zeta)}{\min\left(A(S_1), A(S_2)\right)}, \tag{2.20}$$

where $\mathcal{L}$ indicates the length of the boundary ζ and A indicates the area (Lebesgue measures in the first and second dimension, respectively) of S_1 and S_2; the sets separated by ζ (Froyland 2015, Hadjighasem et al. 2017). Suppose we divide Ω into $M = M_x \times M_y$ rectangular boxes of equal size as was done for the transfer operator in the previous section. Now suppose we have some function $f(x, y)$ defined over the domain Ω. Define the vector

$$\mathbf{f} = [f_{1,1}\ f_{2,1}\ \cdots\ f_{M_x,1}\ f_{1,2}\ f_{2,2}\ \cdots\ f_{M_x,2}\ \cdots\ f_{1,M_y}\ f_{2,M_y}\ \cdots\ f_{M_x,M_y}],$$

where $f_{i,j}$ is the value of the function at the central point of the i–th box in the x–direction and the j–th box in the y–direction. The elements of the Laplace operator matrix $\mathbf{\Delta}$ are the finite element coefficients used to compute the Laplacian $\Delta f = \nabla^2 f$ for the function defined over this box configuration for Ω. We obtain a vector of these Laplacian values through the matrix multiplication $\mathbf{\Delta f}$. In the corresponding eigenvalue problem

$$\mathbf{\Delta v} = \lambda \mathbf{v}, \tag{2.21}$$

the first eigenvalue $\lambda_1 \approx 0$, and gives an approximation for the static Cheeger constant χ (Froyland & Junge 2018). All remaining eigenvalues are negative and decreasing in order tending to negative infinity, i.e. $0 \approx \lambda_1 > \lambda_2 > \lambda_3 > \cdots$. Like with the transfer operator method, the second eigenvector $\mathbf{v}_2$ of the eigenvalue problem (2.21) is used to find two almost invariant coherent sets, which are separated by a zero contour or the contour corresponding to the aforementioned Cheeger constant estimate λ_1 (Froyland & Junge 2018).

For a non–autonomous dynamical system such as (2.1), the dynamic Laplace operator now applies the diffusion effect on top of the dynamics of the flow. In this case, the Cheeger constant now takes the form

$$\chi^D = \min_{\zeta} \frac{\mathcal{L}\left(\zeta\right) + \mathcal{L}\left(\mathbf{F}_{t_0}^{t_0+T}\left(\zeta\right)\right)}{2 \min\left(A(S_1), A(S_2)\right)}, \tag{2.22}$$

(Froyland 2015), and the dynamic Laplace operator is defined as

$$\mathbf{\Delta}^D = \frac{1}{2}\left(\mathbf{\Delta}_0 + \mathbf{P}^\top \mathbf{\Delta}_T \mathbf{P}\right), \tag{2.23}$$

where $\mathbf{P}$ is the finite–dimensional discretisation of the transfer operator as defined in the previous section, and $\mathbf{\Delta}_0$ and $\mathbf{\Delta}_T$ are Laplacian operator matrices defined over the box configurations of Ω at time t_0 and $\widetilde{\Omega}$ at time $t_0 + T$, respectively. The elements of each of these Laplacian operators are found using the finite difference formula

$$\begin{aligned} \frac{1}{4h_x^2}\frac{u_{m+1,n}}{u_{m,n}} f_{m+2,n} + \frac{1}{4h_x^2}\frac{u_{m-1,n}}{u_{m,n}} f_{m-2,n} + \frac{1}{4h_y^2}\frac{u_{m,n+1}}{u_{m,n}} f_{m,n+2} + \frac{1}{4h_y^2}\frac{u_{m,n-1}}{u_{m,n}} f_{m,n-2} \\ - \left(\frac{1}{4h_x^2}\frac{u_{m+1,n} - u_{m-1,n}}{u_{m,n}} + \frac{1}{4h_y^2}\frac{u_{m,n+1} - u_{m,n-1}}{u_{m,n}}\right) f_{m,n}, \end{aligned}$$

where h_x and h_y are the lengths of each box in the x and y directions respectively, f is the function for which we would compute the Laplacian with discrete values expressed within the vector $\mathbf{f}$ defined above, and each u is taken from a vector representing the areas of each box defined at time t_0 (or $t_0 + T$) with similar dimensions to $\mathbf{f}$ (Froyland & Kwok 2017). We then solve a similar eigenvalue problem to (2.21)

$$\mathbf{\Delta}^D \mathbf{v} = \lambda \mathbf{v}, \tag{2.24}$$

which produces an analogous spectrum of eigenvalues and their corresponding eigenvectors to those produced from (2.21). Like with the transfer operator, we can obtain two coherent sets for Ω using the second eigenvector $\mathbf{v}_2$, obtain a different pair of coherent sets using a subsequent eigenvector such as $\mathbf{v}_3$, or find k coherent sets using techniques such as k–means clustering applied to the first k eigenvectors obtained from the eigenvalue problem (2.24) (Froyland & Junge 2018).

While we will use the matrix definition of the dynamic Laplace operator (2.23) for the majority of flow systems considered in this study, we will also use a sparse computational algorithm for this LCS detection method conceptualised by Froyland & Junge (2018). This method comes in useful for computing two coherent sets for a dynamic system defined by flow trajectory data over a sparse collection of initial points, a non–rectangular or otherwise awkwardly shaped flow domain, or obstruction of certain portions of the flow domain by land in an oceanographic flow system; features of flow data which would otherwise make producing a Laplacian matrix operator over a uniform grid of rectangular boxes difficult or impossible. The method has been derived using finite element techniques and does not require the prior computation of the transition matrix $\mathbf{P}$ as the relevant information obtained from this matrix is embedded within the finite element definitions of the operator. Suppose we have N initial points within our domain Ω, defined in a uniformly spaced grid or in a sparse fashion. We then obtain, or are given (depending on the flow data available), flow trajectories relevant to each of these points defined over $K+1$ time steps $t_k = t_0 + k\Delta t$, where $k = 0, 1, 2, \dots, K$ and $K\Delta t = T$. Each initial point, along with corresponding points along its flow trajectory, is assigned an index ranging from 1 to N. At each time step t_k, we take the formation of all flow trajectory points as they stand within Ω at that time, and produce a triangular mesh of these points either using Delaunay triangulation or an Alpha Triangulation method in MATLAB. We take one of these triangles, which we will denote by Υ, with vertices $\mathbf{x}_a = (x_a, y_a)$, $\mathbf{x}_b = (x_b, y_b)$ and $\mathbf{x}_c = (x_c, y_c)$, where a, b and c are the respective indices for these points corresponding to a particular flow trajectory. This triangle will carry three finite element basis functions respective to each vertex, which we can define by the formulae

$$\phi_a(x, y) = \frac{1}{|\mathbf{X}|}\left((y_b - y_c)\,x + (x_c - x_b)\,y + (x_b y_c - x_c y_b)\right),$$

$$\phi_b(x, y) = \frac{1}{|\mathbf{X}|}\left((y_c - y_a)\,x + (x_a - x_c)\,y + (x_a y_c - x_c y_a)\right),$$

$$\phi_c(x, y) = \frac{1}{|\mathbf{X}|}\left((y_a - y_b)\,x + (x_b - x_a)\,y + (x_a y_b - x_b y_a)\right),$$

where

$$|\mathbf{X}| = \det\begin{bmatrix} 1 & x_a & y_a \\ 1 & x_b & y_b \\ 1 & x_c & y_c \end{bmatrix}.$$

We use these formulae to produce a mass matrix $\mathbf{M}$ and a stiffness matrix $\mathbf{D}$ for this triangle. The mass matrix $\mathbf{M}_{ij}$ contains the inner products of the basis functions ϕ_i and

ϕ_j defined over this triangle. After some hand calculation, this matrix is found to take the form

$$\mathbf{M} = \frac{A\left(\Upsilon\right)}{12}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

where $A\left(\Upsilon\right)$ denotes the area of the triangle. The stiffness matrix $\mathbf{D}_{ij}$ contains the Lebesgue integrals of the dot products of the gradient vectors of the basis functions ϕ_i and ϕ_j over the area of this triangle. After some further hand calculation, the elements on the diagonal of this matrix turn out to equal

$$\begin{aligned} \mathbf{D}_{11} &= \frac{A\left(\Upsilon\right)}{|\mathbf{X}|^2}\left(\left(y_b - y_c\right)^2 + \left(x_c - x_b\right)^2\right), \\ \mathbf{D}_{22} &= \frac{A\left(\Upsilon\right)}{|\mathbf{X}|^2}\left(\left(y_c - y_a\right)^2 + \left(x_a - x_c\right)^2\right), \\ \mathbf{D}_{33} &= \frac{A\left(\Upsilon\right)}{|\mathbf{X}|^2}\left(\left(y_a - y_b\right)^2 + \left(x_b - x_a\right)^2\right), \end{aligned}$$

which look the like the squares of the distances between the vertices at which each respective basis function vanishes multiplied by the area of the triangle and divided by the square of the above defined matrix determinant. For the remaining matrix elements,

$$\begin{aligned} \mathbf{D}_{12} = \mathbf{D}_{21} &= \frac{A\left(\Upsilon\right)}{|\mathbf{X}|^2}\left(\left(y_b - y_c\right)\left(y_c - y_a\right) + \left(x_c - x_b\right)\left(x_a - x_c\right)\right), \\ \mathbf{D}_{13} = \mathbf{D}_{31} &= \frac{A\left(\Upsilon\right)}{|\mathbf{X}|^2}\left(\left(y_b - y_c\right)\left(y_a - y_b\right) + \left(x_c - x_b\right)\left(x_b - x_a\right)\right), \\ \mathbf{D}_{23} = \mathbf{D}_{32} &= \frac{A\left(\Upsilon\right)}{|\mathbf{X}|^2}\left(\left(y_c - y_a\right)\left(y_a - y_b\right) + \left(x_a - x_c\right)\left(x_b - x_a\right)\right). \end{aligned}$$

We then consider a larger stiffness matrix $\mathbf{D}^{(k)}$ and a larger mass matrix $\mathbf{M}^{(k)}$, each of size $N \times N$ and relevant to the time step $t_0 + k\Delta t$. We add the elements of the first row of $\mathbf{D}$ to the (a, a)-th, (a, b)-th and (a, c)-th entries of $\mathbf{D}^{(k)}$ respectively, the elements of the second row of $\mathbf{D}$ to the (b, a)-th, (b, b)-th and (b, c)-th entries of $\mathbf{D}^{(k)}$ respectively, and the elements of the third row of $\mathbf{D}$ to the (c, a)-th, (c, b)-th and (c, c)-th entries of $\mathbf{D}^{(k)}$ respectively. We carry out a similar procedure for adding the elements of $\mathbf{M}$ to the entries of $\mathbf{M}^{(k)}$. After completing this for one triangle at time t_k, we repeat this procedure for all of these triangles to complete the matrices $\mathbf{D}^{(k)}$ and $\mathbf{M}^{(k)}$. If two points $\mathbf{x}_i$ and $\mathbf{x}_j$ do not form a triangle with a third point at time t_k, $\mathbf{D}^{(k)}_{ij} = \mathbf{M}^{(k)}_{ij} = 0$. After producing these matrices for each time step t_k, we define the matrices

$$\mathbf{D}_{t_0}^{t_0+T} = -\frac{1}{K+1}\sum_{k=0}^{K}\mathbf{D}^{(k)},$$

$$\mathbf{M}_{t_0}^{t_0+T} = \frac{1}{K+1}\sum_{k=0}^{K}\mathbf{M}^{(k)}$$

and solve the generalised eigenvalue problem

$$\mathbf{D}_{t_0}^{t_0+T}\mathbf{v} = \lambda\mathbf{M}_{t_0}^{t_0+T}\mathbf{v}. \tag{2.25}$$

The dynamic Laplace operator $\mathbf{\Delta}^D$ is discretised into the matrices $\mathbf{D}_{t_0}^{t_0+T}$ and $\mathbf{M}_{t_0}^{t_0+T}$ using finite element techniques. Upon solving this system, we obtain an analogous collection of eigenvalues $0 \approx \lambda_1 > \lambda_2 > \lambda_3 > \cdots$ and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots$, which are used to obtain coherent sets for Ω as was the case with the dynamic Laplace operator defined over a rectangular grid of boxes.

2.7 Fuzzy C–Means Clustering

All of the aforementioned discussed LCS detection methods (aside from the finite element discretisation of the dynamic Laplace operator) share one common thread - each of these methods rely on knowledge of analytic or Eulerian velocity data representing the flow system (2.1). In some instances of real world data, only sparsely defined flow trajectories are available without the underlying velocity data, and in these instances none of the above methods can be used to find LCSs. We therefore need to consider the use of sparse LCS detection methods which group together batches of flow trajectories based on the consistent closeness of flow trajectories over the duration of a flow or familiar mixing behaviour observable between groups of these trajectories. One of these methods is the fuzzy c–means (FCM) clustering method, which groups together a series of sparsely defined flow trajectories into a pre-determined number of coherent clusters.

Bezdek et al. (1984, 1987) conceptualised the FCM clustering algorithm to cluster together groups of individual Cartesian data points by iteratively updating the centres of a number of pre-determined clusters and the membership probabilities of a point belonging to any of these clusters. This algorithm was subsequently expanded to cluster discretely defined flow trajectories instead of just grid points based on their proximity and the likelihood of groups of flow particles maintaining a similar shape (Froyland & Padberg-Gehle 2015, Allshouse & Peacock 2015). This method can be executed using more continuous definitions of flow trajectories (Froyland & Padberg-Gehle 2015); however in this research we will only focus on the simplest version of this method, which involves the use of discretely defined Lagrangian flow trajectories.

We begin with a collection of flow trajectories for the system (2.1) defined over the time interval $[t_0, t_0 + T]$ which begin from N distinct initial points $\mathbf{x}_0^{(j)} \in \Omega$, $j = 1, \ldots, N$, arranged either sparsely or in a uniform grid. Each flow trajectory is arranged into a vector of the form

$$\mathbf{X}_j = \left[\mathbf{F}_{t_0}^{t_0}\left(\mathbf{x}_0^{(j)}\right), \mathbf{F}_{t_0}^{t_0+\Delta t}\left(\mathbf{x}_0^{(j)}\right), \mathbf{F}_{t_0}^{t_0+2\Delta t}\left(\mathbf{x}_0^{(j)}\right), \ldots, \mathbf{F}_{t_0}^{t_0+T}\left(\mathbf{x}_0^{(j)}\right)\right], \tag{2.26}$$

for n equispaced time steps $\Delta = T/n$. We then pre-determine an appropriate number of clusters K expected to exist within Ω based on the behaviour of (2.1) over this time interval, either through a priori information if available or by arbitrarily guessing a value. We submit the trajectories $\mathbf{X}_j$ to MATLAB's in–built `fcm` algorithm, which implements the FCM clustering algorithm to group these trajectories into K clusters or groups. The algorithm achieves this by iteratively updating the "central trajectory" of each cluster,

$$\mathbf{C}_k = \frac{\sum\limits_{j=1}^{N} (p_{k,j})^m \mathbf{X}_j}{\sum\limits_{j=1}^{N} (p_{k,j})^m}, \tag{2.27}$$

and the membership probability, or the likelihood of a trajectory belonging to each cluster,

$$p_{k,j} = \left[\sum_{i=1}^{K} \left(\frac{\|\mathbf{X}_j - \mathbf{C}_k\|}{\|\mathbf{X}_j - \mathbf{C}_i\|}\right)^{\frac{2}{m-1}}\right]^{-1} \tag{2.28}$$

(Allshouse & Peacock 2015). The process continues until the quantity

$$\sum_{k=1}^{K}\sum_{j=1}^{N} (p_{k,j})^m \|\mathbf{X}_j - \mathbf{C}_k\|^2$$

has been minimised with respect to an acceptable threshold (Froyland & Padberg-Gehle 2015). In the above equations, the parameter m is a sharpness parameter which determines the resolution of each FCM cluster. The value of m must be greater than 1, and larger values of m will result in poorer resolution and the formation of clusters which will have much less of a well-defined shape. $m = 1.5$ is a typical value for this parameter (Allshouse & Peacock 2015, Froyland & Padberg-Gehle 2015), while stronger resolution can be obtained in some instances by lowering m to between 1.1 and 1.25 (Froyland & Padberg-Gehle 2015, Hadjighasem et al. 2017). It is not recommended to assign this parameter a value greater than 2 (Froyland & Padberg-Gehle 2015). Once the procedure has been completed, K coherent clusters are obtained for Ω by grouping the sparse or uniform flow trajectories $\mathbf{X}_j$ by their membership probabilities against a probability threshold p_0. If the likelihood of a trajectory $\mathbf{X}_j$ belonging to cluster K is greater than or equal to this probability threshold, the trajectory is a "member" of this coherent cluster. The choice of p_0 depends on a number of factors pertinent to the velocity system being investigated, though in some instances it can be selected by arbitrarily guessing a threshold or selecting one following some experimentation with the parameter (Allshouse & Peacock 2015).

2.8 Coherent Structure Colouring

A major drawback with the FCM clustering detection method is that the number of coherent clusters expected to exist within Ω under (2.1) over the time interval $[t_0, t_0 + T]$ has to be pre-determined before implementing the algorithm, either through some a priori known information or just by making an arbitrary guess. In some simplistic flow systems, such as the Double Gyre system, pre-determining the number of clusters is straightforward (Allshouse & Peacock 2015). However, this will not be true for the vast majority of real world type flow systems where "true" LCSs are known to exist, and implementing the FCM clustering method will require experimentation with the number of pre-determined clusters with no guarantee of a cohesive or meaningful result to emerge from this. Instead, one can use an alternative sparse LCS detection method which involves the use of coherent structure colouring (CSC).

The CSC method embodies techniques from graph theory to group together flow trajectories for (2.1) over $[t_0, t_0 + T]$ based on their kinematic similarities. What this involves is quantifying how similarly flow trajectories move through a system, and how they braid, mix or weave together with other trajectories as a flow progresses; by way of defining a network graph involving the flow particles at a series of time steps between t_0 and $t_0 + T$; and quantifying how this network changes shape and formation with each time step. By solving a generalised eigenvalue problem related to these metrics, one is able to divide Ω into a range of shapes and sets where all trajectories beginning within these coherent shapes move together in a more or less similar fashion (Schlueter-Kuck & Dabiri 2017).

To carry out this method, we begin by once again producing a set of N flow trajectories $\mathbf{X}_j$ in a scattered or uniform formation and arrange these in the form (2.26) used for the previous method. Then at each time step $t_k = t_0 + k\Delta t$, $k = 0, 1, 2, \ldots, K-1$, we define a network graph for each of the trajectory points at time t_k, where the trajectory points form the nodes, and the edge weights $r_{ij}(t_k)$ are the Euclidean distances between node i and node j at time t_k (Schlueter-Kuck & Dabiri 2017). The distances between flow trajectories is an easy choice for the quantity used to measure the kinematic similarities between these trajectories, though there may be scope for using other quantities such as the absolute difference in the vorticities of these trajectories. Next, we construct the $N \times N$ adjacency matrix

$$\mathbf{A}_{ij} = s_{ij}/\overline{r_{ij}}, \tag{2.29}$$

where $\overline{r_{ij}}$ is the mean distance between trajectories i and j over all time steps t_k, and

$$s_{ij} = \frac{1}{\sqrt{K}} \sqrt{\sum_{k=0}^{K-1} \left(r_{ij}(t_k) - \overline{r_{ij}}\right)^2};$$

which is essentially the standard deviation of the distances r_{ij}. The adjacency matrix entries can be calculated all at once, but if the number of trajectories N and/or the number of time steps K are too large, resulting in system memory issues in MATLAB, the quantities $\overline{r_{ij}}$ and s_{ij} can be computed iteratively using the formulae

$$\begin{aligned}\overline{r_{ij}}^{(k)} &= \frac{1}{k}\left(r_{ij}\left(t_k\right) + (k-1)\,\overline{r_{ij}}^{(k-1)}\right),\\ \sigma_{ij}^{(k)} &= \frac{k-2}{k-1}\,\sigma_{ij}^{(k-1)} + \frac{1}{k}\left(r_{ij}\left(t_k\right) - \overline{r_{ij}}^{(k-1)}\right)^2,\end{aligned}$$

for each time step t_k, where $\sigma_{ij} = s_{ij}^2$ (Welford 1962). We then sum the rows of $\mathbf{A}$ to produce the entries for the diagonal matrix $\mathbf{D}$, use both of these matrices to calculate the Graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$, and solve the generalised eigenvalue problem

$$\mathbf{L}\mathbf{v} = \lambda\mathbf{D}\mathbf{v}. \tag{2.30}$$

We seek the largest eigenvalue $\lambda_{\max}$ of (2.30), and its corresponding eigenvector $\mathbf{v}_{\max}$. The values of $\mathbf{v}_{\max}$ are coherent structure colouring coefficients assigned to each of the N trajectories, in the order in which they were indexed to produce the adjacency matrix $\mathbf{A}$. Clusters of flow trajectories which correspond to similar values of this coefficient form LCSs for the domain Ω, and can take many shapes from coherent jet type streams to elliptic vortices and coherent flow bundles of ambiguous shape (Schlueter-Kuck & Dabiri 2017). To visualise these coherent shapes, we either directly produce a scalar field of the coefficients if the initial points $\mathbf{x}_0$ were uniformly distributed, or if these points were initially scattered, we produce a uniform grid of points which covers (but does not go beyond) the range of initial points $\mathbf{x}_0$ in both the x and y directions, and interpolate the coherent structure colouring coefficients over the points in this grid. If the option is available, it is preferable to define a uniform grid of initial points at time t_0 to avoid this extra calculation and the interpolation errors associated with it, however if the initial data set consists only of scattered flow trajectory data with no information regarding the underlying velocity data, this extra step cannot be avoided (Schlueter-Kuck & Dabiri 2017).

Chapter 3

Resultant Coherent Structures From Each Method

In this Chapter, we document the results of implementing each of the LCS detection methods described in Chapter 2 on a range of flow systems. First, we display LCS results for a range of analytically defined idealised flow systems to familiarise ourselves with each of the methods and (where possible) compare our results with those published in the literature. The analytically defined flows which we will consider here are the Double Gyre flow, a non-autonomous version of the Stuart vortex system, and the Bickley Jet flow. We then turn our attention to three flavours of "real–world" type data, including two versions of velocity data simulated under the Navier–Stokes equations featuring Kelvin–Helmholtz vortex layers, and an oceanographic data set representing the well known Gulf Stream system in the North Atlantic Ocean. We have written our own code in MATLAB to implement each of these methods, with evidence of validation of this code on both analytically and numerically defined velocity data sets available to view in Appendix A. Table 3.1 shows abbreviations for some of the flow systems introduced in this Chapter.

Table 3.1: Flow systems first defined in Chapter 3, and their abbreviations.

Flow System	Abbreviation	Equation(s)	Defined in/at
Double Gyre	DG	(3.1)	Section 3.1
Unsteady Stuart Vortex	USV	(3.2)	Section 3.2
Bickley Jet	N/A	(3.3)	Section 3.3
Kelvin–Helmholtz Version 1	KH1	(3.4)	Section 3.4
Kelvin–Helmholtz Version 2	KH2	(3.5)	Section 3.5
Gulf Stream	N/A	N/A	Section 3.6

3.1 Double Gyre

The first system which we will consider is the Double Gyre (DG) flow, which has regularly been used for LCS analysis in many published studies. The Double Gyre system takes the form

$$\dot{\mathbf{x}} = \mathbf{u}\left(x, y, t\right) := \begin{bmatrix} -\pi A \sin\left(\pi f\left(x,t\right)\right)\cos\left(\pi y\right) \\ \pi A \cos\left(\pi f\left(x,t\right)\right)\sin\left(\pi y\right)\frac{\partial f}{\partial x} \end{bmatrix}, \tag{3.1}$$

where

$$f\left(x,t\right) = x^2\left(\epsilon_* \sin\left(\omega t\right)\right) + x\left(1 - 2\epsilon_* \sin\left(\omega t\right)\right).$$

The flow is defined over a closed domain $\Omega = [0,2] \times [0,1]$ and is characterised by two counter–rotating gyres centred at $(1/2, 1/2)$ and $(3/2, 1/2)$. If the perturbation parameter $\epsilon_* = 0$, the sizes of the gyres do not change and a consistent vertical flow barrier centred along $x = 1$ separates Ω into two square "gyre chambers" as these gyres rotate in an autonomous version of this system. However if $\epsilon_* > 0$, the gyres stretch and shrink in horizontal size causing this vertical flow barrier to deform and fluid particles to move between gyre chambers. The parameter ω determines the time over which one period of this full sliding action occurs, which is $2\pi/\omega$, while A determines the scale or amplitude of the velocity (Shadden et al. 2005, Balasuriya 2020*a*). The existence of flow barriers with wild and unpredictable shapes alongside the more coherent rotating gyres within this system; along with the relative simplicity of this method from its analytical velocity definition, makes this a popular method for use at the very least as a first experiment for LCS detection using a wide variety of methods (Shadden et al. 2005, Farazmand & Haller 2012, Tallapragada & Ross 2013, Allshouse & Peacock 2015, Froyland & Padberg-Gehle 2015, Onu et al. 2015, Balasuriya 2020*a*,*b*). We choose a similar set of parameters to those used in Shadden et al. (2005) and Farazmand & Haller (2012), which are $A = 0.1$, $\epsilon = 0.1$ and $\omega = \pi/5$. We choose the same flow period used in Farazmand & Haller (2012), which ranges from $t_0 = 0$ to $t_0 + T = 20$.

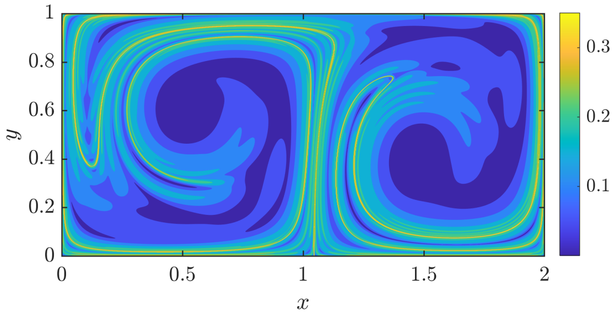

Figure 3.1: The FTLE Φ_0^{20} for the DG flow.

We begin with the FTLE for this flow, which is displayed in Figure 3.1. We compute this field using Lagrangian flow data defined over a uniform grid of 1000 × 1000 points.

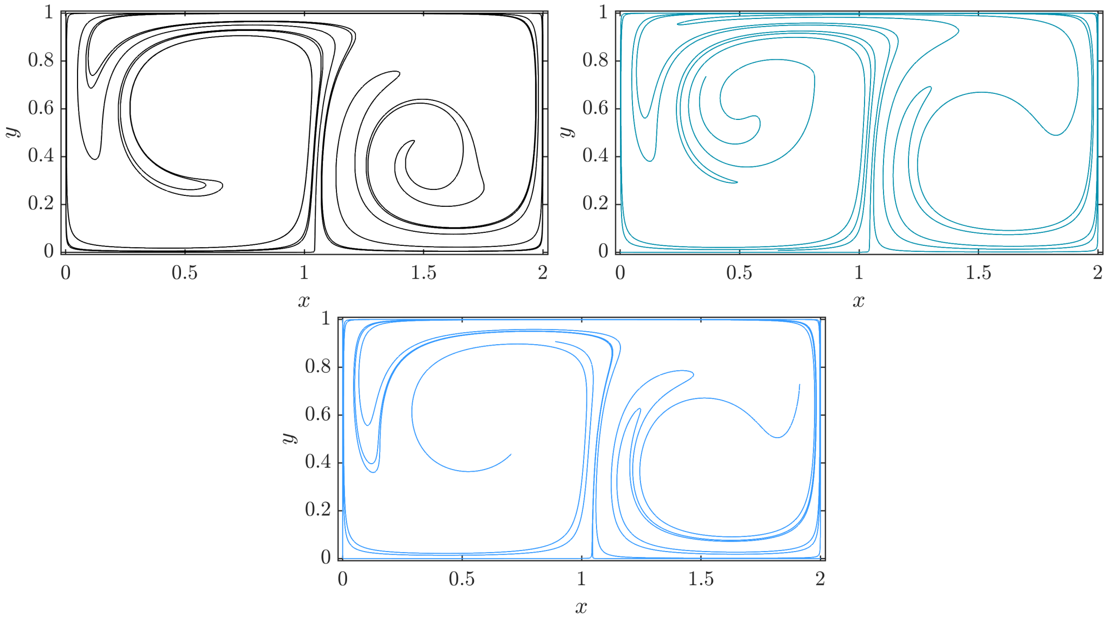

Figure 3.2: Hyperbolic VLCSs for the DG flow computed using the VLCS–A (top–left), VLCS–B and VLCS–C (bottom) algorithms.

The maximal ridges of this field indicate how the vertical flow barrier of the Double Gyre system has warped in shape under the influence of the gyres which expand and shrink as the flow progresses. We do not have several maximal ridges of the FTLE within Ω, but rather one continuous curve which begins from approximately the point $(1.05, 0)$, wraps around the left gyre from above, produces a parabolic loop with a vertex positioned at approximately the point $(0.15, 0.38)$, traverses the upper edge of the domain, curls around the right gyre from below, produces another parabolic loop with approximate vertex $(1.4, 0.75)$, curls back around the right gyre, traverses the upper and left edges of the domain, before traversing the lower edge of the left gyre chamber, wrapping around the left gyre again and terminating about the point $(0.6, 0.35)$. More coherent flow is observable from the regions of the domain corresponding to the lowest FTLE values, which include the two gyres themselves and long, thin “tails” which emanate from these circular regions.

Next, we consider the VLCS detection method, with the results from our three detection algorithms featured in Figure 3.2. To carry out this method, we once again compute the Lagrangian flow map for 1000×1000 initial points $\mathbf{x}_0 \in \Omega$. For VLCS–A, we randomly sample 500 points from the acceptable set $\mathcal{G}_0$ to begin integration of the $\boldsymbol{\xi}_1$ eigenvector field, set the allowance parameter $\mathcal{L}_f = 2$, and isolate the five strainlines of largest size to avoid overcrowding the domain. In VLCS–B, we choose $\rho = 0.2$ and $\mathcal{L}_{\max} = 30$, and in VLCS–C we maintain $\rho = 0.2$ and $\mathcal{L}_{\max} = 30$ while reducing $\mathcal{L}_f = 0.5$ due to the

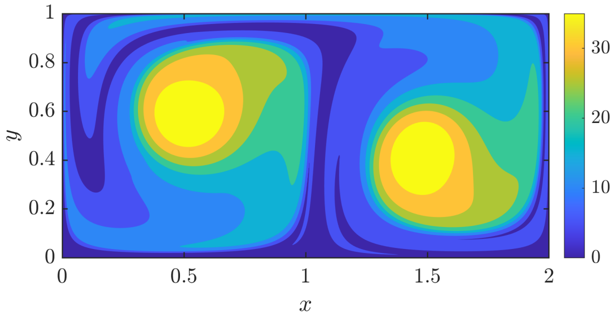

Figure 3.3: The LAVD Θ_0^{20} for the DG flow.

smoothness of the FTLE in comparison to $\mathcal{G}_0$. Each of these algorithms has detected multiple flow barriers of similar shape to the maximal ridge of the FTLE field identified in Figure 3.1, though several key differences are observable. For instance, the parabolic loops produced vary in width, these loops rotate at different angles around the gyres (particularly in the case of the right gyre), and hyperbolic LCSs appear to have been generated within the gyres themselves; though these structures are "locally maximal" by the definitions of the method (Farazmand & Haller 2012, Onu et al. 2015). The methods each produce a slightly different foliation of VLCSs, due to variations in the initial points of eigenvector field integration and differences in the regulation of these computations between each algorithm.

We next consider the LAVD method, which we calculate by solving (2.7) for the same uniform grid of 1000×1000 initial points within the domain Ω. The resultant scalar field can be viewed in Figure 3.3. Convex structures corresponding to large values of the LAVD are LCSs which exhibit strong rotational coherence within Ω, which for this flow are the olive coloured, bean shaped structures surrounding each gyre. Interestingly, the lowest values of the LAVD are observable along the vertical line $x = 1$ and along a shape not too dissimilar to the flow barriers extracted from the previous two methods. It must be stressed that this is a special case of the LAVD, and that structures of low level LAVD as a rule do not necessarily coincide with flow barriers or structures of large particle repulsion capability in all dynamic velocity systems (Haller et al. 2016).

The S^2 for the DG system is then computed over the same initial grid of points defined for the previous three methods, using the end–time grid expansion technique for computing the gradient as Ω is closed. In Figure 3.4, we plot the natural logarithm of the scaled S^2 quantity (2.14), along with three robust sets for this system subject to three different lengthscales L. As the Double Gyre is a non–dimensional flow and we have no a priori estimate regarding the velocity uncertainty given in non–dimensional coordinates, we choose $\epsilon = 1 \times 10^{-2}$. The Scaled S^2 shown in Figure 3.4 looks almost identical to the FTLE scalar field shown in Figure 3.1 bar some minor differences, such as the relative smoothness of the "tails" emanating from the gyres themselves, and considerably low

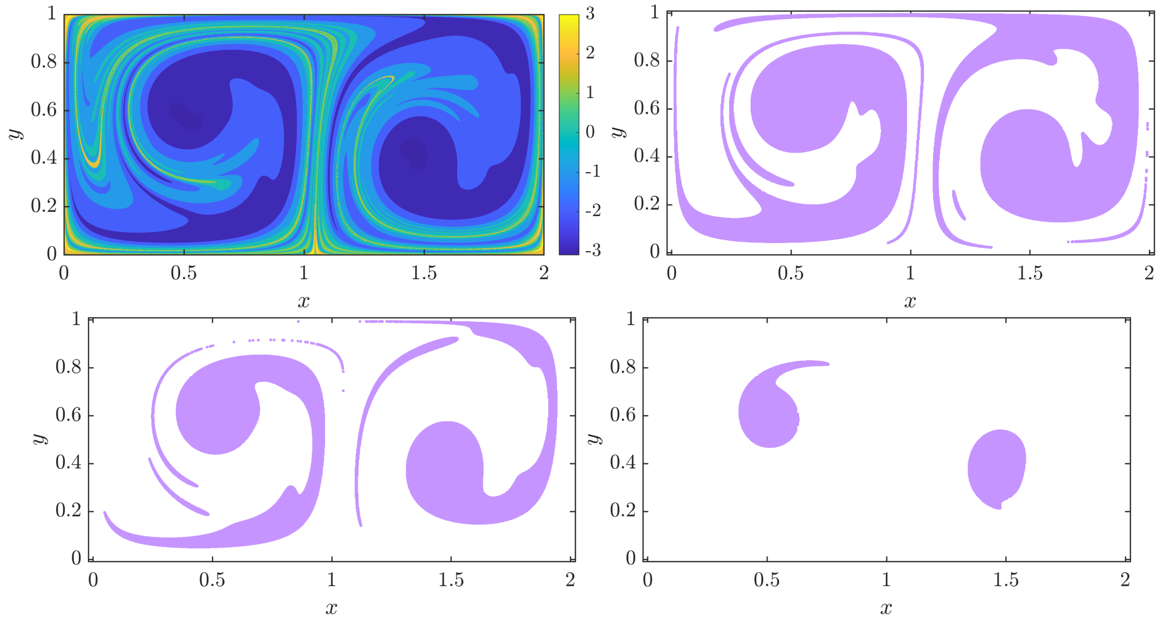

Figure 3.4: The Scaled S^2 for the DG flow, with $\epsilon = 1 \times 10^{-2}$ (top–left), and three robust sets for this flow corresponding to lengthscales $L = 0.2$ (top–right), 0.125 (bottom–left) and 0.075 (bottom–right).

values of the Scaled S^2 within the epicentre of the gyres. It is not unusual for the FTLE field and the Scaled S^2 field to have similar or even identical features, but this is not true for all systems, and will not be true for systems with particularly large amounts of flow turbulence spread across the whole domain (Balasuriya 2020*a*).

As stated earlier, the robust sets indicate more coherent flow regions of our domain Ω by highlighting regions where the Scaled S^2 (the lengthscale of uncertainty) is less than a threshold L. In our first example, we choose $L = 0.2$ and generate the gyres along with thick tails which emanate from these structures as “foetal–shaped” LCSs, along with some thin ridge type structures which wrap around the edges of the left and right gyre chambers. Upon reducing the threshold to 0.125, we maintain the gyres along with thinner tails emanating from them; and when L is reduced to 0.075, only small elliptical structures representing the centres of the gyres remain, though the left hand structure takes a tadpole shape as it still retains some of its tail. These structures correspond to the lowest levels of S^2, and therefore flow trajectories beginning from within these sets exhibit the most self–consistency and robustness against noise or uncertainty within the velocity data representing this system.

We next turn our attention to the transfer operator method, which we compute for this system by dividing Ω into 240×120 rectangular boxes of equal size at both the initial time $t_0 = 0$ and the final time $t_0 + T = 20$, and computing the flow maps of

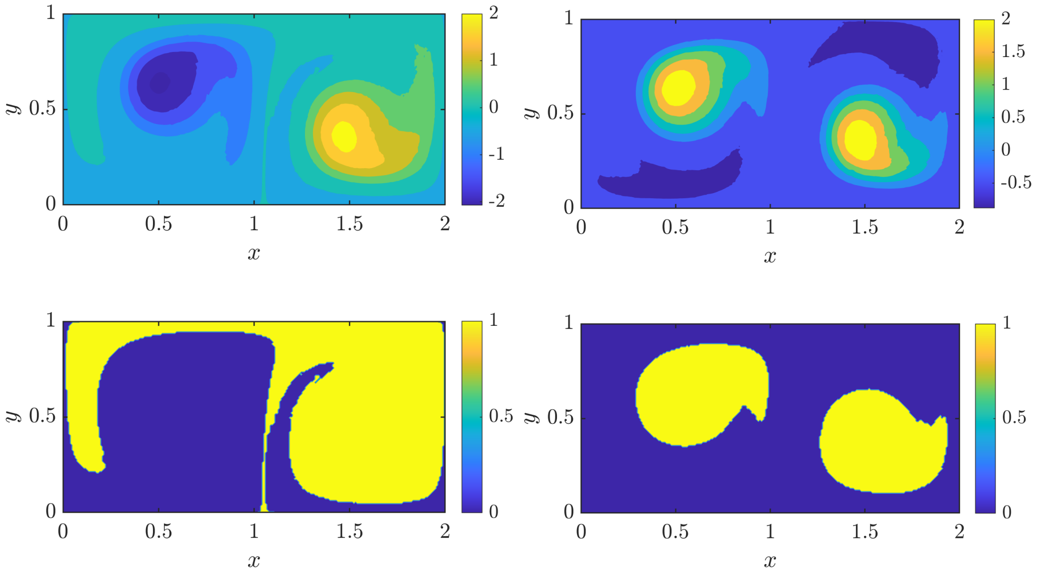

Figure 3.5: Scalar fields of the second (top–left) and third (top–right) right singular vectors of $\mathbf{P}$ for the DG flow and two coherent sets for the flow obtained by dividing Ω along the zero contour lines of each respective field (bottom Figures).

1000 uniformly distributed initial points within each box to estimate the entries of this operator. After carrying out this method, the first six singular vectors of $\mathbf{P}$ calculated were 1, 0.9997, 0.9996, 0.9994, 0.9992 and 0.9991. Figure 3.5 shows the optimal vector $\mathbf{v}_0^*$ computed using both the second and third right singular vectors of $\mathbf{P}$ and two coherent sets obtained from each of these vectors. Even though the spectral gap between the singular values is not remarkably large, we consider the third singular vector of $\mathbf{P}$ based on the results obtained by Tallapragada & Ross (2013). From the two coherent sets obtained from the second singular vector, we more or less define each gyre chamber as a coherent set, with parabolic hook type structures which wrap around each gyre included due to particle flow which crosses the vertical barrier along $x = 1$ into opposite chambers. Using the third singular vector, we instead identify the two gyres as one disjoint coherent set, while the remainder of the domain forms the other coherent set. Both of these results are claimed to be correct, seeing as parcels of flow trajectories sampled within each set will not mix with trajectories from opposing sets. The spectral gap between singular values in most cases is a good guide for determining which singular vector to use to produce these coherent sets (Froyland et al. 2010).

We next consider the dynamic Laplace operator method, computed using the same box configurations as those defined above for the transfer operator method. The first six eigenvalues generated for this operator are 0, -34.3621, -60.4858, -102.9318, -120.9574 and -159.2938. Figure 3.6 shows the scalar fields for the second and third eigenvectors of $\mathbf{\Delta}^D$ and the two coherent sets extracted from each of these eigenvectors. The results

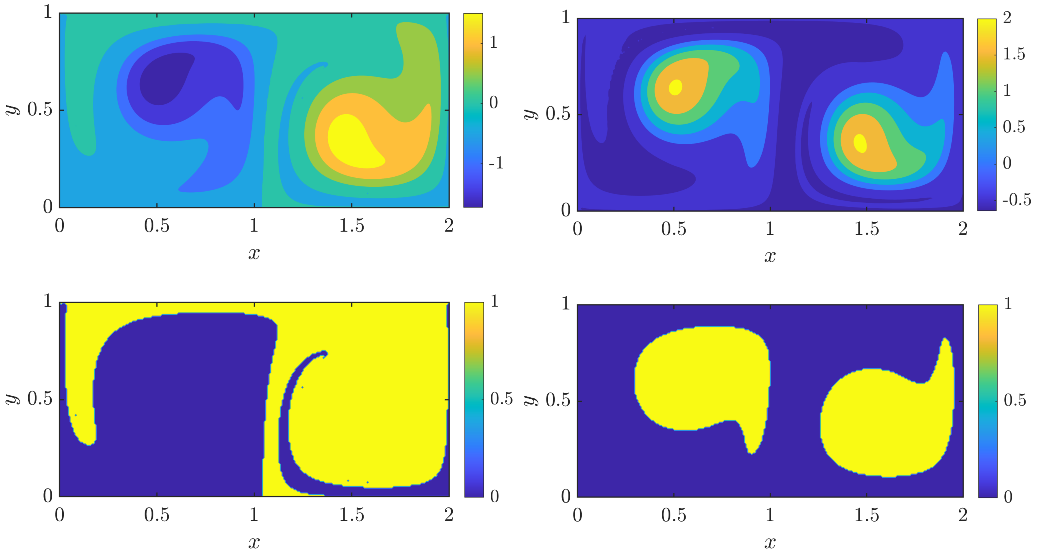

Figure 3.6: Scalar fields of the second (top–left) and third (top–right) eigenvectors of $\mathbf{\Delta}^D$ for the DG flow and two coherent sets for the flow obtained by dividing Ω along the zero contour lines of each respective field (bottom Figures).

for the second eigenvector are consistent with the coherent sets produced in Figure 3.5 using the second right singular vector of $\mathbf{P}$, with each gyre chamber and an adjoining parabolic hook which ventures over into the other chamber forming each set. When we consider the third eigenvector of the operator, we once again manage to isolate the gyres to form a disjoint coherent set separate from the rest of Ω. Again, like with the transfer operator results, both versions of these pairs of coherent sets are claimed to be correct and mixing will not occur between particles with initial positions located within opposing sets (Froyland & Junge 2018).

We next consider the FCM clustering method, using two different a priori guesses for the number of clusters expected to exist within the DG system. We apply the algorithm to the previously defined uniform grid of 1000×1000 initial flow points, and we set the sharpness parameter $m = 1.5$. We start by looking for $K = 2$ clusters, with the resultant membership probability fields and the clusters obtained from a membership probability threshold of 80% on display in Figure 3.7. One coherent cluster is formed approximately from the left gyre chamber with a parabolic shape missing from the left hand end of the domain. This shape forms part of the other coherent cluster, along with the right gyre chamber without a hook shape curve curling around the right gyre from its lower end. Particles which begin within this hook shaped structure seem to have a 50–50 chance of ending up in either cluster, indicating that particle flow within this structure is incoherent with no guarantee of ending up in either of the clusters.

We next increase the number of clusters K to 4, and run the algorithm again with all

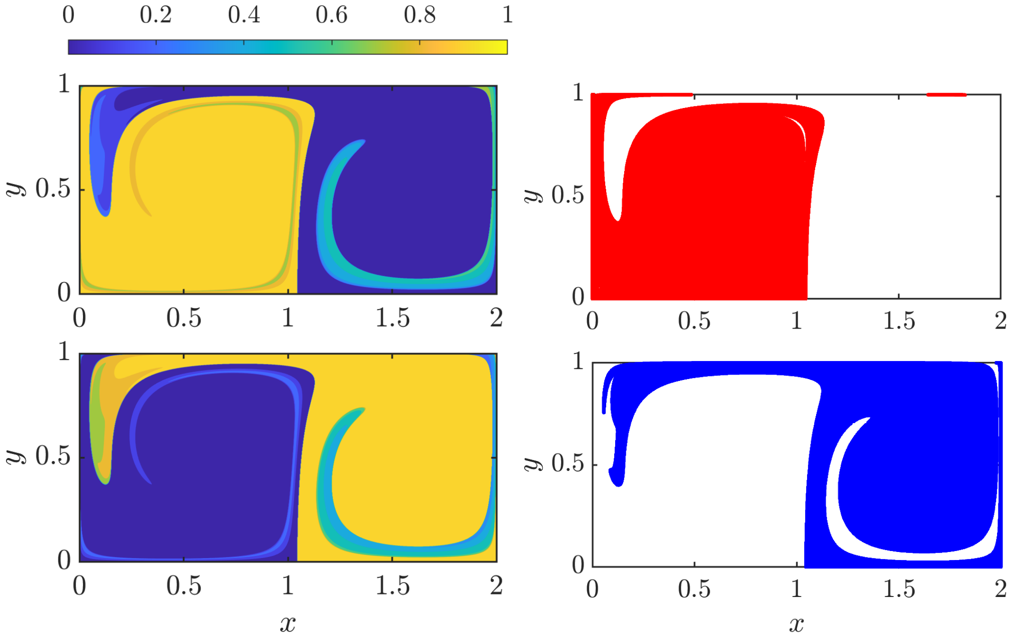

Figure 3.7: Membership probability fields for two FCM clusters of the DG flow (left Figures) and two clusters for this flow obtained at the 80% membership probability level (right Figures).

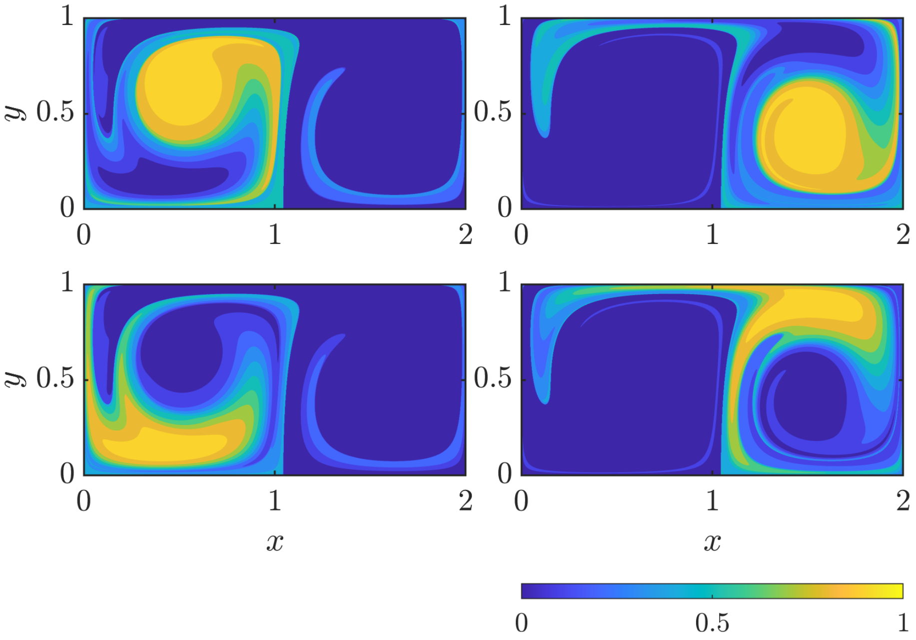

Figure 3.8: Membership probability fields for four FCM clusters of the DG flow.

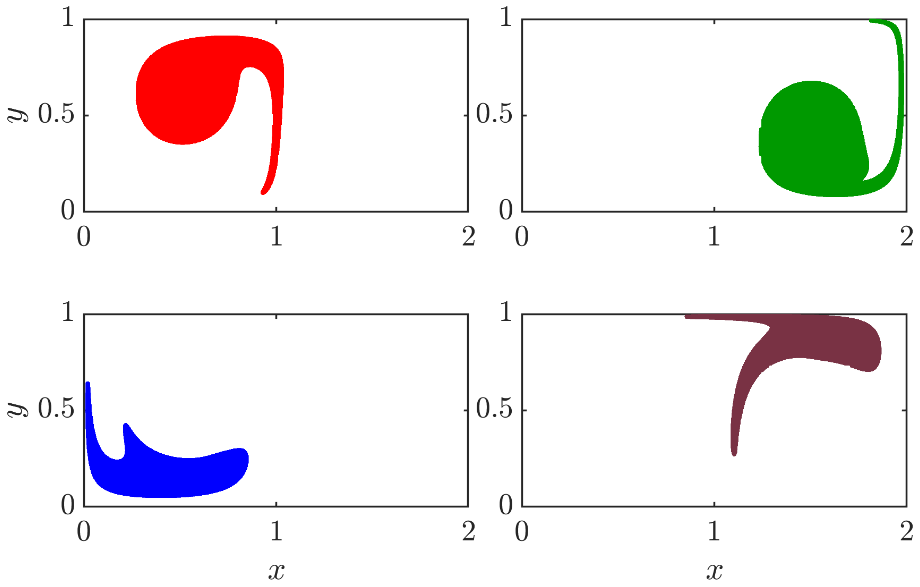

Figure 3.9: Four FCM clusters for the DG flow obtained at the 80% membership probability level.

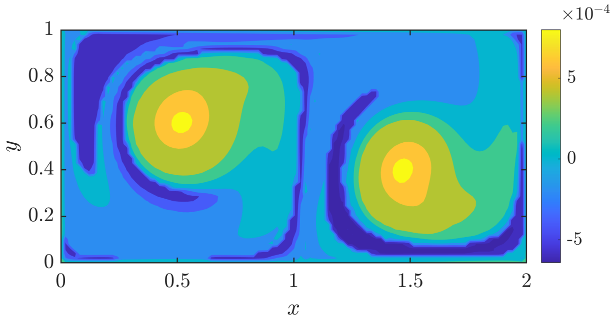

Figure 3.10: The CSC scalar field for the DG flow.

remaining computational parameters remaining consistent. The membership probability scalar fields are on display in Figure 3.8 and Figure 3.9 isolates four coherent clusters within our domain based once again on the 80% membership probability threshold. Here, we produce similar results to those published by Allshouse & Peacock (2015) where, rather than dividing our domain based on the two gyre chambers existent within it, we find smaller and more detailed coherent structures, such as the gyres themselves rather than the full gyre chambers. The four clusters identified at an 80% membership probability level (visible in Figure 3.9) are the two gyres with thin tails emanating from them, and one primarily oblong shaped cluster in each gyre chamber.

Finally, we carry out the coherent structure colouring algorithm on a uniform grid of 100×50 initial points $\mathbf{x}_0 \in \Omega$. We use a uniform grid in most of our computations for this method to avoid the extra step of numerical interpolation to produce a uniform scalar field from scattered data, along with the numerical estimation errors that come with it. Figure 3.10 shows the resultant scalar field for the CSC coefficients. From strongly positive values of these coefficients, we are able to identify the gyres themselves once again, while strongly negative valued CSC coefficients are used to identify hook shaped structures which wrap around each of the gyres, as well as the parabolic shaped structure which sits next to the left hand gyre.

From the results displayed above, all of these detection methods were able to detect at least some of the most definitive structures observable within this system. The most obvious coherent structures within this system are the counter rotating gyres, which were detected using the LAVD, transfer operator, dynamic Laplace operator and CSC methods most comprehensively. We also detect these structures using the S^2 robust sets and the FCM clustering method, though in these cases a thin “tail” emerges from each of the gyres. The gyres and their tails are also identifiable from smaller values of the FTLE and S^2, but are not identifiable from the VLCS foliations. Had we have produced code to generate elliptic VLCSs, there is a good chance that we may have been able to identify these gyres from this method (Karrasch et al. 2014, Onu et al. 2015). In some instances, a gyre chamber (the left or right hand end of the domain Ω) is identified as a coherent structure or set, using the transfer operator or dynamic Laplace operator methods; or by generating two FCM clusters. We do not generate rectangular sets in these cases, rather each one contains a parabolic hook shaped structure from particles which spill over from one gyre chamber to the other, due to the deformation of the vertical barrier positioned along $x = 1$. Also visible within this system are warped versions of this vertical barrier, identifiable from the FTLE, S^2 and VLCS methods. The VLCS algorithms each produce large foliations of closely bunched flow barriers due to the variation in seeding initial points for eigenvector field integration and how this integration is regulated by each algorithm.

3.2 Unsteady Stuart Vortex

The next analytically defined toy model which we concentrate our attention towards is the unsteady Stuart vortex (USV) system, a non–autonomous extension of the well known steady Stuart vortex flow (Stuart 1967, Crowdy 2004). The USV flow takes the form

$$\dot{\mathbf{x}} = \mathbf{u}\left(x, y, t\right) := \frac{1}{a \cosh\left(y\right) + b \cos\left(x + ct\right)} \begin{bmatrix} a \sinh\left(y\right) \\ b \sin\left(x + ct\right) \end{bmatrix}, \tag{3.2}$$

where a is arbitrary, $b = \sqrt{a^2 - 1}$ and we define c as a horizontal translation parameter for the Stuart vortices. If $c = 0$, we obtain the original, autonomous Stuart vortex system characterised by a vortex layer consisting of two Stuart vortices, one being centred at $(-\pi, 0)$ and the other at $(\pi, 0)$; which rotate in place over the whole flow. However if $c \neq 0$, the vortices flow at a consistent rate to the left or right of the domain (depending

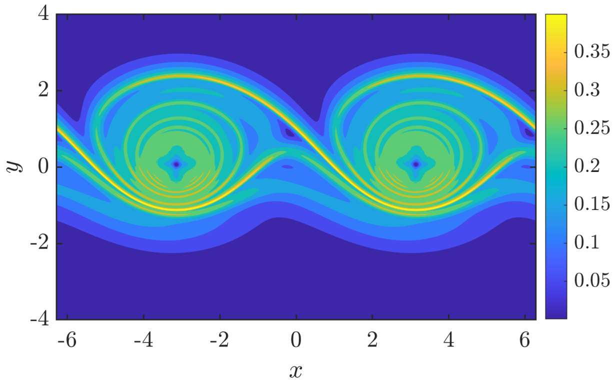

Figure 3.11: The FTLE $\Phi_0^{4\pi}$ for the USV flow.

on the sign of c) as they rotate in a style reminiscent of a real world vortex layer such as a Kelvin–Helmholtz layer (which we will consider later on in this Chapter). For our LCS analysis, we define this system over the domain $\Omega = [-2\pi, 2\pi) \times [-r, r]$ with $r = 4$ (arbitrary), which is periodic in the x–direction. We set $a = 2$, $b = \sqrt{a^2 - 1} = \sqrt{3}$ and $c = -0.5$ (which translates the vortices in the positive x direction); and consider the period of flow ranging from $t_0 = 0$ to $t_0 + T = 4\pi$, which represents one full period of the velocity system in t given our choice of c.

The FTLE for this system was computed from a uniform initial grid of 600 × 400 initial points, with the resultant scalar field produced on display in Figure 3.11. The key structures which can be extracted from this field are sinusoidally shaped maximal ridges tucked under each vortex, which curl to the left of each vortex, traverse towards the other vortex producing a barrier between the two, and wrap around the other vortex until it sits just above it. One can also detect another sinusoidal curve which supports the vortex layer from below, though this is a ridge that is mostly of a medium level of FTLE. The lowest values of the FTLE are observable at the exact central point of each vortex, small patches tucked between the vortex support wave and the vortex separation ridge (approximately half way between the centres of each vortex) and the regions of the domain above and below the vortex layer, where the flow velocity has almost zero magnitude.

We next consider the VLCS detection method, generated from Lagrangian flow data defined over a uniform initial grid of 600 × 400 points once more. The results obtained from each algorithm are on display in Figure 3.12. In VLCS–A, we randomly sample 500 candidate points from $\mathcal{G}_0$, set $\mathcal{L}_f = 2$ and choose the strainlines corresponding to $\mathcal{L}_{\min} = 2.5$. In VLCS–B, we choose a separation radius $\rho = 2\pi/3 \approx 2.1$, and set $\mathcal{L}_{\max} = 30$. We reuse the parameter choices for ρ and $\mathcal{L}_{\max}$ in VLCS–C, and choose $\mathcal{L}_f = 0.5$. From the results in Figure 3.12, VLCS–A struggles to identify a VLCS of considerable length, which is most likely due to the presence of thin ridges within the acceptable set $\mathcal{G}_0$ (not shown) which hyperbolic strainlines slide off of very quickly. By contrast, VLCS–B produces a

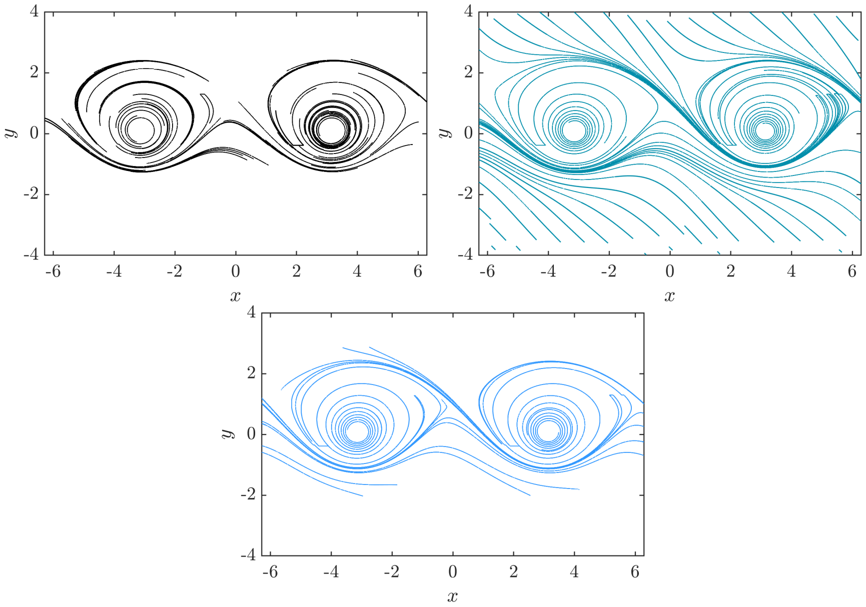

Figure 3.12: Hyperbolic VLCSs for the USV flow computed using the VLCS–A (top–left), VLCS–B (top–right) and VLCS–C (bottom) algorithms.

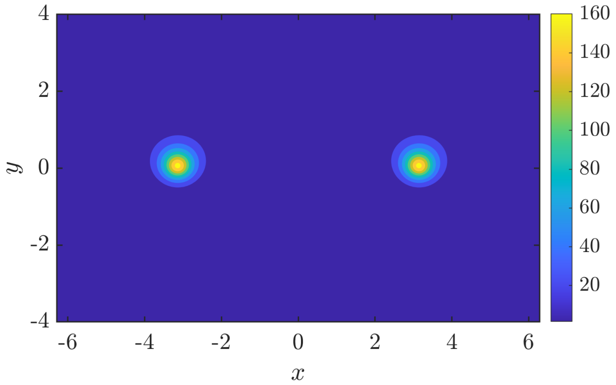

Figure 3.13: The LAVD $\Theta_0^{4\pi}$ for the USV flow.

large number of VLCSs which do coincide with the maximal and more medium level ridges of the FTLE field seen in Figure 3.11, but also produces a large number of VLCSs in the stagnant flow regions of Ω above and below the vortex layers. The results of VLCS–C are much more useful, with hyperbolic VLCSs produced of a reasonable length which do not enter regions of the domain where particle flow is more robust or non–existent.

Next, we consider the LAVD for this system computed on the same initial grid of 600×400 uniformly spaced points by solving the extended ODE system (2.7). The resultant scalar field for the system is shown in Figure 3.13. As expected, only two

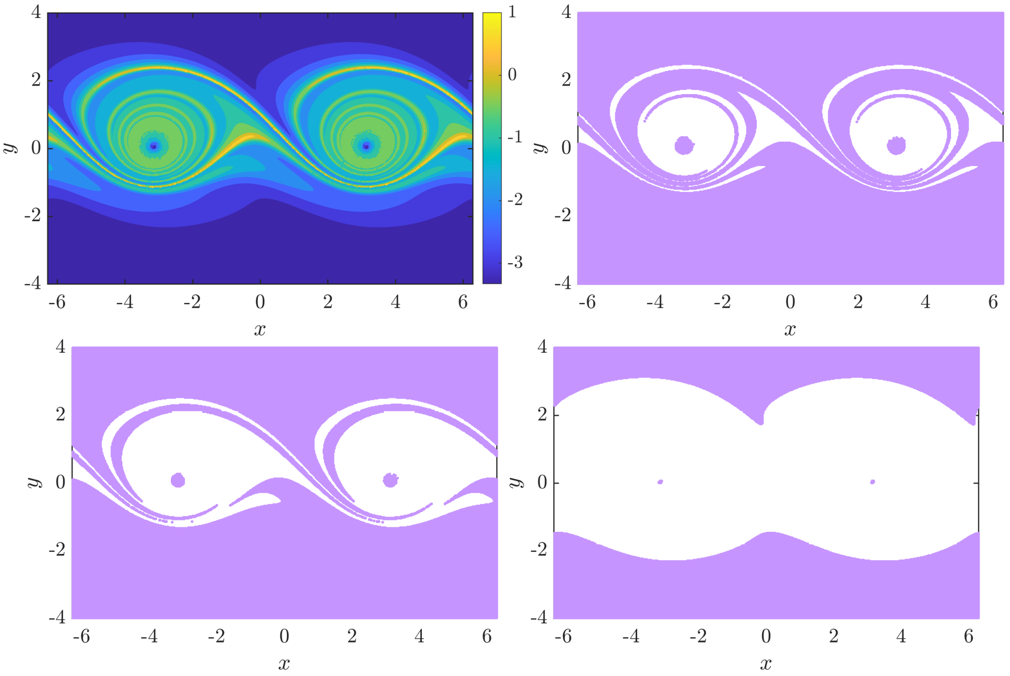

Figure 3.14: The Scaled S^2 for the USV flow, with $\epsilon = 1 \times 10^{-2}$ (top–left), and three robust sets for this flow corresponding to lengthscales $L = 0.35$ (top–right), 0.25 (bottom–left) and 0.05 (bottom–right).

small circles representing the epicentres of the Stuart vortices are found to be the most rotationally coherent structures within the system. The LAVD is low throughout the remainder of the scalar field, as there is no rotational coherence observable within the rest of our flow domain.

The next method to consider is the S^2 method, where this quantity is calculated again using the same initial grid of 600 × 400 uniformly spaced points and by expanding the grid of advected points at time $t_0 + T$ to compute the entries for the matrix Λ. The scalar field for the Scaled S^2 is shown in Figure 3.14, with $\epsilon = 1 \times 10^{-2}$ like with the DG flow discussed in the previous Section; along with robust sets produced from three different lengthscale tolerances L, taking values of 0.35, 0.25 and 0.05 in non–dimensional units. Like with the DG flow, the Scaled S^2 field looks almost identical to the FTLE field shown in Figure 3.11 bar some minor differences, such as the round parcels of low FTLE centred approximately at the points $(0, 0.5)$ and $(2\pi, 0.5)$ not being found in the Scaled S^2 field. Letting $L = 0.35$ for the first robust set, most of the domain Ω seems to be included within this set, except for annulus shaped structures within the vortices and surrounding their centres; triangular shaped fluid parcels located to the right of each vortex, and thin ridges which wrap around each vortex in a shape similar to the maximal ridges obtained from the FTLE and Scaled S^2 fields. Reducing the lengthscale threshold to 0.25 makes a

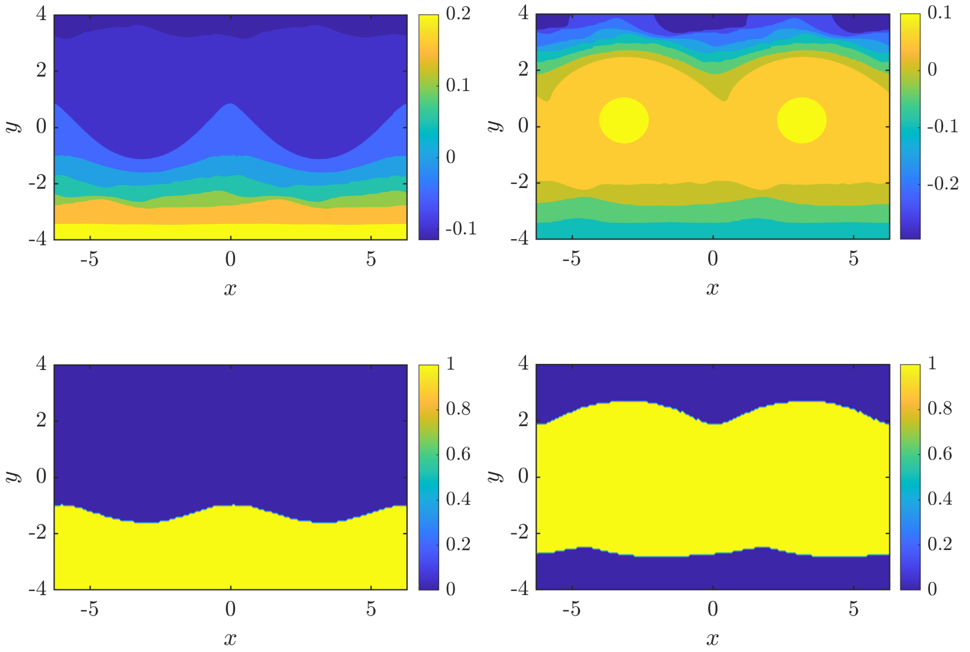

Figure 3.15: Scalar fields of the second (top–left) and third (top–right) right singular vectors of $\mathbf{P}$ for the USV flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

minor change to the coverage of the robust set, however when we reduce the threshold to 0.05, only the stagnant flow surrounding the vortex layer is included within the set, along with very tiny circular sets positioned at the centres of each of the Stuart vortices.

We next focus on computing the transfer operator for this system, achieved by dividing Ω at time t_0 into 180×120 rectangular boxes of equal size, and $\widetilde{\Omega} = \mathbf{F}_{t_0}^{t_0+T}(\Omega)$ at time $t_0 + T$ into 210×140 rectangular boxes also of equal size. Seeing as Ω is periodic in x, we map particles which spill over one horizontal end of the domain back into the domain on the other side, while extending the domain slightly in the y direction to estimate $\widetilde{\Omega}$. 1000 points are uniformly sampled within each box and flowed forward to estimate the entries of $\mathbf{P}$. Upon completion of the algorithm, the first six singular values produced for $\mathbf{P}$ are 1, 0.9999, 0.9998, 0.9997, 0.9996 and 0.9996. The optimal vector $\mathbf{v}_0^*$ computed using both the second and third right singular vectors of $\mathbf{P}$ and the corresponding coherent sets produced from each vector are featured in Figure 3.15. Using the second right singular vector of $\mathbf{P}$, we obtain the upper half of the domain and the entire Stuart vortex layer as one coherent set with the remaining lower portion of the domain forming the other coherent set. If we use the third singular vector instead, we obtain the vortex layer and surrounding flows as one coherent set, while the stagnant flow regions at the upper and lower edges of the domain form a disjoint second coherent set.

We next consider the dynamic Laplace operator method, using the same computational

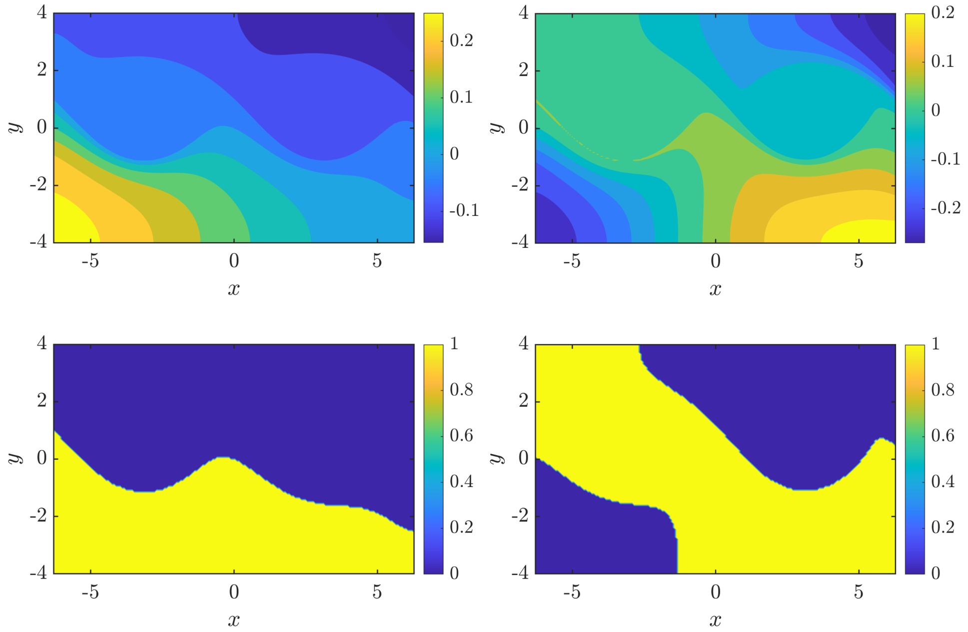

Figure 3.16: Scalar fields of the second (top–left) and third (top–right) eigenvectors of $\boldsymbol{\Delta}^D$ for the USV flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

parameters defined above for the transfer operator method including the application of periodic boundary conditions in the x direction. In this case, the first six eigenvalues produced for the dynamic Laplace operator came out to be -0.0000, -0.1106, -0.2266, -0.3453, -0.4756 and -0.6237. In Figure 3.16, we display the second and third eigenvector scalar fields for the dynamic Laplace operator and the two coherent sets produced from each of these fields. The second eigenvector has divided the domain Ω in a similar fashion to the transfer operator method (see Figure 3.15), while if we consider the third eigenvector, the coherent sets obtained each contain one of the Stuart vortices, and portions of the flow domain subsequently advected along with these vortices as they move through the system.

Next, we use the FCM Clustering method to divide Ω for this system into $K = 2$ coherent clusters. We again seed 600×400 initial points in a uniform grid, and set the sharpness parameter $m = 1.5$. The membership probabilities for these clusters are shown in Figure 3.17, along with two FCM clusters isolated at an 80% probability level. From these results, we have been able to divide the flow domain Ω in half in a similar fashion to the coherent sets obtained from the second eigenvector of $\boldsymbol{\Delta}^D$ (see Figure 3.16). The first cluster contains the upper half of the domain including the vortex layer, with the two vortices separated by a wave–shaped structure which sits between the vortices. The other cluster contains the lower half of the domain, with its upper boundary approximately

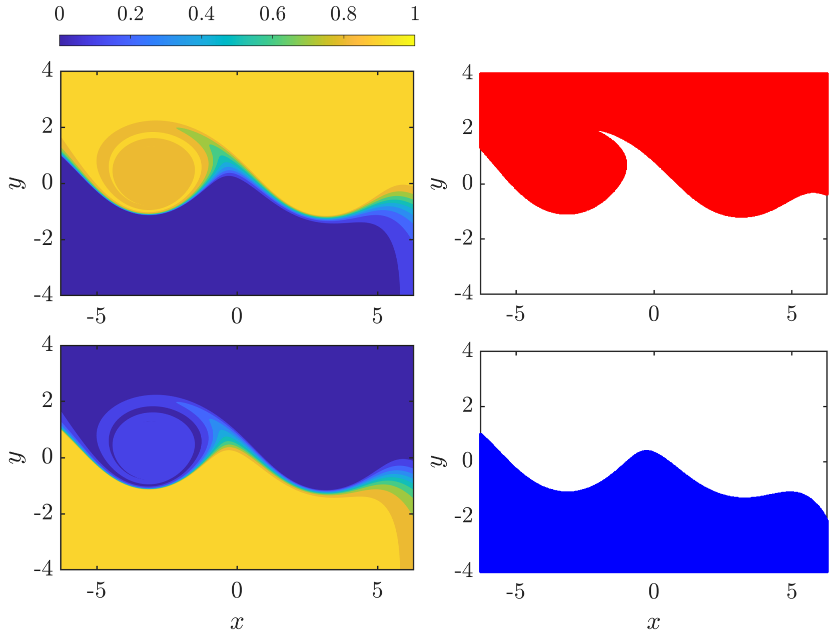

Figure 3.17: Membership probability fields for two FCM clusters of the USV flow (left Figures) and two clusters for this flow obtained at the 80% membership probability level (right Figures).

forming the shape of the vortex layer support curve, until it begins to deviate away from this support barrier in similar fashion to the zero contour level curve of the dynamic Laplace operator's second eigenvector field.

We then repeat the FCM Clustering method, this time seeking $K = 6$ clusters. The membership probability scalar fields for these six clusters are shown in Figure 3.18, and six coherent clusters isolated at the 80% probability level are shown in Figure 3.19. By choosing $K = 6$ clusters, we are able to identify the Stuart vortices and their immediate surrounds as coherent structures, as noted by the red and blue clusters in 3.19. We also extract four more coherent shapes located in each corner of the domain. These coherent shapes hardly move, as they are located above and below the Stuart vortex layer where particle flow is stagnant. However, as they retain their shape they are therefore deemed coherent shapes.

Finally, in Figure 3.20 we consider the coherent structure colouring method applied to this system, on a uniform grid of 90 × 60 initial points. From the positive valued CSC coefficients, we are able to identify the centres of the Stuart vortices along with their elliptic surrounds; combined with the top half of the domain and hook shaped structures which curl into the elliptic regions containing the vortices. From the negative valued CSC coefficients, we identify the sinusoidal wave type structure which supports the vortex layer along with the lower half of the domain. We also identify a hook type structure which

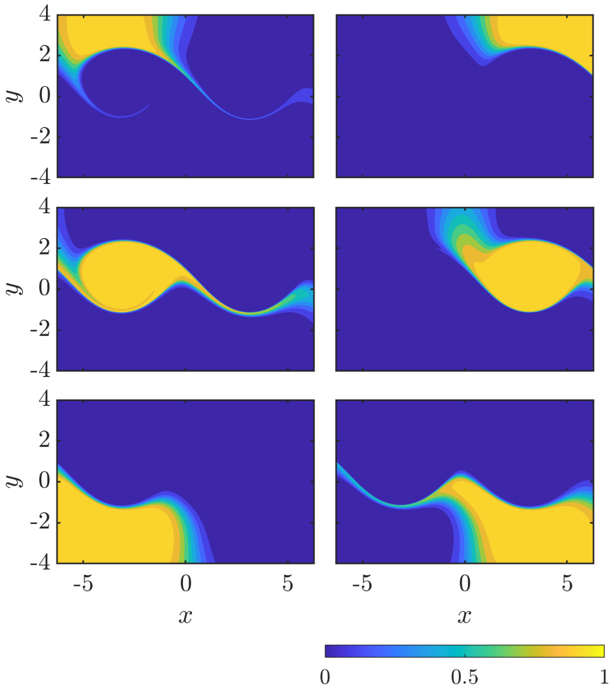

Figure 3.18: Membership probability fields for six FCM clusters of the USV flow.

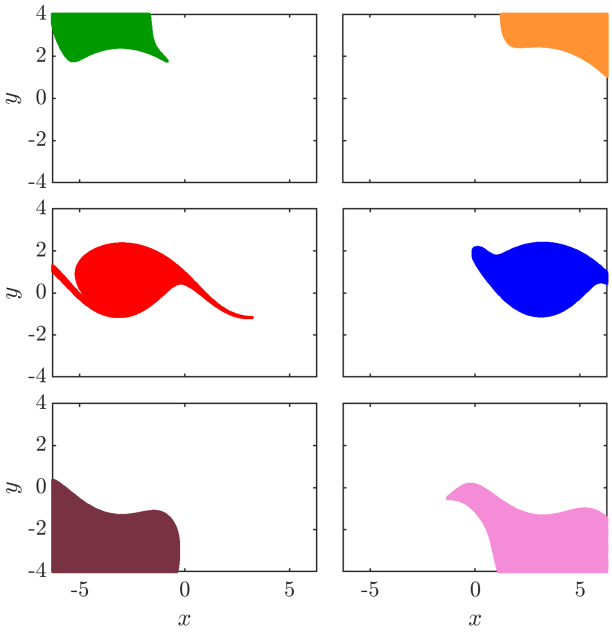

Figure 3.19: Six FCM clusters for the USV flow obtained at the 80% membership probability level.

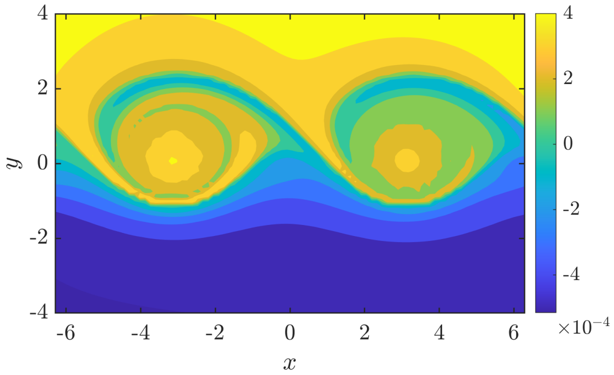

Figure 3.20: The CSC scalar field for the USV flow.

folds in towards the centre of the Stuart vortex as the flow progresses.

By applying a time–dependent horizontal translation to the Stuart vortex velocity system in the fashion seen in (3.2), the ordinarily stationary vortices form a transport layer that is also defined by parcels of flow particles dragged along by the vortices as they rotate and move through the system. The Stuart vortices themselves are still identifiable from the LCS analysis undertaken above, featuring prominently in the CSC scalar field, S^2 robust sets, and the LAVD field (in the latter case, these vortices are all that is identified). Low values of the FTLE and Scaled S^2 fields identify only the centres of these vortices, but their boundaries can be identified from medium values of these quantities, along with some of the hyperbolic VLCSs identified which tend towards the centres of these objects. Key structures that are also identifiable within this field are elliptic shaped structures formed from the vortices and the surrounding flow which gravitates towards their centres. These are identifiable from the CSC and FCM Clustering methods, and are outlined by the FTLE, VLCSs, and S^2. The latter three methods are also useful in determining flow barriers which separate the two vortices within the layer, as well as a sinusoidal wave which is tucked under the vortex layer and acts as a support barrier for these vortices. The transfer operator and dynamic Laplace operator methods, on the other hand, only manage to divide this flow domain into two sets in different ways based on the evolution of the vortex layer. The disparities observed between the results obtained from these two methods is most likely as a result of a lack of significant spectral gap between the second and third singular values of $\mathbf{P}$, and the second and third eigenvalues of $\boldsymbol{\Delta}^D$. If a lack of a spectral gap between these singular values/eigenvalues is observable, this indicates that the flow domain in question cannot be cohesively split into two coherent sets, and such disparities may therefore arise between the results from each of these methods. We note a similar observation with the Taylor–Green vortex system in Appendix A.2, which is rectified upon selecting a subsequent singular vector/eigenvector of these operators where a larger spectral gap can be identified.

3.3 Bickley Jet

The third and final analytically defined toy flow which we will consider in this Chapter is the Hamiltonian Bickley Jet (3.3), a system sold as an analytically defined, quasi–realistic flow system with characteristics that mimic a stratospheric polar vortex system with an impenetrable zonal jet (del Castillo-Negrete & Morrison 1993, Rypina et al. 2007) which is used regularly for LCS analysis in the literature. The version of this system which we will consider is the same as that used by Froyland & Junge (2018) and Froyland et al. (2019); which takes the form

$$\dot{\mathbf{x}} = \mathbf{u}(x,y,t) := \begin{bmatrix} u(x,y,t) \\ v(x,y,t) \end{bmatrix}, \tag{3.3}$$

where

$$u(x,y,t) = U_0 \operatorname{sech}^2\left(\frac{y}{L_0}\right) + 2U_0 \operatorname{sech}^2\left(\frac{y}{L_0}\right) \tanh\left(\frac{y}{L_0}\right) \sum_{i=1}^{3} A_i \cos(k_i(x - c_i t)),$$

$$v(x,y,t) = -U_0 L_0 \operatorname{sech}^2\left(\frac{y}{L_0}\right) \sum_{i=1}^{3} A_i \sin(k_i(x - c_i t)),$$

with parameters $U_0 = 62.66\text{ms}^{-1} = 5.4138\text{Mmd}^{-1}$, $L_0 = 1.770\text{Mm}$, $A_1 = 0.0075$, $A_2 = 0.15$, $A_3 = 0.3$, $k_i = (2i)/r_e$, $i = 1,2,3$, $r_e = 6.371\text{Mm}$, $c_2 = 0.205U_0$, $c_3 = 0.461U_0$ and $c_1 = c_3 + ((\sqrt{5}-1)/2)(k_2/k_1)(c_2 - c_3) \approx 0.1446U_0$. The system is defined over the flow domain $\Omega = [0,20) \times [-3,3]$, which is defined in units of megametres (Mm) and is periodic in the x-direction. We consider the time interval which ranges from $t_0 = 0$ to $t_0 + T = 40$ days (d). There are many different versions of this system which can be conceptualised based on the time interval or parameters chosen.

The key feature of this system is a sinusoidal jet type structure centred along $y = 0$ Mm which forms a solid, impermeable border between two large vortex layers. Each of these layers contain a set of equally spaced vortices which retain their shape and move through the domain in a consistent fashion across the interval of flow. This is in contrast to the flow behaviour observable between the vortices, which is considerably unpredictable in nature. The vortices exist in this system based on the parameters chosen above, but can be broken apart or dissolved by choosing different values for the A_i parameters (Rypina et al. 2007, Froyland et al. 2010). Like with the DG system, the Bickley Jet flow is a popular analytical flow model for use in LCS detection in the literature (Haller & Beron-Vera 2012, Hadjighasem et al. 2017, Schlueter-Kuck & Dabiri 2017), particularly when it comes to testing the transfer operator or dynamic Laplace operator methods as the two vortex layers in this system form two very well defined almost invariant coherent sets (Froyland et al. 2010, Froyland 2013, Froyland & Junge 2018, Froyland et al. 2019).

We start once more with the FTLE for this system, generated on a uniform grid of 900×300 initial points within our domain Ω and displayed in Figure 3.21. Maximal ridges

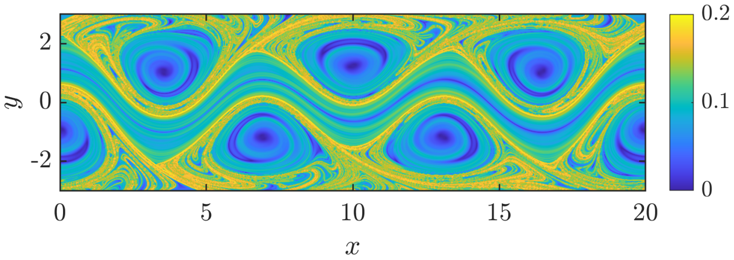

Figure 3.21: The FTLE Φ_0^{40} for the Bickley Jet flow.

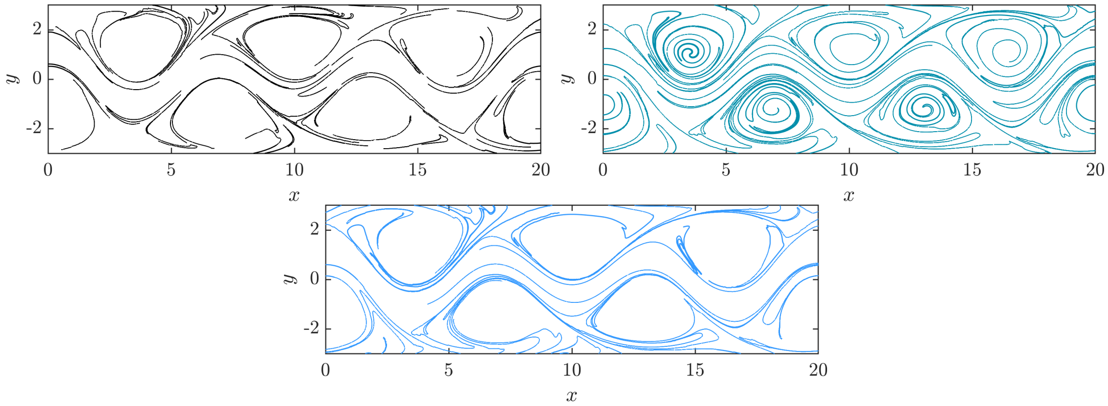

Figure 3.22: Hyperbolic VLCSs for the Bickley Jet flow computed using the VLCS–A (top–left), VLCS–B (top–right) and VLCS–C (bottom) algorithms.

of the FTLE are observable along the edges of the zonal jet as expected, as well as within the vortex layers wrapping around the vortices in parabolic and hook type shapes, but not strictly forming boundaries of these vortices. The resolution of the ridges observable within the vortex layers are grainy in resolution due to the unpredictable flow behaviour observable within these vortex layers, with small parcels of low FTLE observable between these ridges. The lowest FTLE values are observable along the zonal jet itself, along with the centres of the coherent vortices and crescent–shaped objects which surround these centres.

The hyperbolic VLCS results from our three detection algorithms are illustrated in Figure 3.22. We compute the Lagrangian flow map and its corresponding Cauchy–Green eigenbasis for 900×300 initial points organised in the same fashion as done above to carry out the FTLE method. For VLCS–A, we randomly select 500 initial points for eigenvector field integration from the acceptable set $\mathcal{G}_0$, set the allowance parameter $\mathcal{L}_f = 2$ and dispose of all strainlines with an arc length less than 2.5 Mm. In VLCS–B, we set the local radius $\rho \approx 0.556$ Mm and $\mathcal{L}_{\max} = 30$. These parameters are maintained in VLCS–C, however we reduce the allowance parameter $\mathcal{L}_f$ to a value of 0.5. The VLCS–A algorithm produces a significant number of small hyperbolic VLCS segments which attempt to form

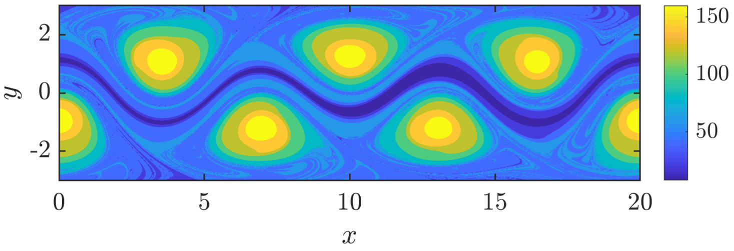

Figure 3.23: The LAVD Θ_0^{40} for the Bickley Jet flow.

the boundaries of the Bickley Jet vortices and the shapes of the sinusoidal ridges which form above and below the zonal jet; which is likely due to the presence of thin ridges within the acceptable set $\mathcal{G}_0$ (not shown) which hyperbolic strainlines slide off of quickly. By contrast, the hyperbolic VLCSs produced from VLCS–B are longer and cover a much greater area of the domain, including the centres of the coherent vortices. It can be argued that the VLCSs produced using VLCS–C form the middle ground between the results of the previous two algorithms, with borders of the vortices and the zonal jet identifiable, but only two structures observable within the zonal jet and no structures seen to be present inside the vortices.

The LAVD scalar field for this system can be seen in Figure 3.23, and was computed by solving the extended ODE system (2.7) for the same uniform grid of 900×300 initial points in our domain Ω as those used for the previous two methods. As anticipated, the largest values of the LAVD are observable within the clearly identifiable Bickley Jet vortices. The lowest LAVD values are observable along the zonal jet core itself, seeing as this structure is not associated with any rotational coherence, but rather extensional coherence as a sinusoidal flow barrier. We observe plenty of ridges of low to medium LAVD value between the vortices, which can most likely be explained by high Lagrangian shear and low rotational coherence.

We now turn our attention to the S^2 method once again, where the relevant quantity is computed on the same initial grid of 900×300 uniformly spaced points and this time using the star grid method to evaluate the backwards time flow gradients due to memory constraints. In Figure 3.24, we display the Scaled S^2 scaled by $\epsilon = \sqrt{hv_r}$, where $h \approx 2.22 \times 10^{-2}$ Mm and $v_r \approx 1.4 \times 10^{-3}$ $\text{Mm} \cdot \text{d}^{-1}$, thereby letting $\epsilon = 5.6 \times 10^{-3}$ $\text{Mm} \cdot \text{d}^{-\frac{1}{2}}$. We also display three robust sets for this system in this Figure, using the same value of ϵ and lengthscale thresholds L of 0.5, 0.3 and 0.1 Mm. Once again, the Scaled S^2 field is qualitatively similar to the FTLE field (see Figure 3.21) bar again some exceptions, such as the absence of the crescent shaped structures within the vortices corresponding to low values of the FTLE from the Scaled S^2 field, and the presence of medium level ridges in the Scaled S^2 field which curl around the vortex centres, and which either aren't observable in the FTLE field or are only partially observable.

Beginning with the lengthscale $L = 0.5$ Mm, the Robust Set for this flow is able to

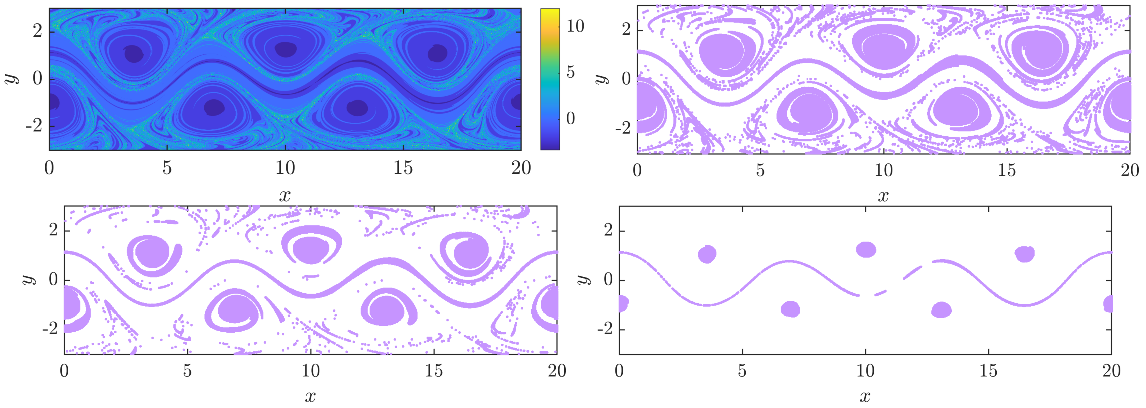

Figure 3.24: The Scaled S^2 for the Bickley Jet flow, with $h \approx 2.22 \times 10^{-2}$ Mm and $v_r \approx 1.4{\times}10^{-3}$ Mm$\cdot$d^{-1} (top–left); along with three robust sets for this flow corresponding to lengthscales $L = 0.5$ Mm (top–right), 0.3 Mm (bottom–left) and 0.1 Mm (bottom–right).

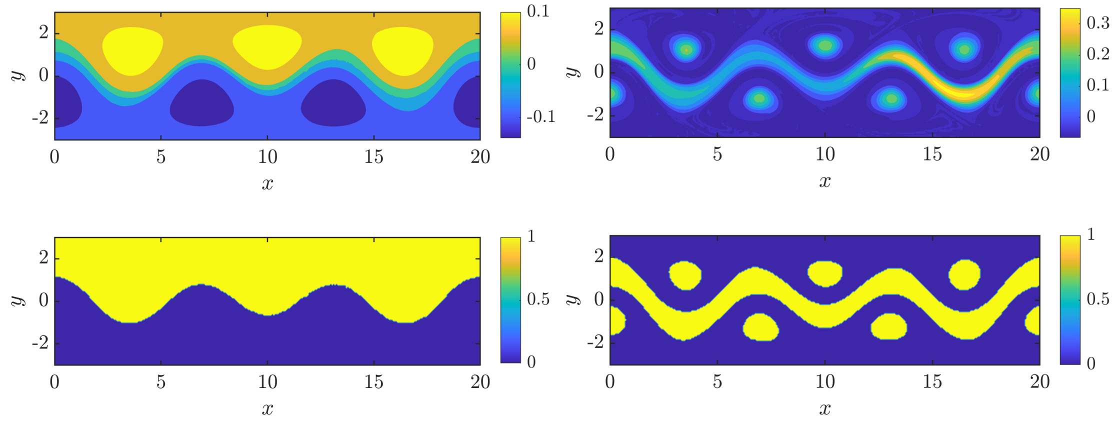

Figure 3.25: Scalar fields of the second (top–left) and ninth (top–right) right singular vectors of $\mathbf{P}$ for the Bickley Jet flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

capture the zonal jet as a coherent structure as well as the Bickley jet vortices, though with a thin line curling towards the centre of each vortex excluded. Also captured by the set are sparse and incoherent clusters of points which appear to surround each of the coherent vortices. As we reduce the lengthscale L to 0.3 Mm, the zonal jet shrinks slightly and the vortices have decreased in size, with an increase in the deviation observable between the centres of the vortices and their fringes. Eventually as L is reduced to 0.1 Mm, all that remains within the robust set are the centres of the Bickley jet vortices and a very thin zonal jet with noticeable gaps.

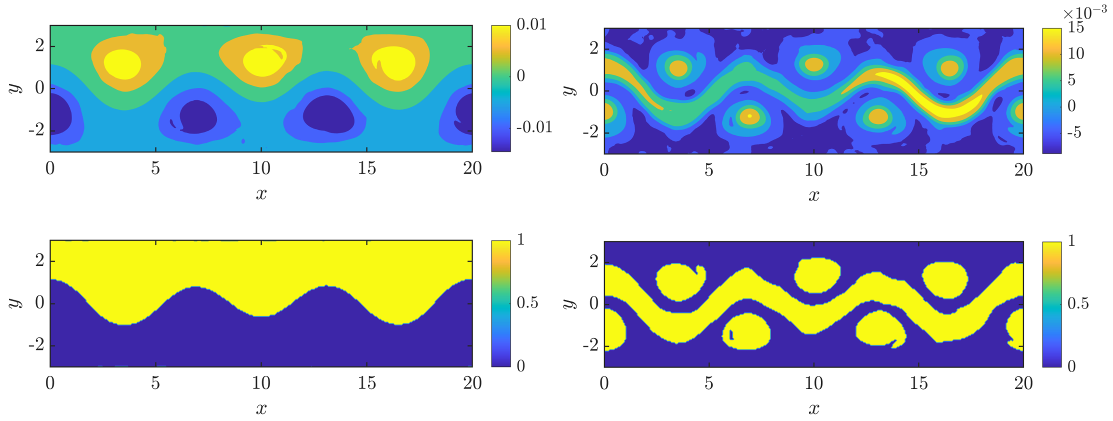

Figure 3.26: Scalar fields of the second (top–left) and ninth (top–right) eigenvectors of $\mathbf{\Delta}^D$ for the Bickley Jet flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

We next consider the transfer operator method, which we execute on this flow by dividing the domain Ω at time t_0 into 300×90 rectangular boxes of equal size, and seeding 1000 uniformly distributed initial flow points within each of these boxes. In determining the domain at the final flow time $t_0 + T$, we treat Ω as periodic in x, extend the domain in y from $[-3, 3]$ to $[-4.5, 4.5]$ and define 300×120 rectangular boxes of equal size. After estimating the entries for the transfer operator $\mathbf{P}$, we compute the first ten singular values for this matrix which have turned out to equal 1, 0.9996, 0.9984, 0.9984, 0.9984, 0.9983, 0.9980, 0.9976, 0.9954 and 0.9940. Figure 3.25 shows the scalar fields representing the optimal vectors generated from the second and ninth right singular vectors of $\mathbf{P}$ and two coherent sets extracted from each of these vectors. We can clearly identify two coherent sets from the second singular vector field, which are the upper and lower vortex layers separated by the zero contour which (as expected) lies along the zonal jet. From the field of the ninth right singular vector, we identify one coherent set consisting of the zonal jet itself along with all of the coherent vortices; and the remainder of the domain as the other coherent set.

We execute the dynamic Laplace operator method for this system using the same set of parameters defined above for the transfer operator method. The first ten eigenvalues produced for the dynamic Laplace operator were calculated to be 0, -3.9665, -6.4184, -6.4748, -6.5077, -6.5522, -7.9136, -13.9962, -17.4218 and -22.4814. Figure 3.26 shows different pairs of coherent sets obtained from both the second and ninth eigenvectors of $\mathbf{\Delta}^D$. The two coherent sets obtained in Figure 3.26 from the second eigenvector are practically identical to those obtained from the second right singular vector of $\mathbf{P}$ (see Figure 3.25), with the domain divided into two coherent vortex layers separated by the zonal jet. The ninth eigenvector of $\mathbf{\Delta}^D$ also produces the same coherent sets as the ninth singular vector of $\mathbf{P}$ (see Figure 3.25), albeit with a minor lack of smoothness observable

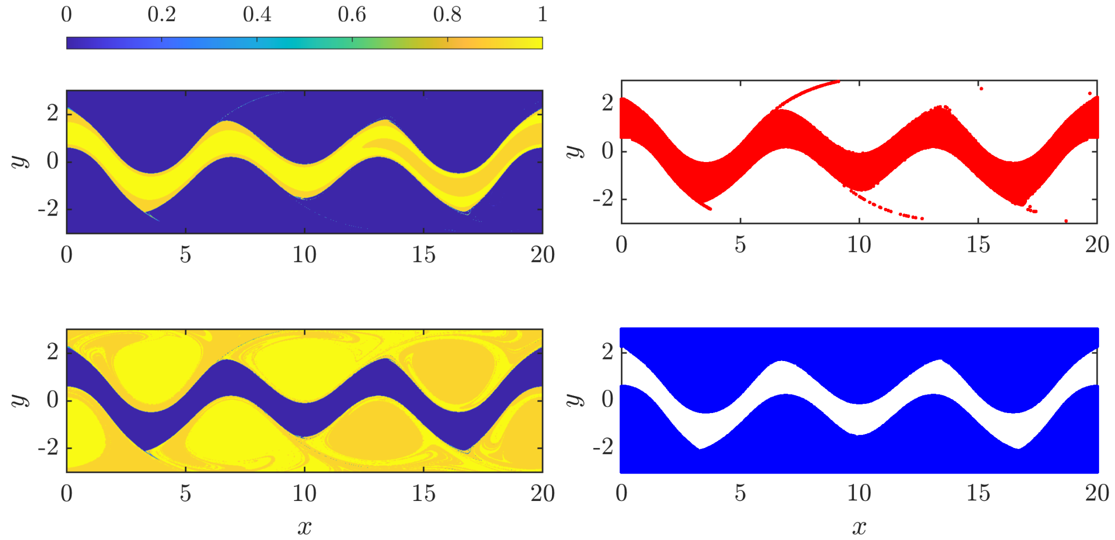

Figure 3.27: Membership probability fields for two FCM clusters of the Bickley Jet flow (left Figures) and two clusters for this flow obtained at the 80% membership probability level (right Figures).

in the eigenvector scalar field.

Next, in Figure 3.27, we apply the FCM clustering method to this flow seeking $K = 2$ clusters. We use a uniform grid of 900×300 initial points and set the sharpness parameter $m = 1.1$. For this number of clusters, the method shows no difficulty in dividing the domain into two coherent clusters, with the first being the zonal jet (along with some excess lines which leak into the vortex layers), and the second comprising of both vortex layers in a disjoint fashion. We then increase the number of clusters K to 3, with Figure 3.28 showing the membership probability fields and Figure 3.29 showing the clusters extracted at the 80% threshold. The method demonstrates no difficulty in identifying the zonal jet as a coherent cluster once more, with the other two clusters produced from the greater elliptic regions surrounding the coherent vortices, and the partitions of the vortex layers lying between these vortices.

Finally, we apply the CSC method to the Bickley jet flow using a uniform initial grid of 80×80 points to produce the scalar field shown in Figure 3.30. Strongly positive CSC coefficients from this field enable us to identify the coherent Bickley jet vortices once more, while the strongest negative values of the CSC coefficients help visualise the zonal jet which divides the domain into two coherent vortex layers. Between each of the coherent vortices, we see an incoherent scatter of CSC coefficient values which lie close to zero, which may indicate that particle movement in these regions is rapid and highly incoherent.

To some degree, each of the methods detailed above has been able to identify the

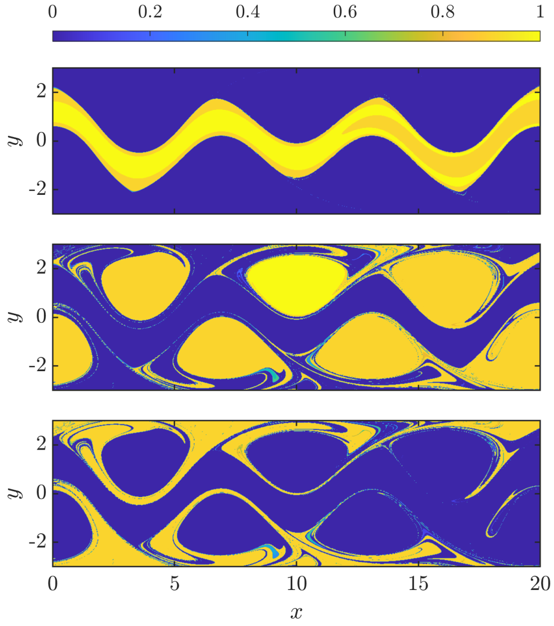

Figure 3.28: Membership probability fields for three FCM clusters of the Bickley Jet flow.

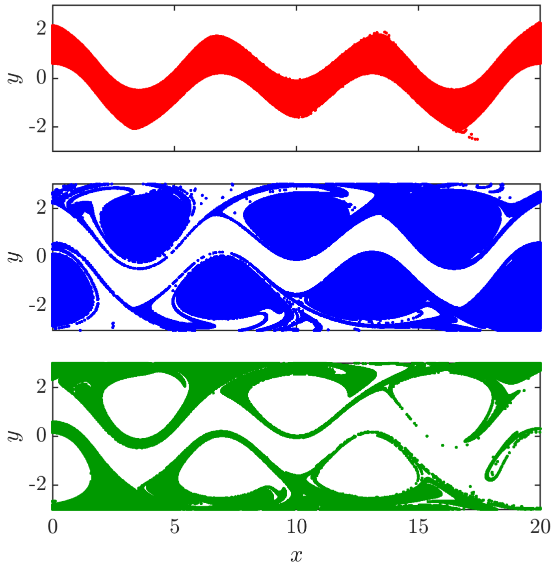

Figure 3.29: Three FCM clusters for the Bickley Jet flow obtained at the 80% membership probability level.

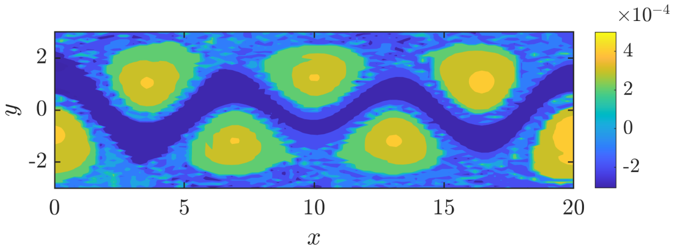

Figure 3.30: The CSC scalar field for the Bickley Jet flow.

same collection of structures within the Bickley jet system. It must be noted that this is a rare occurrence when investigating Lagrangian coherent structures in general, as the Bickley jet is a highly idealised flow with sturdy, coherent flow objects highly unlikely to be detected within realistic flow systems. The definitive coherent structure of this system is the zonal jet which separates the flow domain into two vortex layers and is itself coherent by keeping flow particles within a sinusoidal fragment of the domain. All of the methods were able to identify this structure either as a boundary between the vortex layers (FTLE, VLCSs, dynamic Laplace operator, transfer operator, S^2) or as a coherent flow object in its own right (LAVD, CSC, S^2 Robust Sets, FCM Clustering, transfer operator, dynamic Laplace operator). The vortices are another key feature of this system and were identified practically by every method except for the VLCS method, though the foliations of hyperbolic VLCSs identified (particularly those from VLCS–C) wrap around these vortices to form disjoint boundaries for these structures. Other LCSs identified include the vortex layers themselves (identified from the transfer operator, dynamic Laplace operator and FCM clustering), hyperbolic flow barriers which weave around the vortices (from the FTLE, VLCS and S^2), and the flow regions between the vortices (from FCM clustering, which defines these regions as consistent in the sense that fluid particles initially positioned here do not enter the vortices or cross the zonal jet).

3.4 Kelvin–Helmholtz Version 1 (KH1)

We have examined how each of the LCS detection methods we consider in this study detect coherent flow objects within analytically defined “toy models”, which have well defined, idealised coherent flow structures that are unlikely to exist within more realistic flows. Hence, we now turn our attention to more realistic flavours of flows to see if our LCS detection methods will be just as effective in detecting coherent flow objects here as they were when applied to the analytically defined flows discussed in the previous three sections. Our first two “realistic”–type flows will be simulated through solving the Navier–Stokes equations through spectral methods and techniques to obtain two versions of computational fluid dynamics (CFD) velocity data representing flow systems which contain Kelvin–Helmholtz type vortex layers, based on previous work in this endeavour (Lesieur et al. 1988, Metcalfe et al. 1987, Lee & Kim 2015, Schroeder et al. 2019). Our first version of this system (which we will refer to as Kelvin–Helmholtz (Version 1), or KH1) is obtained by solving the Navier–Stokes equations with the initial condition

$$\mathbf{u}_0\left(x,y\right)=\begin{bmatrix}\frac{1}{2}\left(\tanh\left[\frac{1}{\delta}\left(y-\frac{\pi}{2}\right)\right]-\tanh\left[\frac{1}{\delta}\left(y-\frac{3\pi}{2}\right)\right]\right)\\ \frac{1}{100}\sin\left(mx\right)\end{bmatrix}, \tag{3.4}$$

where $\delta = 0.05$ and $m = 5$. We generate the velocity data for this system by solving the Navier–Stokes equations using a fast Fourier transform algorithm which we have coded (and verified) ourselves in MATLAB. The equations are solved over a uniform initial grid of 1024×1024 equally spaced points within the 2π periodic domain $\Omega = [0, 2\pi) \times [0, 2\pi)$ using a third–order Runge–Kutta integration scheme in spectral wavenumber space, with $\Delta t = 0.001$ and Eulerian velocity data saved after one tenth of a unit of time for a total of ten units of time. We define this velocity data to have units, with length measured in metres, time measured in seconds and velocity measured in metres per second, with each respective unit scaled by a factor of 1. We also apply viscosity to these simulations using a Reynolds number of 10^3, computed from the lengthscale of 1 metre, maximal velocity of 1 ms^{-1} and kinematic viscosity $\nu = 10^{-3}\mathrm{m}^2/s$.

These simulations produce a flow system which consists of a wide transient flow channel centred along $y = \pi$ metres and approximately π metres in length, where particles flow consistently to the right of the domain Ω at a velocity no larger than 1 ms^{-1}. The vorticity of fluid particles along $y = \pi/2$ metres and $y = 3\pi/2$ metres results in the gradual construction of Kelvin–Helmholtz vortex layers, with consistent vortices forming in shape and becoming more well defined as the flow proceeds. In the remaining portions of the domain, where $y < \pi/2$ metres and $y > 3\pi/2$ metres, flow velocity is virtually nil and fluid particles remain more or less stagnant. For the LCS analysis to follow, we consider the initial time $t_0 = 5$ seconds, by which time sizeable Kelvin–Helmholtz vortices would have formed, and we consider the time interval which extends to $t_0 + T = 8$ seconds.

We begin once again with the FTLE, which was computed over the full grid of 1024×1024 uniformly spaced initial points used to simulate the velocity for this system and is on display in Figure 3.31. The maximal ridges of this field are sinusoidal curves which sit between the vortex layers and the central flow channel and act both as flow barriers between these components of the flow and also as support curves for the vortices themselves. In addition to these maximal ridges, we also observe more grainy ridges of high level FTLE between the vortices and the flow channel which reduce in FTLE value as we tend towards the channel, and medium level ridges which weave towards the centres of each of our vortices. This is to be expected as, unlike in a flow system such as the Double Gyre (see Figure 3.1) or Bickley Jet (see Figure 3.21), the vortices in this system are not present immediately at $t_0 = 0$. Instead, the vortex layers begin as straight lines and the lines change shape, forming vortices as the flow progresses. Lower values of the FTLE are observable within the central flow channel itself, the stagnant flow regions

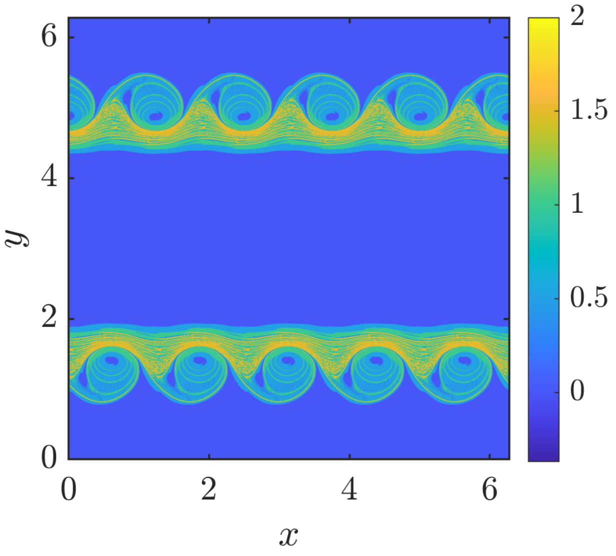

Figure 3.31: The FTLE Φ_5^8 for the KH1 flow.

present around the upper and lower boundaries of the domain, and the centres of the vortices themselves.

Next, we consider all three of our VLCS detection algorithms on the full set of 1024×1024 initial points, as was the case with the FTLE method. The results from each of these algorithms can be seen in Figure 3.32. For VLCS–A, we again randomly sample 500 initial points from the acceptable set $\mathcal{G}_0$ and choose the allowance parameter $\mathcal{L}_f = 1$ m. For VLCS–B and VLCS–C, we choose a local separation parameter $\rho = 25\pi/512 \approx 0.153$ m and a maximal length for the structures $\mathcal{L}_{\max} = 30$; and in VLCS–C we reduce the allowance parameter $\mathcal{L}_f$ to 0.5. The foliation of hyperbolic VLCSs produced from each of these algorithms consists primarily of wave shaped curves whose curved crests sit to the left of each vortex, providing a flow barrier between each vortex while also separating off the vortex layer and the stagnant outer flow regions from the central channel. VLCS–A not only produces these structures, but also extracts curves which spiral towards the centres of the coherent vortices in similar fashion to the medium value FTLE ridges from Figure 3.31. VLCS–B detects this type of hyperbolic structure for only one vortex (though it could do it for more vortices if a parameter such as ρ were to be changed), and VLCS–C does not detect these at all, as the FTLE values are only of medium range (and therefore, below the threshold). Both VLCS–A and VLCS–B have extracted curves which veer into the stagnant flow regions of the domain, while VLCS–B produces curves which begin to tend towards the centre of the flow channel where particle flow is supposed to be linear and robust. VLCS–C appears to produce the simplest and cleanest picture of hyperbolic VLCSs in this system, with only the aforementioned wave shaped structures which separate the vortex layers from the central flow channel detected.

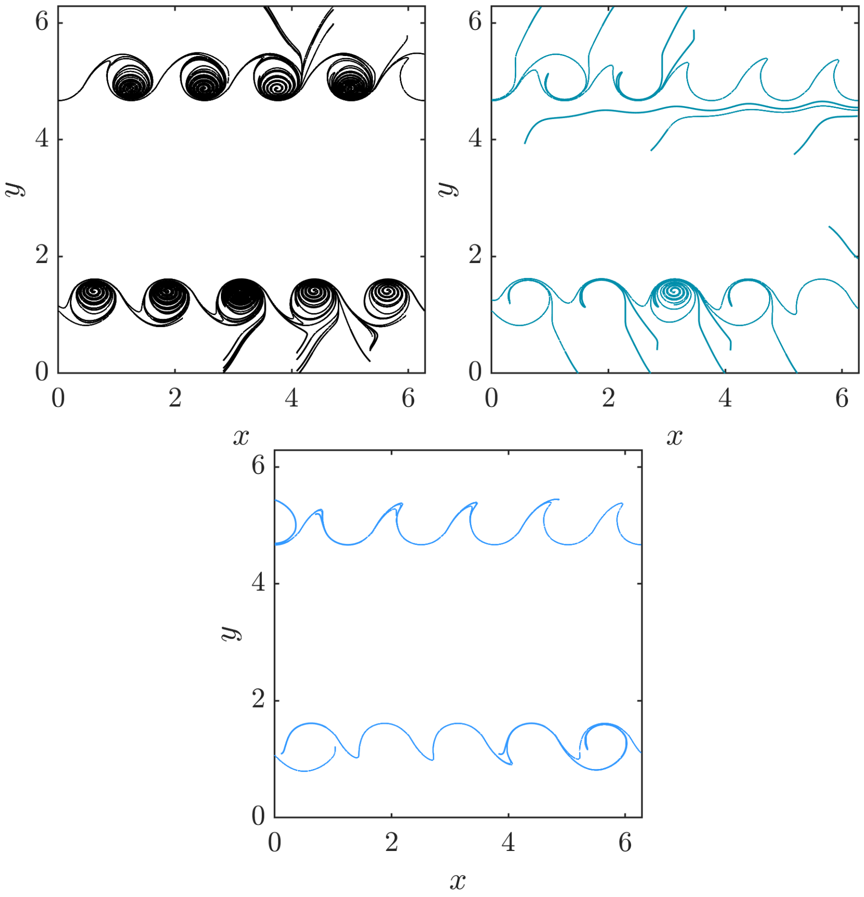

Figure 3.32: Hyperbolic VLCSs for the KH1 flow computed using the VLCS–A (top–left), VLCS–B (top–right) and VLCS–C (bottom) algorithms.

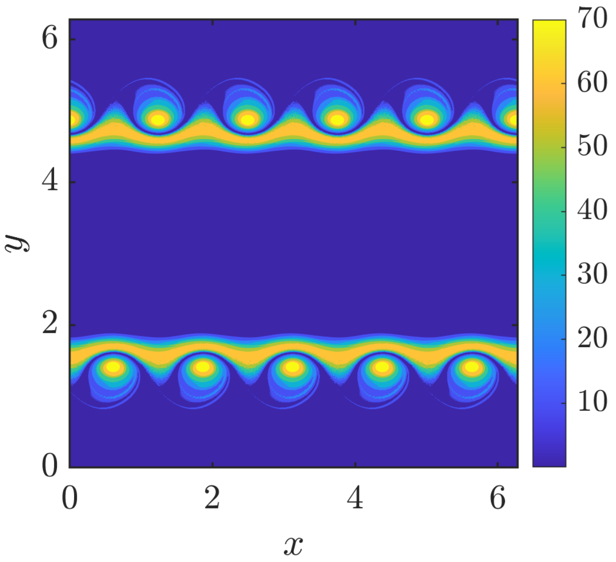

Figure 3.33: The LAVD Θ_5^8 for the KH1 flow.

The LAVD method will next be applied to this system, again by solving (2.7) for the full initial grid of 1024×1024 uniformly spaced points. The resultant scalar field produced can be seen in Figure 3.33. As expected, the largest LAVD values are observable within the coherent vortices of this system, allowing these structures to be easily identifiable from the scalar field. However, in a surprising result, we also observe large values of the LAVD within the flow barriers which sit under (and support) the vortices and separate these from the central flow channel. The likely reason we observe this is that Lagrangian shear is very strong within these portions of the domain, and even though these sinusoidal–shaped surfaces are not elliptical in shape, flow particles within these structures exhibit rotational coherence not unlike what we would expect to see within the bounds of a vortex. The LAVD is low throughout the remaining portions of the domain, including the central flow channel which is an extensionally (not rotationally) coherent structure for this system.

Next we consider the stochastic sensitivity method, computed on the same initial grid of 1024×1024 uniformly spaced points using the star grid method for the backwards time flow gradients once more due to memory constraints in MATLAB. Figure 3.34 shows the scalar field of the Scaled S^2, where $h = 2\pi/1024 \approx 6.1 \times 10^{-3}$ m and $v_r \approx 1.63 \times 10^{-2}$ ms^{-1}, corresponding to a precise value for ϵ of 10^{-2} $\mathrm{m} \cdot \mathrm{s}^{-1/2}$. A subsequent collection of robust sets are on display in the same Figure, which have been produced using the same values of h and v_r and with the lengthscales L chosen to be 0.45 m, 0.2 m and 0.08 m. Like with all of the other flow systems considered thus far, the Scaled S^2 field looks almost identical to the FTLE field (see Figure 3.31) in that it carries the same features and identifies very similar coherent structures. The biggest difference between these fields is that in the Scaled S^2 field we do not obtain a maximal ridge of the stochastic sensitivity which wraps

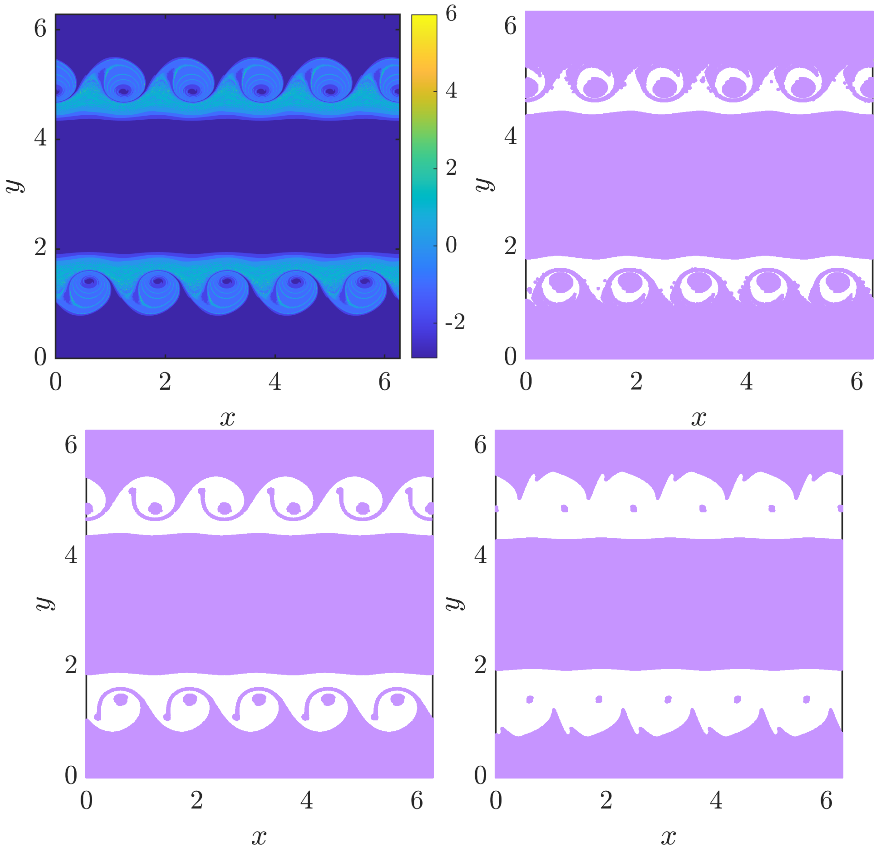

Figure 3.34: The Scaled S^2 for the KH1 flow, with $h = 2\pi/1024 \approx 6.1 \times 10^{-3}$ m and $v_r \approx 1.63 \times 10^{-2}$ ms^{-1} (top–left); along with three robust sets for this flow corresponding to lengthscales $L = 0.45$ m (top–right), 0.2 m (bottom–left) and 0.08 m (bottom–right).

around the vortices and instead larger values of this quantity are more or less distributed evenly throughout the shear layers.

In analysing the robust sets, one thing that stands out quite clearly from these results is that pretty much regardless of the value of L that we choose in this Figure, the robust set will practically always contain the central flow channel and the upper and lower stagnant flow regions. These structures are particularly consistent and robust within our Kelvin–Helmholtz flow system (with the stagnant regions robust in the sense that flow particles hardly move), and only undergo minor alterations when L is changed, such as a minor decrease in the width of the central flow channel. The biggest changes we observe within this set as L decreases are visible within and close to the vortex layers themselves. Beginning with $L = 0.45$ m, the centres of the vortices along with their boundaries are contained within the set, with the shear layer below the vortices excluded along with a small annular region surrounding the centres of each vortex. As L is reduced to 0.2 m, the shear layer along with some wave type shapes which curl into the vortices are excluded from the set, with only small centres of the vortices and hook type curves surrounding them remaining within the set. Reducing L even further to 0.08 m reduces the robust set to just the central flow channel, the stagnant flow regions and the centre–most points of the vortices.

We next consider the transfer operator method, which we carry out by dividing the domain Ω into 150×150 boxes of equal size and sampling 1000 uniformly distributed initial points within each box once more to compute the elements of the transfer operator matrix $\mathbf{P}$. In consideration of the final time t_0+T, we define periodic boundary conditions in both the x and y directions and divide the domain $\widetilde{\Omega}$ into 150×150 rectangular boxes of equal size once again. After producing the transfer operator matrix $\mathbf{P}$, the first ten singular values of this matrix we calculated to be 1, 0.9999, 0.9999, 0.9997, 0.9997, 0.9995, 0.9995, 0.9995, 0.9994 and 0.9994. Figure 3.35 shows the scalar fields of the second and third right singular vectors of $\mathbf{P}$ and two pairs of coherent sets obtained from these singular vectors. The second singular vector has divided the flow domain Ω almost precisely in half, with a zero contour line along $y = \pi$ acting as the boundary between each half of the domain Ω. Because we have defined periodic boundary conditions in y, some "spilling over" of portions of these coherent sets from the upper boundary to the lower boundary of Ω (or vice versa) can be observed. In contrast, when using the third right singular vector we detect the central flow channel as one coherent set while the domain above and below the channel (each containing a vortex layer) forms the second coherent set.

We obtain similar results when we apply the dynamic Laplace operator method to this flow system. We use the same parameters as those defined earlier for the transfer operator method, and the first ten eigenvalues produced for $\mathbf{\Delta}^D$ have come out to be 0, -0.2664, -0.5747, -1.1543, -1.7192, -1.8084, -2.1355, -2.1759, -2.6637 and -3.0581. In Figure 3.36, we divide Ω into a pair of coherent sets in two ways using the second and third eigenvectors produced for $\mathbf{\Delta}^D$. The resultant coherent sets obtained from the second eigenvector

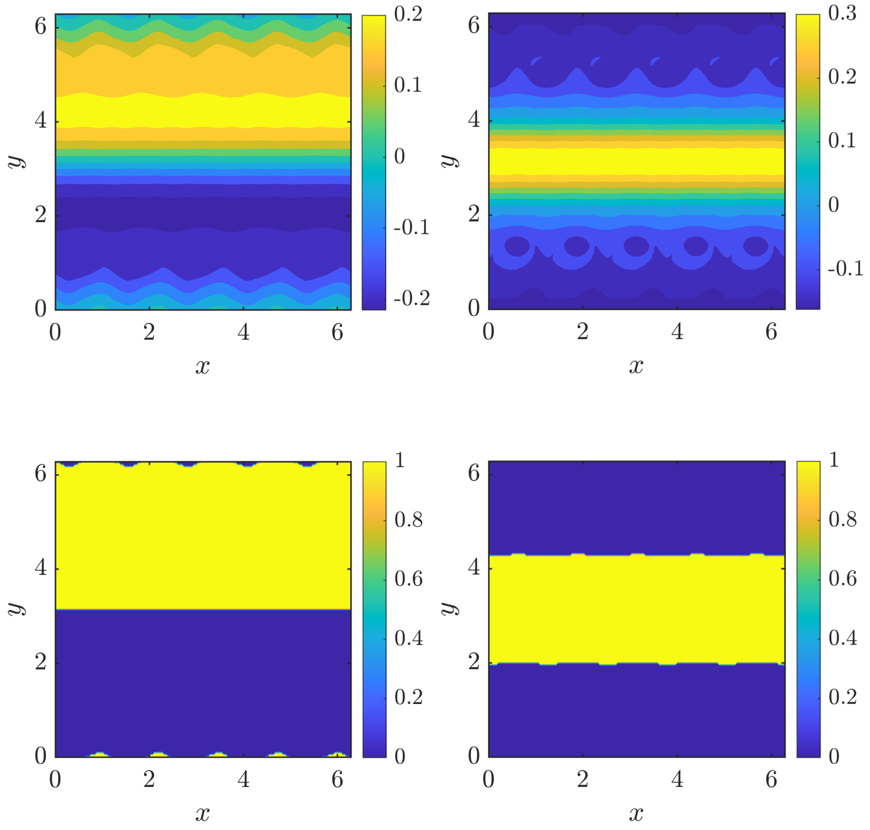

Figure 3.35: Scalar fields of the second (top–left) and third (top–right) right singular vectors of $\mathbf{P}$ for the KH1 flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

have come out to be almost exactly the same as those produced from the second right singular vector of $\mathbf{P}$, with the top half of the flow domain classified as one coherent set and the bottom half of the domain classified as the other. The same can be said for the third eigenvector of $\boldsymbol{\Delta}^D$, which isolates the central flow channel as one coherent set and the remainder of the flow domain which sits above and below this channel as the other coherent set, just as the third right singular vector of $\mathbf{P}$ was able to do.

Next, we apply the FCM clustering method to this flow in order to seek $K = 2$ clusters. We use the full uniform grid of 1024×1024 initial points and set the sharpness parameter $m = 1.5$. Figure 3.37 shows the membership probability scalar fields for these clusters along with two coherent clusters extracted at the 70% membership probability level. The first cluster contains the central flow channel and is bounded by the vortex layers on either side (as demonstrated by the sinusoidal boundaries of the cluster); while the other cluster contains the vortex layers and the surrounding stagnant flow reaching to the closest vertical boundary to each layer. We next increase K to 4, with the membership

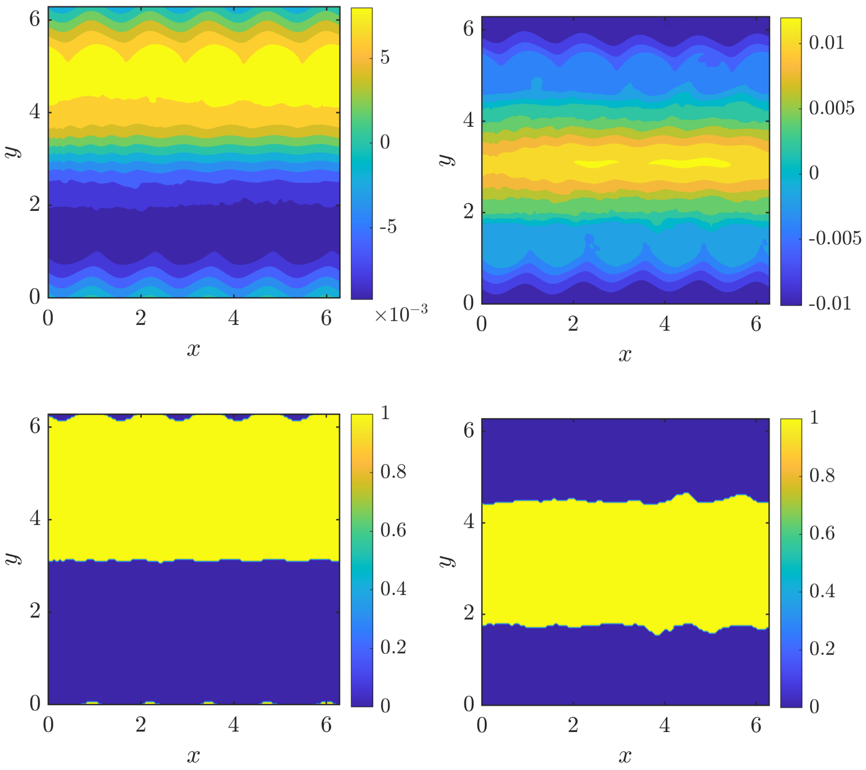

Figure 3.36: Scalar fields of the second (top–left) and third (top–right) eigenvectors of $\mathbf{\Delta}^D$ for the KH1 flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

probability scalar fields shown in Figure 3.38 and the clusters subsequently obtained at the 70% threshold shown in Figure 3.39. We again extract the central flow channel as one coherent cluster for this flow, while the upper and lower vortex layers and their neighbouring stagnant flow regions have been identified as separate coherent clusters. The fourth cluster is a disjoint combination of the upper and lower shear layers which separate the vortex layers from the central flow channel.

Finally, we execute the CSC method on a uniformly spaced grid of 80×80 initial points in Ω to generate the scalar field shown in Figure 3.40. The central flow channel stands out as the key coherent structure of the system from this field, as flow particles corresponding to the strongest positive values of the CSC coefficient cover this region of Ω. Strong negative values of this coefficient can be found in the stagnant flow regions close to the upper and lower edges of the domain, and also within the coherent vortices of the flow. Medium level values of this coefficient are used to identify the shear layers which surround the central flow channel along with wave shaped flow objects that have been produced as the vortex layer changes shape and the coherent vortices take shape.

The key structures which define the KH1 flow are the central flow channel and the

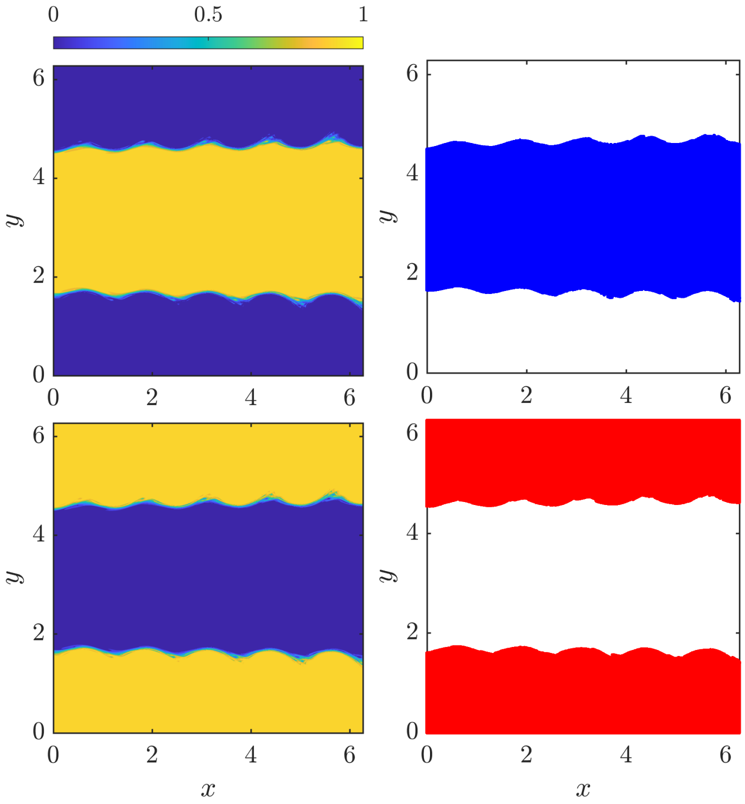

Figure 3.37: Membership probability fields for two FCM clusters of the KH1 flow (left Figures) and two clusters for this flow obtained at the 70% membership probability level (right Figures).

vortex layers which surround it from above and below. Most of the methods were able to locate the central flow channel, however each method had a different way of distinguishing this from the vortex layers. The S^2 robust sets, the transfer operator, dynamic Laplace operator, FCM Clustering and CSC methods were all able to explicitly detect this structure as a rectangular flow channel centred along $y = \pi$ m, though the width of this channel and the shapes of its upper and lower boundaries varied slightly between each of these methods. The FTLE did not explicitly detect this structure, but was able to demonstrate that this was a coherent flow region corresponding to low FTLE values and low particle stretching. The LAVD failed to detect the flow channel as it is extensionally and not rotationally coherent, while the VLCS method also failed to detect the flow channel but was able to find wave–shaped flow barriers which separate the channel from the upper and lower vortex layers. The vortices within this system are more realistic in the sense that they build up as the flow progresses, do not exist immediately as the flow commences and are not defined by a rigid vortex boundary like in the DG or Bickley Jet systems. The LAVD was able to explicitly identify these vortices, as were

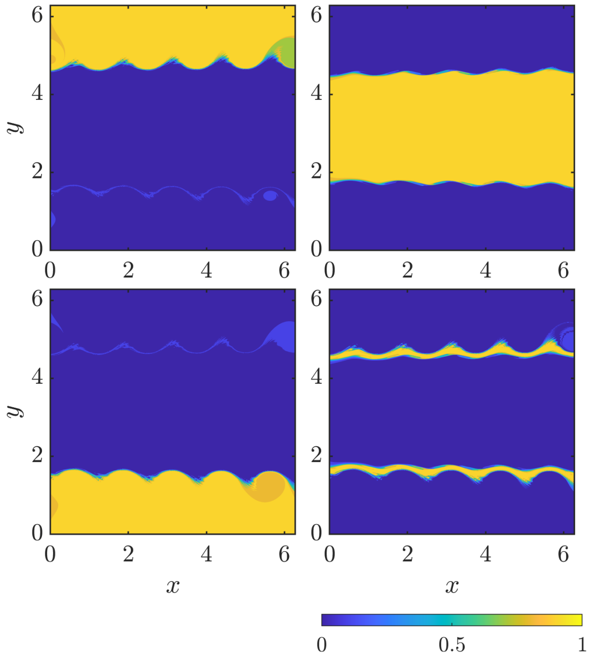

Figure 3.38: Membership probability fields for four FCM clusters of the KH1 flow.

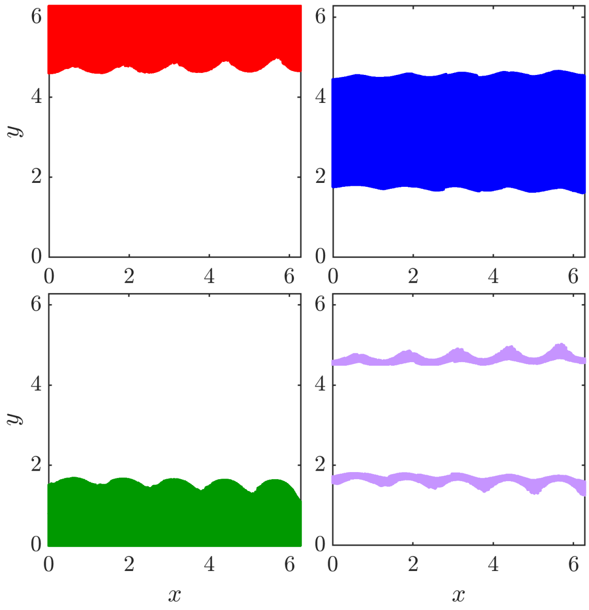

Figure 3.39: Four FCM clusters for the KH1 flow obtained at the 70% membership probability level.

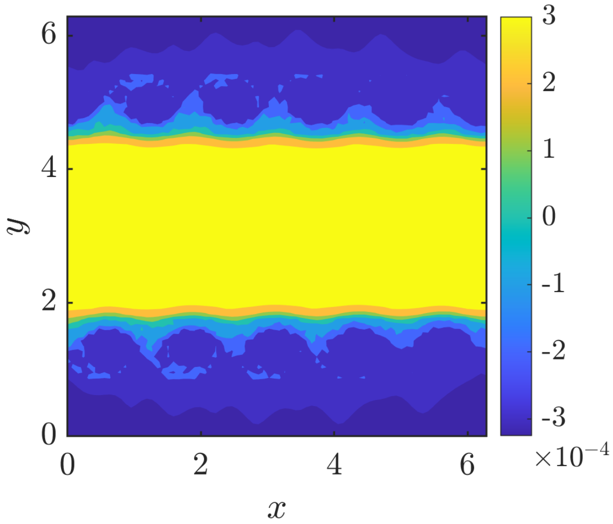

Figure 3.40: The CSC scalar field for the KH1 flow.

the S^2 robust sets for larger lengthscale thresholds L. The S^2 scalar field and the FTLE were also able to identify these structures, along with codimension–1 flow structures of greater flow stretching or uncertainty which folded towards the centres of the vortices. The transfer operator, dynamic Laplace operator, FCM clustering and CSC methods all failed to detect the vortex centres, but were able to identify the full vortex layers as coherent structures to some extent, along with surrounding flow which was advected as the vortex layer took shape. Another interesting feature is the sinusoidal or wave type shear layers which separate the Kelvin–Helmholtz vortices from the central flow channel on either side of the channel. The FTLE, S^2, FCM clustering and CSC methods were all able to identify these features of the flow to some extent, as was the LAVD method where this quantity had a large value within these layers due to the strong Lagrangian shear observable within these structures. The VLCS method found wave shaped flow barriers which form the boundaries of these layers and curl around the vortices, while the transfer operator and dynamic Laplace operator methods both failed to find these layers even as boundaries between the coherent sets identified. The results obtained from our first numerically defined flow system illustrate just how different the LCSs obtained from each detection method can be, particularly as we are no longer working with an analytically defined toy model which features strongly idealised flow structures.

3.5 Kelvin–Helmholtz Version 2 (KH2)

For our next numerically defined system, we consider another simulated flow with Kelvin–Helmholtz vortex layers which move through the flow in a much more unpredictable fashion than in the KH1 flow considered in the previous Section. While the KH1 flow contained two vortex layers within which separate and distinct vortices formed gradually as the flow progressed, in this flow (which we will denote as Kelvin–Helmholtz (Version 2) or KH2) Kelvin–Helmholtz vortices take form more quickly and collide and combine with other vortices as the flow progresses. We simulate the KH2 system by solving the Navier–Stokes equations using the same computational methods and parameters defined in the previous section for the KH1 flow, except in this case we simulate 20 seconds worth of velocity data and we change the initial condition for the velocity to

$$\mathbf{u}_0\left(x,y\right)=\begin{bmatrix}\frac{1}{2}\left(\tanh\left[\frac{1}{\delta}\left(y-\frac{\pi}{2}\right)\right]-\tanh\left[\frac{1}{\delta}\left(y-\frac{3\pi}{2}\right)\right]\right)\\ \frac{1}{100}\left(\sin\left[\alpha x\right]+\frac{1}{2}\sin\left[\left(\alpha-1\right)x\right]+\frac{1}{2}\sin\left[\left(\alpha+1\right)x\right]\right)\end{bmatrix}, \tag{3.5}$$

where $\delta = 0.05$ once more and $\alpha = 9$, the nearest whole number to $0.22/\theta_0 = 0.44/\delta \approx 8.8$ which is an estimate for the most unstable wavenumber computed from θ_0, the initial thickness of the vortex layers which exist within the system (Mattner 2011). In the KH2 flow, we begin at the initial time of $t = 0$ seconds with a flow channel centred along $y = \pi$ m of the same length as that which exists within the KH1 system, two vortex layers centred along $y = \pi/2$ m and $y = 3\pi/2$ m and stagnant flow close to the upper and lower boundaries of the domain $\Omega = [0, 2\pi) \times [0, 2\pi)$. A line of vortices begin to take form along each layer just like in the KH1 flow, however these vortices are smaller in size and are packed closer together along each vortex layer. By the time we have reached $t = 5$ seconds worth of flow, the vortices on each layer begin to move to the right of the domain and collide and merge with each other to form larger vortices and wider vortex layers; causing the width of the central flow channel to decrease. By the time we reach $t = 15$ seconds worth of flow, four large travelling vortices (two on each layer) have been generated, and the width of the central flow channel is now half of its original width. By changing the initial condition to produce vortices that move and collide with each other rather than just gradually taking shape in one spot, the vortex layers in the KH2 system now exhibit much more unpredictable flow behaviour in comparison to the layers in the KH1 system. We implement each of our LCS detection methods on the KH2 system over the flow time interval ranging from $t_0 = 12$ to $t_0 + T = 15$ seconds. All other computational parameters pertinent to each LCS detection method remain the same as those selected for the KH1 flow unless stated otherwise.

We begin with the FTLE method, with the scalar field generated for the KH2 flow on display in Figure 3.41. The maximal ridges of this field are tucked between the vortices and the flow channel in similar fashion to what was observed for the KH1 system. However, this time around, the maximal ridges do not form the boundaries of the flow channel completely due to the enlargement of the vortex layers caused by the merging of these vortices. The shear layers between the vortices and the flow channel have extended in size and changed in shape; with the positions and trace of these layers indicated by ridges of medium level FTLE. In similar fashion to the KH1 flow, medium level ridges of the FTLE are also observable within the vortices themselves, on this occasion traversing towards the

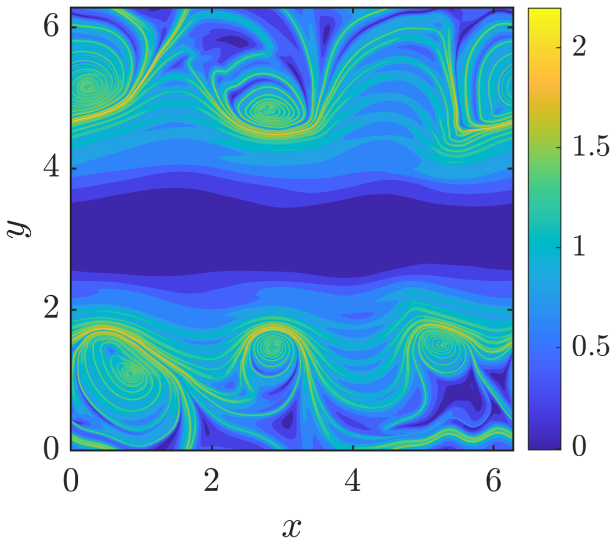

Figure 3.41: The FTLE Φ_{12}^{15} for the KH2 flow.

centre of the vortices. Minimal values of the FTLE are found throughout the centre of the flow channel, which has considerably shrunk in length in comparison to the flow channel present in the KH1 system.

Next, we consider the three VLCS detection algorithms, with the results obtained from each of these algorithms featured in Figure 3.42. Like with the KH1 system, each of the VLCS detection algorithms are able to obtain LCSs that coincide with maximal and medium level ridges of the FTLE scalar field. The structures produced by VLCS–A separate the central flow channel and the surrounding shear layers, and spiral in towards the centres of most of the vortices present after 12 seconds of flow. In a fashion consistent with the other flow systems considered thus far, this algorithm has produced repeated LCSs which lie in similar regions of the domain and are separated from each other by only a small distance. The use of a separation radius ρ prevents this occurring within the results produced by the other two algorithms. The results of VLCS–B separate the vortex layers from the greater flow region containing the central flow channel, feature some (but not all) of the variational structures spiralling towards the centres of the vortices and, as was seen with the KH1 system, include structures of “locally maximal” particle repulsion capability located within the flow channel itself. VLCS–C produces a picture of results that may be viewed as a sensible centre between the results of the previous two algorithms, with the production of structures within the vortex layers excluding those which sit too close and are shaped too similarly to other structures that have already been generated; some structures which spiral in towards the centres of the vortices and no structures present within the central flow channel.

Next, the LAVD scalar field for the KH2 flow is shown in Figure 3.43. This method

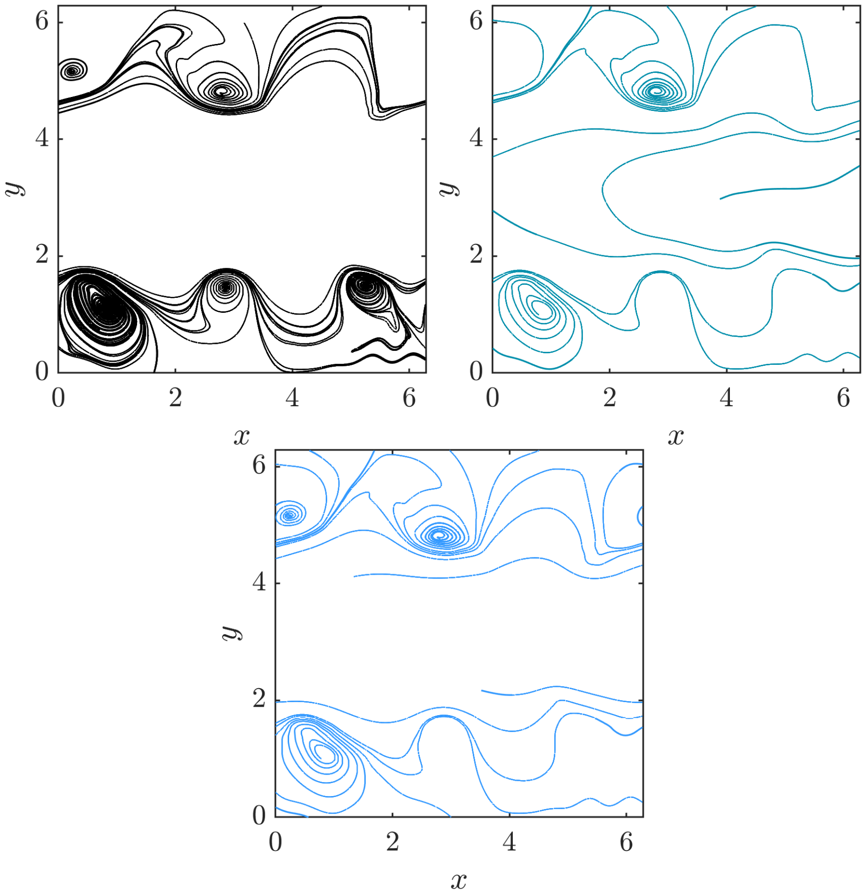

Figure 3.42: Hyperbolic VLCSs for the KH2 flow computed using the VLCS–A (top–left), VLCS–B (top–right) and VLCS–C (bottom) algorithms.

illustrates quite clearly that after 12 seconds of particle flow, collections of smaller vortices have merged to form what appear to be two large vortices in the upper vortex layer (the right most vortex combines with the left most vortex under periodic boundary conditions), and three slightly smaller vortices in the lower vortex layer. In comparison to the LAVD results for the KH1 flow, we were not able to identify any shear layers which support these vortex layers from the LAVD results for this flow. This is most likely due to the expansion and rapid change in the shape of the vortex layers resulting in the boundaries between the vortex layers and the flow channel being defined by strong fluid particle stretching rather than strong Lagrangian shear in this system.

We next consider the S^2 method, with the Scaled S^2 scalar field shown in Figure 3.44, along with three robust sets for this system corresponding to lengthscale thresholds L of 0.4 m, 0.25 m and 0.1 m. The Scaled S^2 field once again appears quite qualitatively similar to the FTLE field, though again there are some minor differences, such as medium

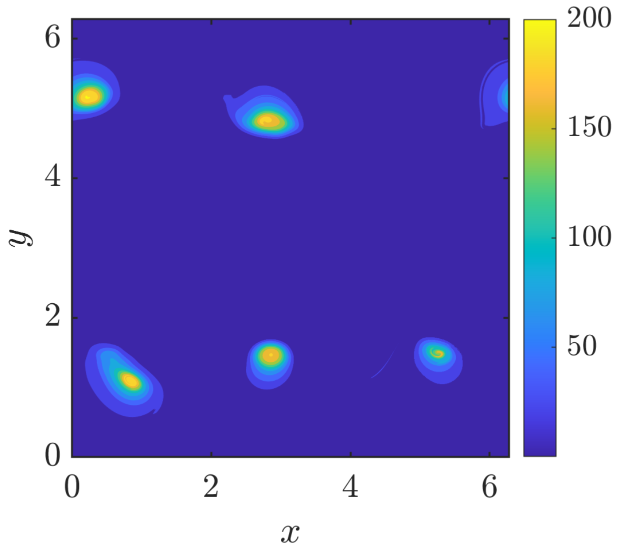

Figure 3.43: The LAVD Θ_{12}^{15} for the KH2 flow.

level ridges of the Scaled S^2 separating the vortices from the central flow channel instead of maximal ridges like in the case of the FTLE field; along with "ripple" type structures observable within the Scaled S^2 field around the points $(x, y) = (2.5, 5.7)$ and $(4.75, 0.5)$; instead of troughs as observed in the FTLE field. Each robust set contains the central flow channel of the KH2 system, though the width of this structure progressively shrinks as the threshold L decreases. For a large enough value of L, the robust set captures the centres of the coherent vortices above and below the channel along with a range of coherent clusters of miscellaneous shape which surround these vortices within the layers.

Next, we focus on applying the transfer operator method to the KH2 system. Upon producing $\mathbf{P}$ for this flow, the first ten singular values produced came out to be 1, 0.9998, 0.9997, 0.9997, 0.9996, 0.9992, 0.9992, 0.9991, 0.9990 and 0.9990. Figure 3.45 shows the scalar fields for the second and third right singular vectors of $\mathbf{P}$ and pairs of coherent sets for this flow generated from each of these vectors. From the second right singular vector, the coherent flow channel is classified as one coherent set and the remainder of Ω is classified as the other coherent set. These results are similar to the coherent sets obtained for the KH1 flow, however those coherent sets were obtained from the third singular vector rather than the second. In contrast, the third right singular vector produces the top half of the domain as one coherent set and the bottom half of the domain as the other. We obtained similar results for the KH1 flow from the second singular vector of $\mathbf{P}$ instead of the third, however in the KH2 flow we see more "spilling" of the coherent sets over the upper and lower edges of the doubly periodic domain than we did in the KH1 flow.

We obtain similar results upon the implementation of the dynamic Laplace operator method on this flow. The first ten eigenvalues generated for $\mathbf{\Delta}^D$ were 0, -0.2034, -0.3477,

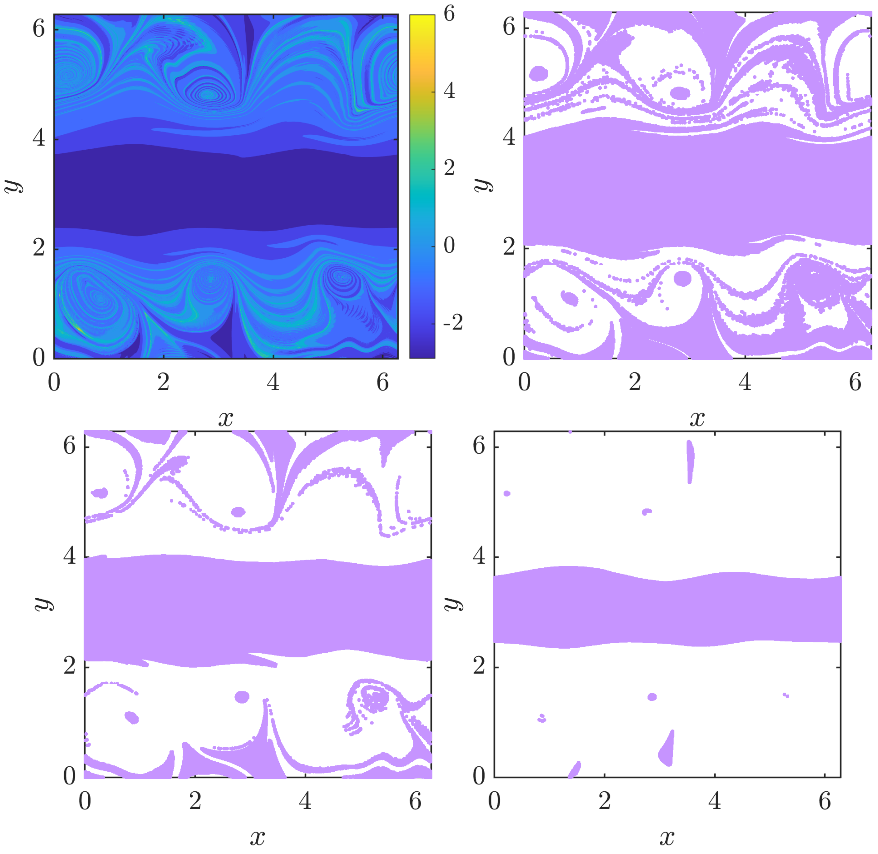

Figure 3.44: The Scaled S^2 for the KH2 flow, with $h = 2\pi/1024 \approx 6.1 \times 10^{-3}$ m and $v_r \approx 1.63 \times 10^{-2}$ ms^{-1} (top–left); along with three robust sets for this flow corresponding to lengthscales $L = 0.4$ m (top–right), 0.25 m (bottom–left) and 0.1 m (bottom–right).

-0.6764, -0.9588, -1.3553, -2.2505, -2.3670, -3.0580 and -3.2800. Figure 3.46 shows the scalar fields of the second and third eigenvectors of $\mathbf{\Delta}^D$ and two coherent sets produced from each of these fields. From the second eigenvector, we have produced an almost identical pair of coherent sets to those obtained from the second singular vector of $\mathbf{P}$, with the central flow channel identified as one coherent set and the remainder of the domain (including the two vortex layers) identified as the other set. Similarly, the results generated from the third eigenvector of $\mathbf{\Delta}^D$ are practically identical to those obtained from the third singular vector of $\mathbf{P}$, bar some minor differences and inconsistencies.

Next, we apply the FCM clustering method with $K = 2$ to this flow to obtain the membership probability scalar fields and coherent clusters shown in Figure 3.47. As we observed with the KH1 flow, the two clusters identified in this case are the central flow channel and a disjoint combination of the two vortex layers above and below this channel; with the biggest observable difference being the sinusoidal boundaries of these clusters.

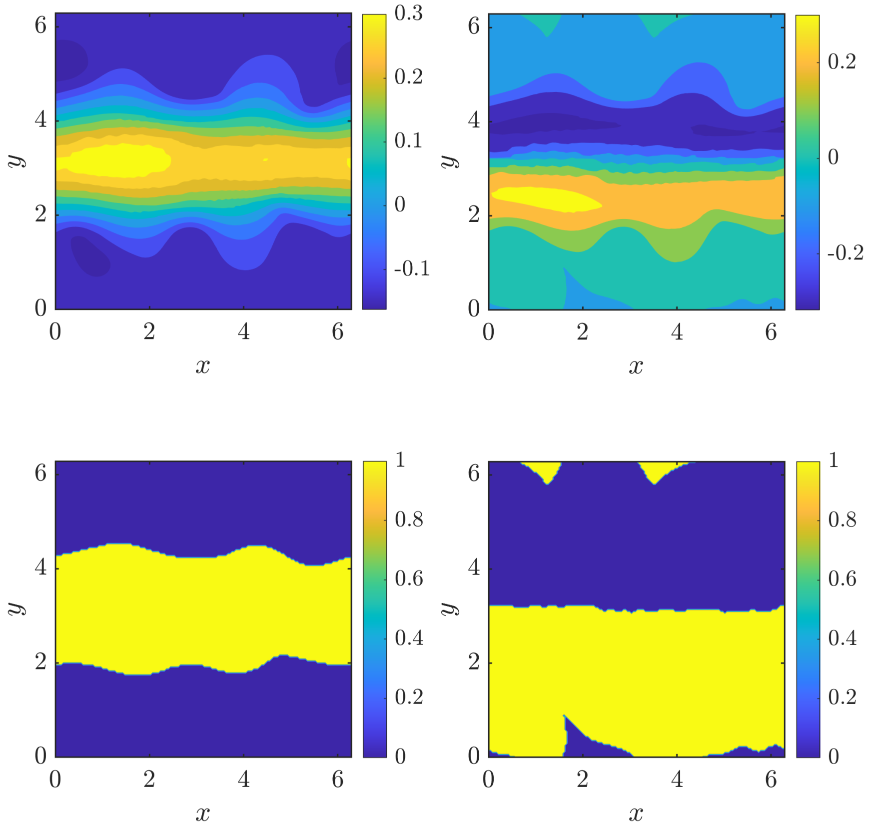

Figure 3.45: Scalar fields of the second (top–left) and third (top–right) right singular vectors of **P** for the KH2 flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

Upon increasing K to 4, we obtain the membership probability fields shown in Figure 3.48 and the four coherent clusters shown in Figure 3.49. We have obtained a similar collection of clusters to those generated from the $K = 4$ case in the KH1 flow, with the central flow channel, the upper and lower vortex layers; and the disjoint shear layers which separate the central flow channel from the vortex layers identified for this flow.

Finally, we apply the CSC method to the KH2 flow and produce the scalar field shown in Figure 3.50. From this result, strongly positive CSC coefficients correspond to the central flow channel in similar fashion to the KH1 flow, while strongly negative CSC coefficients are used to identify the two vortex layers. Unlike in the KH1 flow, the CSC method has been able to identify at least the outlines of the coherent vortices from strongly negative CSC coefficients, along with portions of the domain which were dragged along with the vortices as they merged and the vortex layers expanded. The shear layers between the flow channel and the vortices correspond to CSC coefficients of low magnitude or a value of zero; indicating a lack of coherent flow behaviour within these regions of the

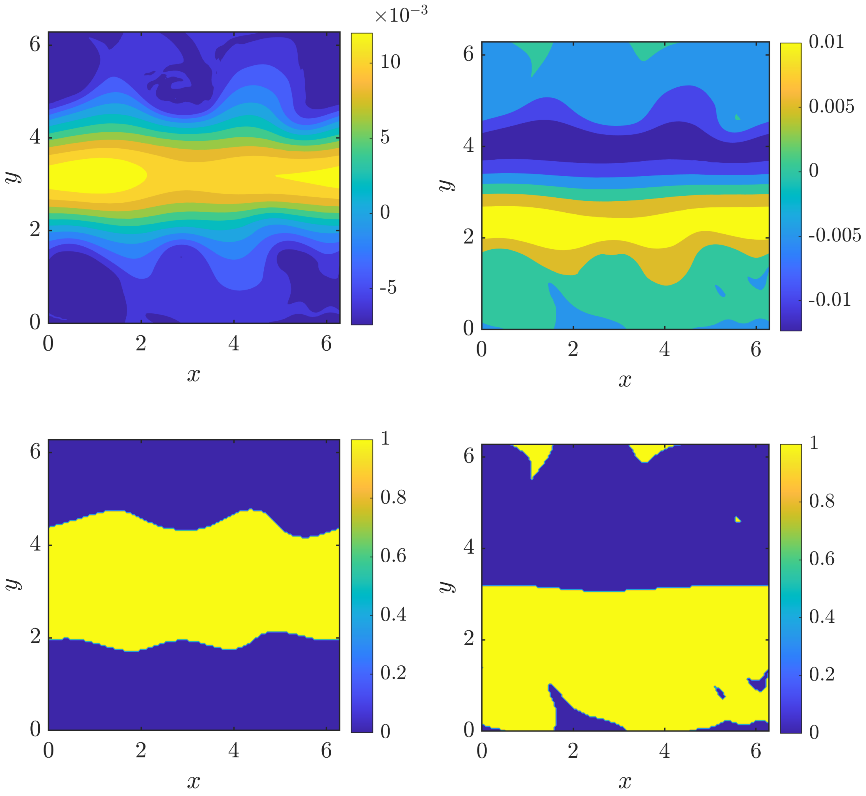

Figure 3.46: Scalar fields of the second (top–left) and third (top–right) eigenvectors of $\boldsymbol{\Delta}^D$ for the KH2 flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

domain.

In comparison to all of the previously discussed dynamic velocity systems, the KH2 flow has been the most volatile and unpredictable dynamic system we have encountered thus far. In light of this, each of our LCS detection methods has detected a range of coherent flow objects with varying shapes, sizes and formations. Like with the KH1 flow, the definitive structure of this flow is the central flow channel, which in this flow is shorter in width due to the more volatile vortex layers. The shape of the central flow channel varies with each of these methods, with the S^2 robust sets, transfer operator, dynamic Laplace operator, FCM clustering ($K = 4$ case) and CSC methods identifying an oblong shaped flow channel which approximately spreads out between $y = 2$ m and $y = 4$ m. Meanwhile the FTLE, Scaled S^2 and FCM clustering (with $K = 2$) methods all identify the flow channel combined with the shear layer between the vortices and the channel as one structure; and the VLCS algorithms detect sinusoidal flow barriers which give this channel a similar shape. Like with the KH1 flow, the LAVD fails to detect the central channel as this structure is extensionally coherent and not rotationally coherent.

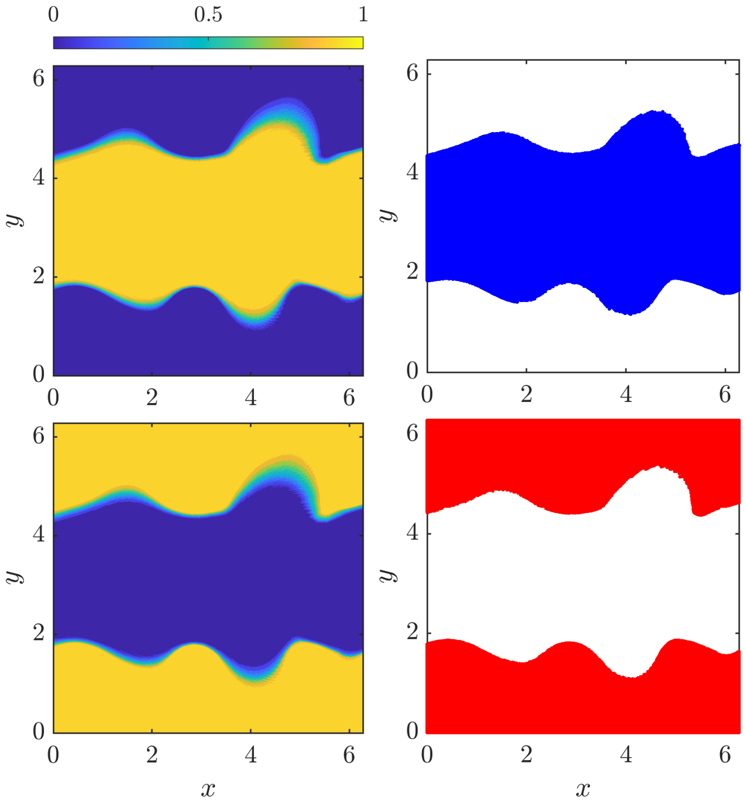

Figure 3.47: Membership probability fields for two FCM clusters of the KH2 flow (left Figures) and two clusters for this flow obtained at the 70% membership probability level (right Figures).

The LAVD method is able to clearly identify the vortices present within this flow, while the S^2 robust sets identify their centres for large enough lengthscale thresholds L and the CSC method identifies the approximate outlines of these vortices. Interestingly, the FTLE, Scaled S^2 and VLCS methods detect curves of large particle stretching which spiral in towards the centres of these vortices; which is likely due to the fact that these vortices are the result of the gradual merging of smaller vortices, and strong stretching of flow trajectories is observable as this occurs. The transfer operator, dynamic Laplace operator and FCM clustering methods do not identify any of these vortices but instead identify the overall vortex layers as coherent clusters or sets, which was similarly observed in the results for the KH1 flow.

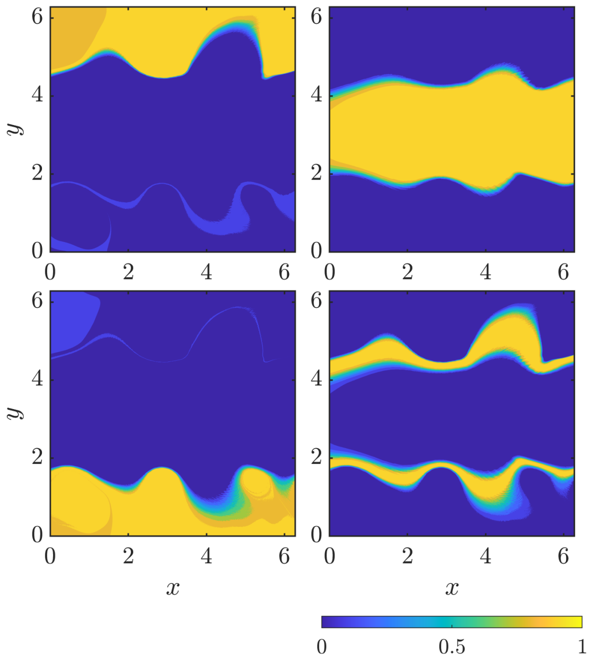

Figure 3.48: Membership probability fields for four FCM clusters of the KH2 flow.

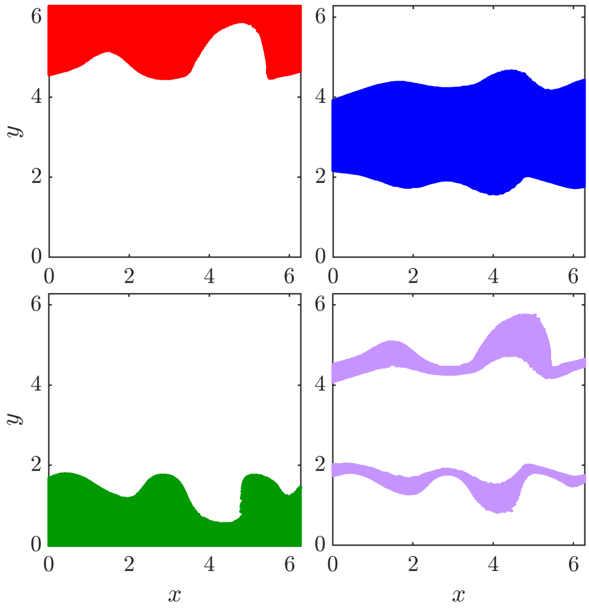

Figure 3.49: Four FCM clusters for the KH2 flow obtained at the 70% membership probability level.

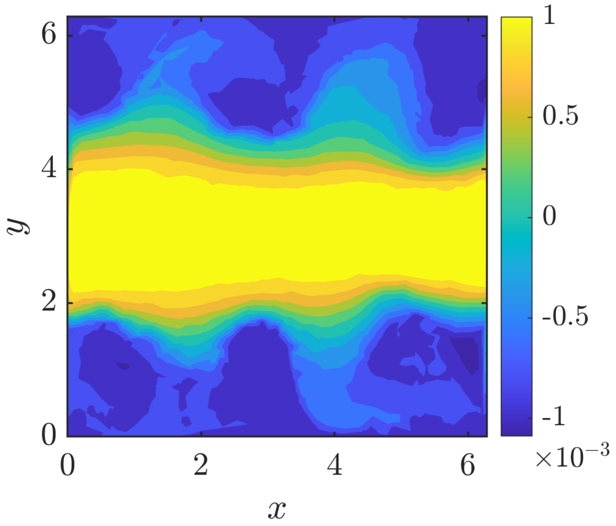

Figure 3.50: The CSC scalar field for the KH2 flow.

3.6 Gulf Stream

The final flow system to be considered in this study is an oceanographic data set focusing on the Gulf Stream, a warm current in the North Atlantic Ocean. The Gulf Stream is a rapidly moving jet stream which changes shape frequently and produces objects such as coherent cyclonic eddies as it evolves over time (Liu et al. 2018, Kang & Curchitser 2013). While this current is considered highly influential regarding climate patterns in the Northern and Western hemispheres (Liu et al. 2018), it has been seldom considered for LCS analysis up until the last few years (Liu et al. 2018, Froyland et al. 2019, Balasuriya 2020*b*). To examine Lagrangian coherent structures within this system, we use absolute geostrophic velocity data compiled by SSALTO/DUACS and supplied by the AVISO Satellite Altimetry Data service (https://www.aviso.altimetry.fr). The relevant velocity data has been defined over the spatial domain which ranges from 82 to 25 degrees West and 30 to 50 degrees North; with Eulerian snapshots of this data recorded daily at midnight (UTC) from January 1st, 2015 to March 31st, 2020 (inclusive). The velocity has been provided in units of metres per second at equidistant points spaced apart by 0.25 degrees latitude/longitude in each direction. For our study, we consider a spatial domain and time parameterisation similar to that used by Froyland et al. (2019) to study the dynamic Laplace operator method. The domain in question $\Omega = [65^\circ \text{ W}, 35^\circ \text{ W}] \times [30^\circ \text{ N}, 50^\circ \text{ N}]$, and the interval of flow considered ranges from t_0 = midnight, January 15th, 2015 to $t_0 + T$ = midnight, April 15th, 2015 (UTC), i.e. covering a period of $T = 90$ days worth of flow. As we will see from the results displayed later on in this Section, it is rather difficult to make out important coherent structures within this system over such a large domain and interval of flow time, however we proceed with these choices for the sake of

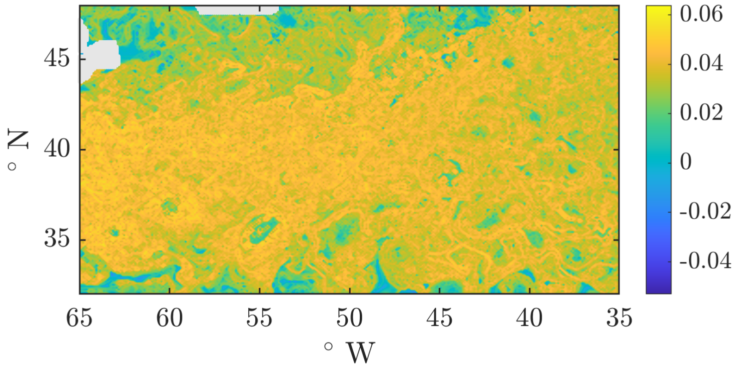

Figure 3.51: The FTLE $\Phi_{t_0}^{t_0+T}$ for the Gulf Stream flow, with t_0 = midnight, January 15th, 2015 (UTC) and $T = 90$ days.

maintaining consistency with Froyland et al. (2019). Froyland et al. (2019) use 291 × 224 uniformly spaced points in Ω to evaluate the dynamic Laplace operator using the sparse method defined in Froyland & Junge (2018). We will endeavour to use the same set up of points in each of our methods where possible, however this will not be possible for all of these methods due to system memory limitations and considerations regarding computation time. To solve the velocity system and produce the necessary flow maps for each of our initial points $\mathbf{x}_0$, we convert the velocity to units of degrees latitude/longitude per day to match the units of spatial distance and the temporal resolution of the data, and assign velocities of 0 to particles initially positioned on land (where the velocity is otherwise defined as NaN, or "Not a Number"). To prevent flow particles from leaving the domain, the boundary of the data set is layered with values of NaN for the velocity, so that flow trajectories "wash up" on the boundaries of the domain rather than leaving the domain altogether. This ensures that a value of each relevant LCS quantity can be computed at all points within our domain, but it can result in the production of some anomalous or unreliable flow patterns as illustrated by the following results.

We begin with the FTLE, with the scalar field produced for this quantity displayed in Figure 3.51. Unlike in the previous flow systems where curves representing maximal ridges of the FTLE quantity were easily identifiable, in this flow system large values of the FTLE are spread out throughout most of our domain. These large FTLE values attempt to represent the overall spread of the Gulf Stream over the domain after the 90–day period of flow, over which time the shape of the Gulf Stream would have changed considerably. We do have some minor indication of the shape of the Gulf Stream from this field, though it is difficult to extract this explicitly amidst a sea of large FTLE values. Smaller values of the FTLE help identify more coherent flow structures within this system, such as coherent eddies and vortices produced by the Gulf Stream. Closer to land, along the upper and lower boundaries of the flow domain, and within the Gulf of St Laurence (in the top–left corner of the domain), FTLE values are small or negative, indicating minimal fluid

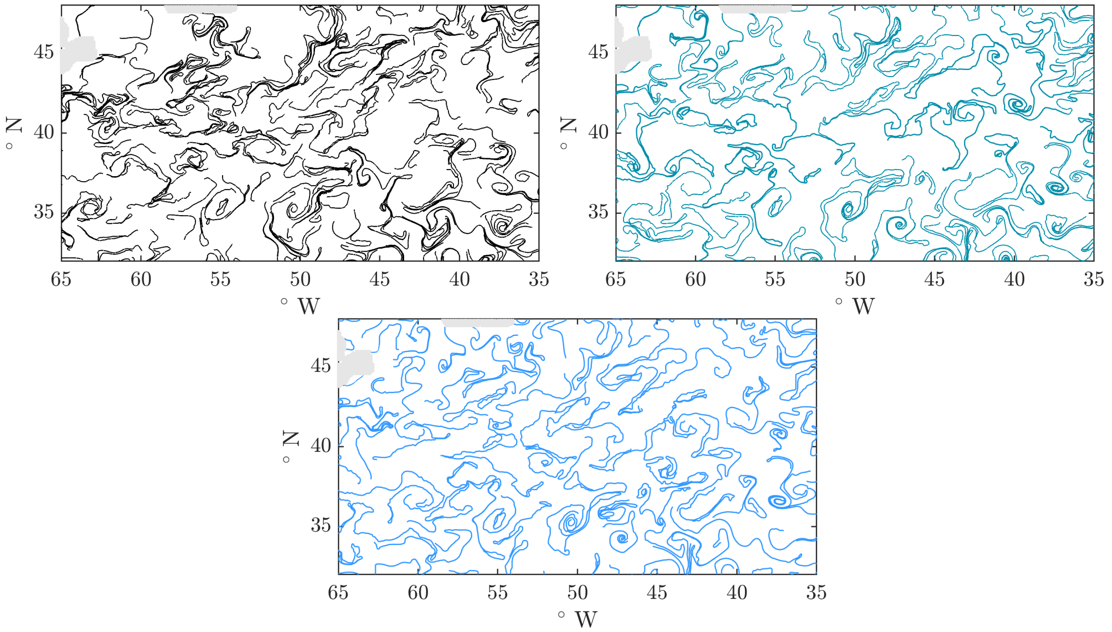

Figure 3.52: Hyperbolic VLCSs for the Gulf Stream flow computed using the VLCS–A (top–left), VLCS–B (top–right) and VLCS–C (bottom) algorithms.

particle stretching or flow folding or compression.

Next, we consider the three hyperbolic VLCS detection algorithms. In VLCS–A, we randomly sample 2000 candidate points for eigenvector field integration from the acceptable set $\mathcal{G}_0$, set the allowance parameter $\mathcal{L}_f = 2$ degrees latitude/longitude, and discard all structures produced which are less than 2 degrees in arc length. For both VLCS–B and VLCS–C, we set the maximal length of the structures $\mathcal{L}_{\max} = 30$ degrees and set the separation radius $\rho = 1$ degree. In VLCS–C, the allowance parameter $\mathcal{L}_f$ is reduced to 1 degree. The resultant foliations of hyperbolic VLCSs produced from each of these algorithms can be viewed in Figure 3.52. These results do not give a very clear picture of the most influential flow structures within this system. Rather than illustrating the boundaries of the Gulf Stream (for instance), the structures produced from this method cover the domain in an almost random fashion, and only occasionally indicate the boundaries of key structures such as the coherent eddies. In all of the other flow systems examined in this Thesis, the VLCS method was able to produce at least some meaningful flow structures, though in the vast majority of these cases these structures were generated within analytically defined toy flows and such idealised structures are highly unlikely to be observed in a more realistic system.

We next consider the LAVD for this system, with the scalar field for this quantity shown in Figure 3.53. From this field, we are able to gain a clear and cohesive picture of the eddies produced from the Gulf Stream as the most rotationally coherent flow

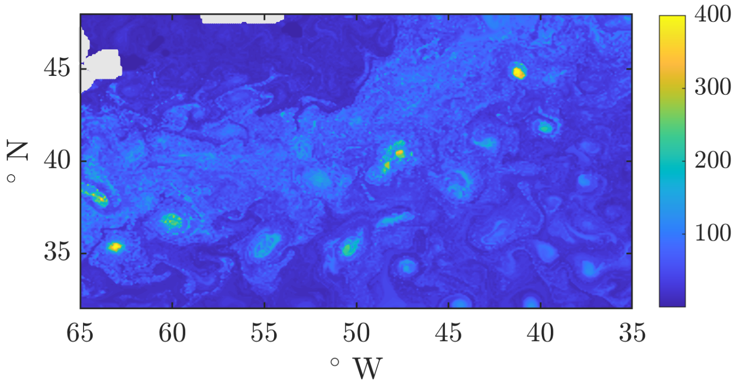

Figure 3.53: The LAVD $\Theta_{t_0}^{t_0+T}$ for the Gulf Stream flow, with t_0 = midnight, January 15th, 2015 (UTC) and $T = 90$ days.

objects within this system. These eddies are clearly identifiable as elliptic objects of medium to large LAVD values, with the LAVD varying in magnitude depending on how strongly coherent these structures are over the whole flow interval. It is likely that the eddies corresponding to larger LAVD values remain as consistent structures over the whole duration of the flow, while those corresponding to slightly lower values may not necessarily last over the complete course of this flow. We can also identify the Gulf Stream to some degree from these results, due to the medium levels of the LAVD observable within the Stream generated by the significant amounts of Lagrangian shear present within this structure. Low values of the LAVD are observable closer to land and within the Gulf of St Laurence, implying that flow behaviour within these portions of the domain is not necessarily incoherent, but rather that this flow is not rotationally coherent in the fashion of a vortex or an eddy.

We now turn our attention to the S^2 method, with the Scaled S^2 field shown in Figure 3.54 along with three robust sets for this flow corresponding to lengthscale thresholds L of 18, 6 and 2 degrees latitude/longitude. For this flow system, the Scaled S^2 field is considerably more grainy and less coherent in comparison to those produced for the flow systems considered earlier on in this Chapter. This is likely due either to the resolution of the velocity data, or the large length of flow T taken for this study. That said, we are still able to identify what appears to be the Gulf Stream from this field, identifiable from the largest values of the Scaled S^2. Structures corresponding to lower S^2 values are difficult to spot due to the graininess of the field, however we can still identify the Gulf Stream eddies or the Gulf of St Laurence in the top left corner of the domain from these values. Due to the aforementioned graininess of the S^2 field, extracting clear coherent structures from the robust sets generated from this system has proven to be difficult. When $L = 18$ degrees, the robust set contains a number of coherent eddies within this system, small parcels of flow surrounding the Gulf Stream and most of the Gulf of St Laurence. Most

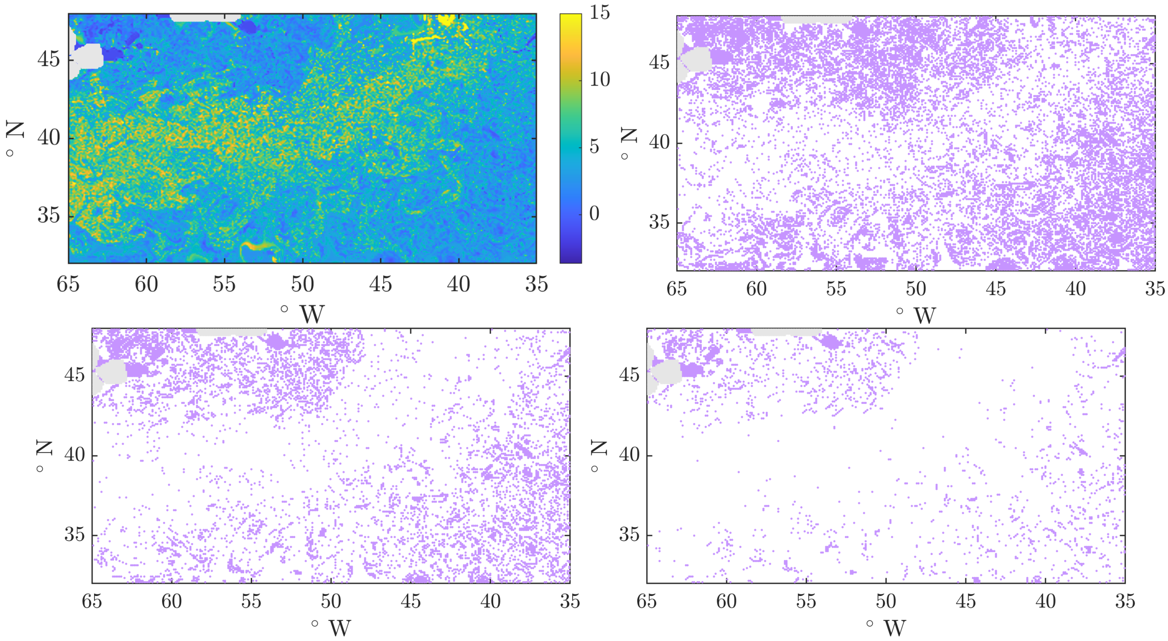

Figure 3.54: The Scaled S^2 for the Gulf Stream flow, with $h = 0.125$ degrees latitude/longitude and $v_r \approx 1.27 \times 10^{-2}$ degrees per day (top–left); along with three robust sets for this flow corresponding to lengthscales $L = 18$ degrees (top–right), 6 degrees (bottom–left) and 2 degrees (bottom–right).

of these structures begin to fade and dissolve as the threshold L is reduced; and by the time we reach a threshold of 2 degrees, all that remains are small parcels of flow within the Gulf of St Laurence and scattered collections of points below the Gulf Stream which do not appear to form any coherent, robust structures.

Next, we apply the transfer operator method to the Gulf Stream flow by dividing the domain Ω into 120×90 rectangular boxes of equal size, and once again sampling 1000 uniformly distributed initial points within each of these boxes; skipping over any boxes where over half of the box contains land. We expand the domain $\widetilde{\Omega}$ at time $t_0 + T$ to cover the full domain over which velocity data is available, i.e. $\widetilde{\Omega} = [82^\circ\ \mathrm{W}, 25^\circ\ \mathrm{W}] \times [30^\circ\ \mathrm{N}, 50^\circ\ \mathrm{N}]$, and divide this into 150×90 boxes. After estimating the entries of the transfer operator matrix $\mathbf{P}$, the first ten singular values of this matrix were computed to be 1, 0.9651, 0.9507, 0.9448, 0.9390, 0.9340, 0.9317, 0.9299, 0.9291 and 0.9259. In Figure 3.55, we produce two pairs of coherent sets for the Gulf Stream flow using both the second and third right singular vectors of $\mathbf{P}$. The results from both of these singular vectors are consistent with the results obtained from the dynamic Laplace operator by Froyland et al. (2019). Using the second right singular vector, the Gulf of St Laurence and its immediate surrounds are identified as one coherent set, and the remainder of the domain is identified as the other coherent set. Considering the third right singular vector instead, we identify

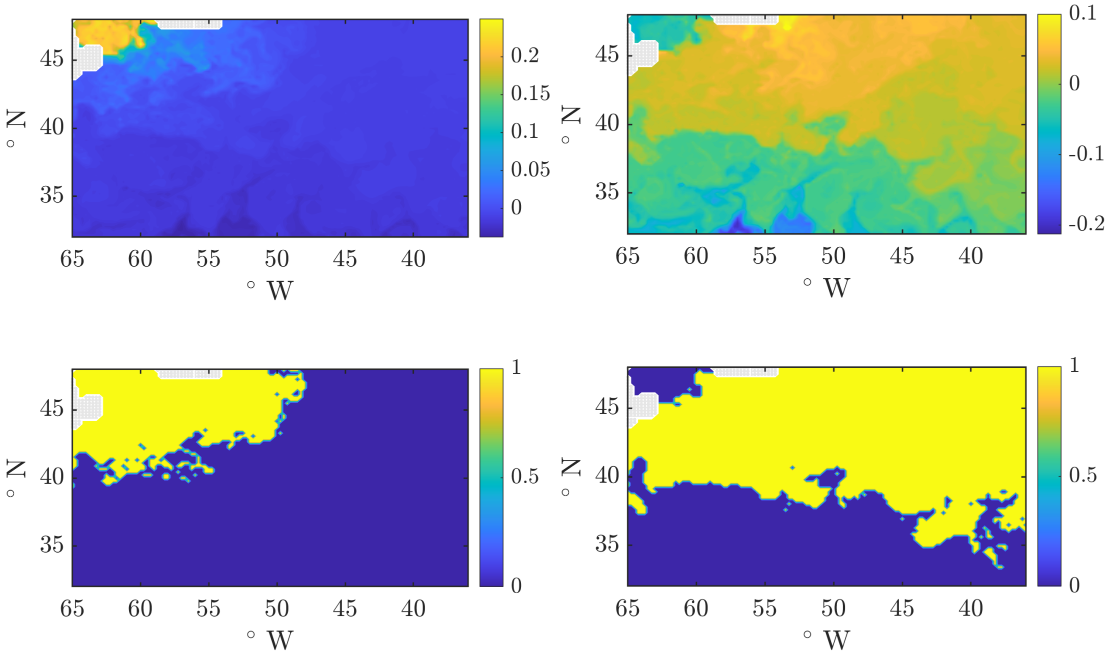

Figure 3.55: Scalar fields of the second (top–left) and third (top–right) right singular vectors of **P** for the Gulf Stream flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

most of the top half of the flow domain as one coherent set, with the other coherent set featuring a disjoint combination of the lower half of the flow domain and the Gulf of St Laurence. Interestingly, the contour which separates the upper half of the domain from the lower half appears to take the shape of part of the Gulf Stream itself, though as we can see in the third singular vector scalar field, larger positive values of the third singular vector entries curl up towards the upper boundary of the domain, in similar fashion to the Gulf Stream shapes produced from previously discussed methods, such as the FTLE, LAVD and S^2.

We obtain similar coherent sets for this flow system using the dynamic Laplace operator method. Defining a matrix operator for the Laplacian over this domain is rather difficult, particularly due to its awkward shape caused by the presence of land in the top–left corner of the domain. Hence, for this flow we will replicate the calculations undertaken by Froyland et al. (2019) using the sparse version of this method and the same set of computational parameters. The first ten eigenvalues produced for $\mathbf{\Delta}^D$ from this method have come out to be 0, -0.7096, -1.5057, -1.8395, -2.1443, -2.8906, -2.9380, -3.5367, -4.2314 and -4.3876. Figure 3.56 shows two pairs of coherent sets produced for the Gulf Stream flow using both the second and third eigenvectors of $\mathbf{\Delta}^D$. Like with the second singular vector of **P**, the first coherent set obtained from the second eigenvector of $\mathbf{\Delta}^D$

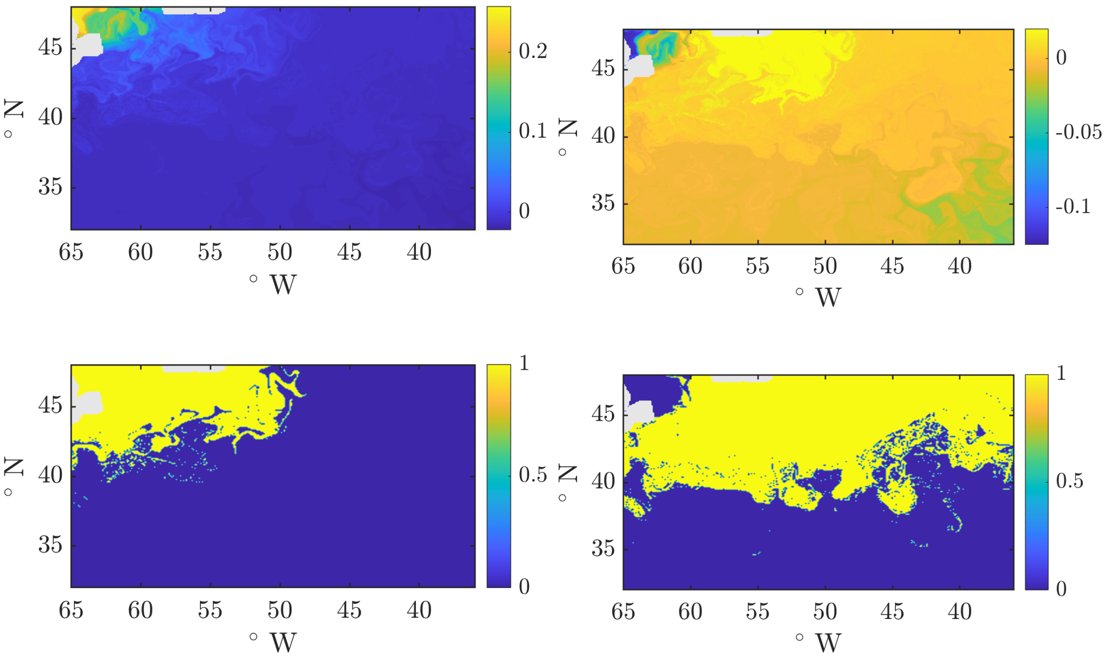

Figure 3.56: Scalar fields of the second (top–left) and third (top–right) eigenvectors of $\mathbf{\Delta}^D$ for the Gulf Stream flow and two coherent sets for the flow obtained by dividing the domain Ω along the zero contour lines of each respective field (bottom Figures).

includes the Gulf of St Laurence and its surrounds, while the second coherent set includes the remainder of the domain. When we use the third eigenvector of $\mathbf{\Delta}^D$, the majority of the top half of the domain forms one coherent set, while the other set contains most of the Gulf of St Laurence along with the bottom half of the domain in a disjoint fashion. Bar some minor differences, the results are consistent with those produced by Froyland et al. (2019).

Next, we apply the FCM clustering method to this system, firstly by seeking $K = 2$ clusters with the sharpness parameter m set to a value of 1.5. The membership probability fields and two clusters obtained at the 75% threshold can be seen in Figure 3.57. Unlike in some of the previously discussed methods, such as the transfer operator and dynamic Laplace operator methods, this method has divided the domain more or less by its latitudinal centre, with most of the left hand side of the domain forming one coherent cluster, and most of the right hand side of the domain forming the other cluster. The second cluster (the red cluster in Figure 3.57) contains a horizontal streak of flow surrounded by the first (blue) cluster, illustrating a portion of the flow domain dragged across to the right hand side of the domain by the Gulf Stream as it evolves over the duration of the flow. We then increase the number of clusters K to 5, with the resultant membership probability fields and five FCM clusters produced again at the 75% threshold shown in

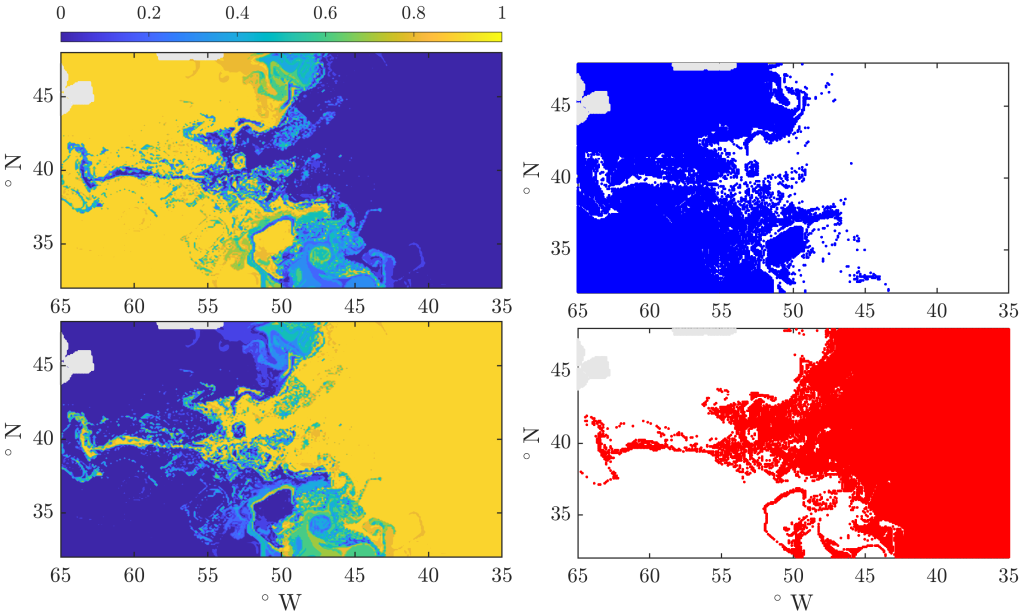

Figure 3.57: Membership probability fields for two FCM clusters of the Gulf Stream flow (left Figures) and two clusters for this flow obtained at the 75% membership probability level (right Figures).

Figure 3.58. This time, the FCM clustering method has divided the domain more or less into five coherent clusters of equal size, and has not been able to identify the Gulf Stream or any other coherent flow objects detected by other methods (such as the eddies).

Finally, we apply the CSC method to the Gulf Stream flow system. Due to system memory constraints in MATLAB, we could only use a smaller initial grid of 210 × 150 uniformly spaced initial points (disregarding all points based on land once more), to execute this method. The resultant scalar field produced from this method can be seen in Figure 3.59. Like with the S^2 scalar field, there is a level of graininess and a slight lack of clarity with the scalar field produced, however it is not as detrimental to the method in this case. From the largest positive CSC coefficients, we are able to visualise a large collection of coherent eddies and eddy–type pockets of flow which exist within this system, along with the majority of flow within or surrounding the Gulf of St Laurence. Negative values of this coefficient identify streams of flow between coherent eddies, particularly in the bottom–right corner of the domain, along with the Gulf Stream itself; though the identification of this is hindered slightly by the graininess of the CSC field within the Stream.

The Gulf Stream system is arguably the most complex flow we have considered in this study, owing to its complicated combination of flow features, from the large and rapidly moving Gulf Stream to the coherent eddies produced from this Stream; all of

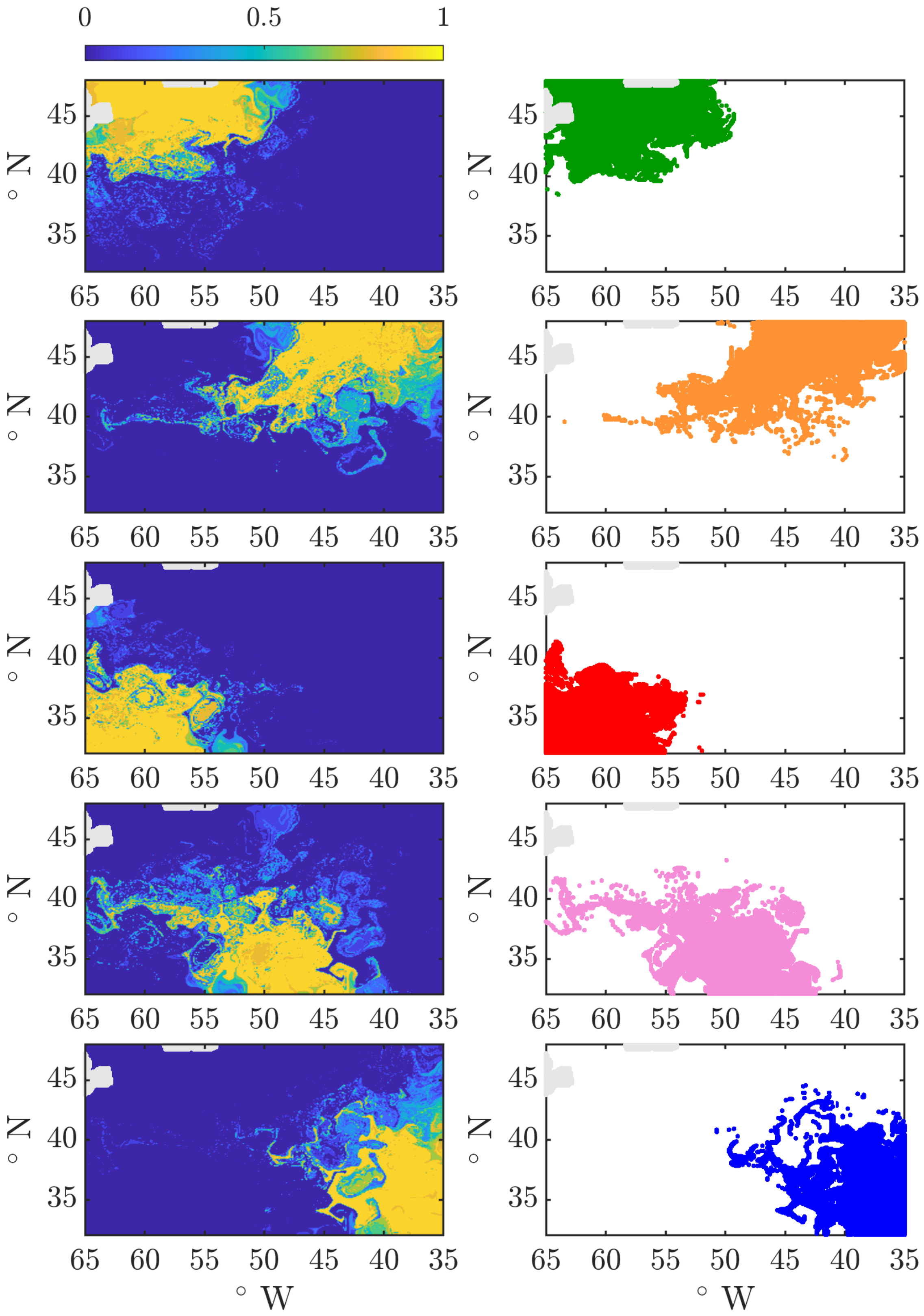

Figure 3.58: Membership probability fields for five FCM clusters of the Gulf Stream flow (left Figures) and five clusters for this flow obtained at the 75% membership probability level (right Figures).

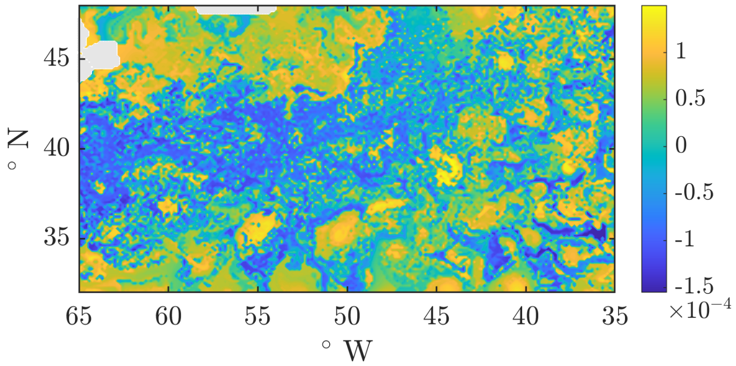

Figure 3.59: The CSC scalar field for the Gulf Stream flow.

which move through the system in unpredictable patterns. As a result of this, each of our methods detect a different collection of LCSs in a variety of different ways. The most influential structure in this flow is the Gulf Stream itself, due to its consistency and influence over the system; even though it is not strictly a "coherent" structure. The Gulf Stream is identifiable from most of the methods considered earlier, including the LAVD (from strong Lagrangian shear), the Scaled S^2 field, the transfer operator, the dynamic Laplace operator, and the CSC field. The FTLE can also be used to find the Gulf Stream, though high particle stretching spread out across the domain makes this difficult to isolate. The eddies produced by the Gulf Stream are also key structures within this flow, and have been identified using the LAVD, CSC, low values of the FTLE and the S^2 robust sets (though L needs to be large and the structures are grainy in resolution). The Gulf of St Laurence was also identified as a coherent flow region as flow trajectories within this region remain separate from the greater Atlantic Ocean due to the Gulf Stream. The transfer operator and dynamic Laplace operator methods explicitly identified this region, though it is also identifiable from the S^2 robust sets, CSC and the FTLE. The two methods which have fared the worst in identifying meaningful coherent structures within this flow are the VLCS and FCM clustering methods. As mentioned earlier, a significant handicap of the FCM clustering method is that the number of clusters expected to exist within a flow needs to be known or guessed. In a real world data set such as this one, predicting the exact number of coherent clusters expected to exist within the flow is difficult to do, and regardless of the value of K chosen, the FCM Clustering method will most likely just divide the flow domain into K clusters of roughly equal size; as was done with this system in both the $K = 2$ and $K = 5$ cases. Each VLCS algorithm has identified an incoherent collection of flow separating barriers spread out across the flow domain, with no meaningful flow objects identifiable from these structures (such as the boundary of the Gulf Stream or a coherent eddy).

Chapter 4

Sensitivity of Lagrangian Coherent Structures to Velocity Uncertainty

In this Chapter, we examine how each of the LCS detection methods discussed in the previous two Chapters respond to noise or uncertainty present within the velocity data which represents a dynamic flow system. The vast majority of observational or numerically defined velocity data sets available will carry some amount of numerical uncertainty, due to observational, computational or experimental errors. Depending on how severe this numerical uncertainty is, this is likely to have an impact on the reliability and accuracy of the results generated from various LCS detection methods. Other studies have investigated the impact of stochastic uncertainty on the functionality of LCS detection methods such as the FTLE (Guo et al. 2016) and Lagrangian descriptors based on trajectory length (Balibrea-Iniesta et al. 2016); with Balasuriya (2020*b*) also formulating an analytically defined quantification of the error for the FTLE using S^2. We therefore aim to add to this work by quantifying the impact of this uncertainty on all of our eight LCS detection methods.

A reminder that a stochastic differential equation (SDE) system takes the form

$$\mathrm{d}\mathbf{x}_t = \mathbf{u}\left(\mathbf{x}_t, t\right)\ \mathrm{d}t + \epsilon\ \boldsymbol{\sigma}\left(\mathbf{x}_t, t\right)\ \mathrm{d}\mathbf{W}_t. \tag{4.1}$$

As has been done in the literature (Balibrea-Iniesta et al. 2016, Guo et al. 2016), we can use the above equation to convert the Eulerian velocity $\mathbf{u}$ of a deterministic dynamical system defined by (2.1) into a stochastic flow system which adds normally distributed uncertainty to our deterministic velocity data through the two–dimensional Brownian motion defined by $\mathbf{W}_t$. How much uncertainty we add to our original velocity data depends on the noise parameter ϵ which, in a dimensional flow system, like all of the flows which will be dealt with in this Chapter, is defined by $\sqrt{h v_r}$. In this expression, h is a lengthscale parameter which represents the resolution of the available velocity data and is determined by the mesh size of this data (or the mesh size of the grid of initial points $\mathbf{x}_0$ if the flow is analytically defined); and v_r represents the scale of uncertainty in this velocity data which can be known or guessed based on the precision of the velocity data. The parameter $\boldsymbol{\sigma}$ in (4.1) represents a diffusion matrix for the stochastic volatility imposed on $\mathbf{u}$. In this research, $\boldsymbol{\sigma}$ is dimensionless as we have chosen ϵ to be dimensional and will simply be set to the default choice $\mathbf{I}_{2\times 2}$, the 2×2 identity matrix (Balasuriya 2020*a*,*b*).

We begin our stochastic noise analysis by converting the Eulerian velocity for three of our flow systems (namely the Bickley Jet, KH2 and Gulf Stream flows) into stochastic flow systems of the form (4.1). The value of the noise parameter ϵ will change with each flow system as the velocity resolution parameter h changes, however we will keep the velocity uncertainty level v_r consistent across all of these flows. We then solve (4.1) for all of our flows over an initial grid of points $\mathbf{x}_0 \in \Omega$ using the Euler–Maruyama integration scheme, ensuring that the same amount of normally distributed stochastic noise is applied to every point within the initial grid so that continuity necessary for the computation of quantities such as the flow map gradient is maintained. We then produce a stochastic realisation of the LCS results from each method using the flow solutions of (4.1), and start by qualitatively assessing how the stochastic LCS results differ from the deterministic LCS results obtained earlier.

We then perform quantitative analysis of these stochastic LCS results by generating 100 realisations of LCS quantities such as the FTLE, the coefficients of the optimal vector of $\mathbf{P}$, or the membership probability of an FCM cluster for each of our flow systems; iteratively calculating the mean and standard deviation of these quantities as these realisations are produced. After some experimentation, 100 was found to be a sufficient number of stochastic realisations needed to obtain convergence of the means and standard deviations for each of these methods applied to all of the flows considered, as we demonstrate in Appendix B of this Thesis. From these results, we should be able to determine common LCS patterns from the mean fields of these quantities, while the standard deviation fields tell us within which areas of the flow domain our LCS methods are more susceptible to this velocity uncertainty.

This approach cannot be used to quantify the impacts of stochastic noise on the VLCS method or the S^2 robust sets, as these methods produce curves or subsets of the flow domain Ω rather than scalar fields. Instead, we estimate the probability/likelihood of a portion of the domain being traversed by a VLCS or being contained within a robust set from 100 stochastic realisations of these methods. In the case of the VLCS method, we divide Ω into “bins” of equal size, record which of these bins each hyperbolic VLCS traverses through for each realisation of the VLCS method and use this to produce a sample probability (likelihood) field that should reveal curve patterns which these hyperbolic VLCSs would be expected to follow. We also include a similar field showing the iteratively calculated variation (standard deviation) of these sample probabilities. We will only carry out this analysis using the VLCS–B algorithm as the underlying computations necessary to produce these VLCSs remain the same for all three algorithms, so there is no point in repeating this analysis for the other two algorithms. To perform quantitative analysis on the realisations of the S^2 robust sets, we generate a probability density function for the robust set based on the points which form each stochastic realisation of this set.

It is worth noting that we do not use one single batch of 100 stochastic flow map realisations to perform quantitative analysis of each of our LCS detection methods. The differences in the calculations required to carry out each of these methods prevent us from doing this. For instance, to compute the FTLE we define a uniform grid of initial points, flow these forward and use the maximal eigenvalues of the Cauchy–Green strain tensor to obtain a scalar field for the FTLE. Compare this with the transfer operator method, where we randomly seed 1000 initial flow points within each of a collection of boxes at time t_0 to estimate the entries of the transition matrix $\mathbf{P}$; and we therefore would not be using the same initial flow points as those used to generate the FTLE. From the results produced, no significant difference between the stochastic LCS means/standard deviations produced from two independent samples of 100 stochastic flow trajectories was observed (we will not document this in the Thesis).

4.1 Bickley Jet

We begin with the Bickley Jet defined over the flow period ranging from $t_0 = 0$ to $t_0 + T = 40$ days, examined in Section 3.3. In performing noise analysis for this system (and all other flows considered in this Chapter), the values of all computational parameters relevant to each individual method (such as the number of initial grid points for flow map computation or the box configurations taken for the transfer operator or dynamic Laplace operator) will remain the same as defined in the previous Chapter. Seeing as the Bickley Jet is an analytically defined flow, we define the horizontal mesh size of the uniform grid of initial points $\mathbf{x}_0$ as the resolution lengthscale parameter $h \approx 2.22 \times 10^{-2}$ Mm. We estimate the velocity uncertainty level $v_r \approx 1.4 \times 10^{-3}$ Mm per day, thereby giving $\epsilon \approx 5.6 \times 10^{-3}$ Mm per square root day for use in (4.1) to produce the stochastic flow map realisations.

We begin with the FTLE method, which has shown to be reasonably self–consistent under the influence of velocity uncertainty as we gather from the results observable in Figure 4.1. For all of the LCS detection methods applied to the Bickley Jet system, we plot twice the standard deviation or variation in a quantity to emphasise key features present within these fields which could not otherwise be made out if just 1× the standard deviation was plotted. Upon application of stochastic noise to the Bickley Jet velocity, the FTLE scalar field proceeds to wobble quite considerably, particularly in the vortex layers between the coherent eddies where particle flow is exceptionally volatile. Despite this, the zonal jet and the centres of the vortices appear to remain consistent in shape. The mean FTLE field appears to look like a smoothened out version of the deterministic field, while the standard deviation is, as expected, larger within the volatile flow regions between the coherent vortices, and considerably small within the vortices themselves along with the

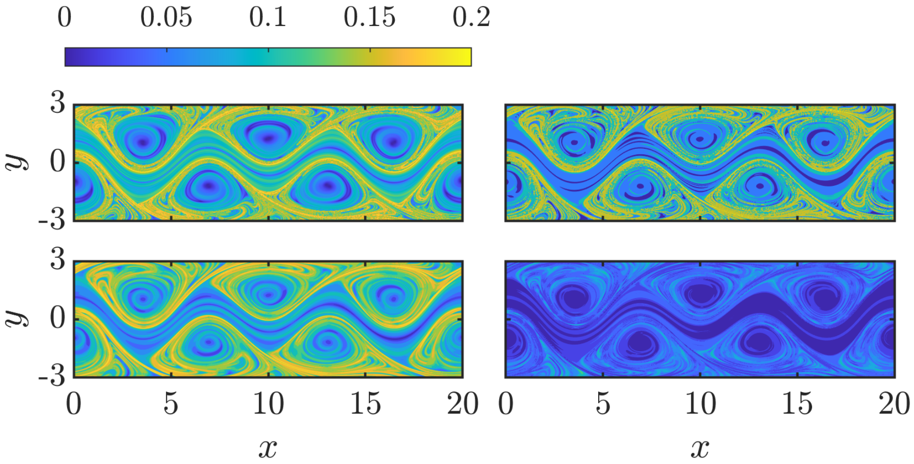

Figure 4.1: The deterministic FTLE of the Bickley Jet flow (top-left), one realisation of this quantity with noise applied to the velocity (top-right) and the mean (bottom-left) and twice the standard deviation (bottom-right) of 100 realisations of the stochastic FTLE.

zonal jet.

Unlike the FTLE method, the VLCSs produced for this system are highly susceptible to velocity noise and irregularity; as the results shown in Figure 4.2 illustrate. The stochastic hyperbolic VLCSs produced completely deform in shape in comparison to their deterministic counterparts. We divide the domain Ω into $180 \times 60 = 10800$ equally sized rectangular bins to produce the likelihood field. Each bin in this field carries a value (between 0 and 1) which indicates the probability of a hyperbolic VLCS generated under stochastic conditions entering that particular bin based on all of the VLCSs produced from each realisation. Some sinusoidal green curves are identifiable within this field, indicating that a VLCS of this shape is somewhat quite robust and (at the very least) a portion of this structure will be identified even with stochastic noise added to the velocity. However, the large amount of light blue coloured streaks of incoherent shape scattered throughout the field demonstrate a lack of consistency in the hyperbolic VLCSs produced for this system, as a large number of deformed curves have been produced for this system under the influence of noise. The variation scalar field illustrates the standard deviation of the likelihood for each of the bins, and it appears that this variation is consistently high throughout this flow domain, which casts further doubt over the validity of the results produced by the VLCS method. The results that we observe here are most likely owed to the fact that the VLCS method relies on calculating numerical flow map gradients, performing interpolation on eigenvector fields to produce solution curves and regulating the integration of these fields as it progresses. These processes will come with a great deal of numerical estimation error and uncertainty; which will only be amplified through the existence of extra uncertainty within the velocity data.

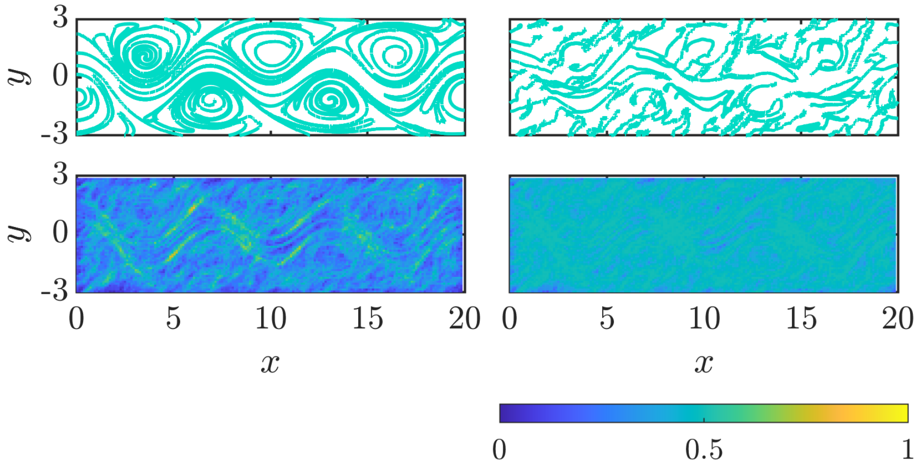

Figure 4.2: Deterministic hyperbolic VLCSs of the Bickley Jet flow (top–left), one realisation of these VLCSs with noise applied to the velocity (top–right) and the likelihood (bottom–left) and twice the variation in this likelihood (bottom–right) of a VLCS passing through one of $180 \times 60 = 10800$ bins computed from 100 stochastic realisations of these VLCSs.

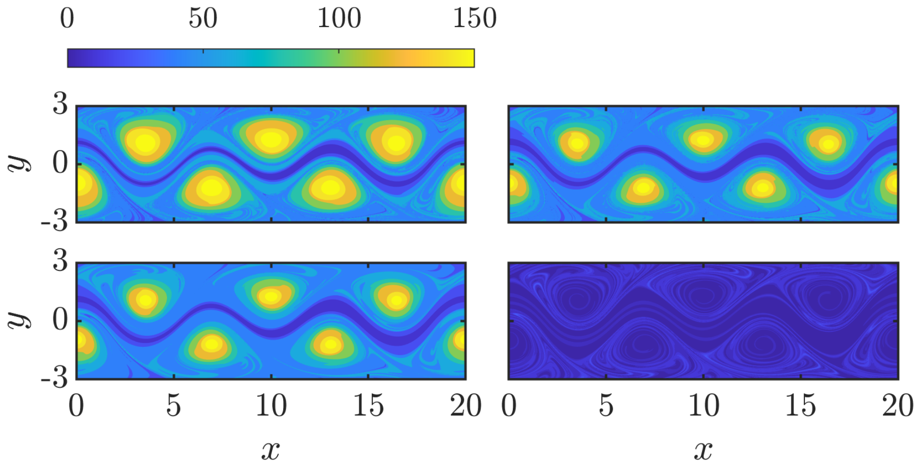

Figure 4.3: LAVD results for the Bickley Jet flow arranged in the same fashion as seen in Figure 4.1.

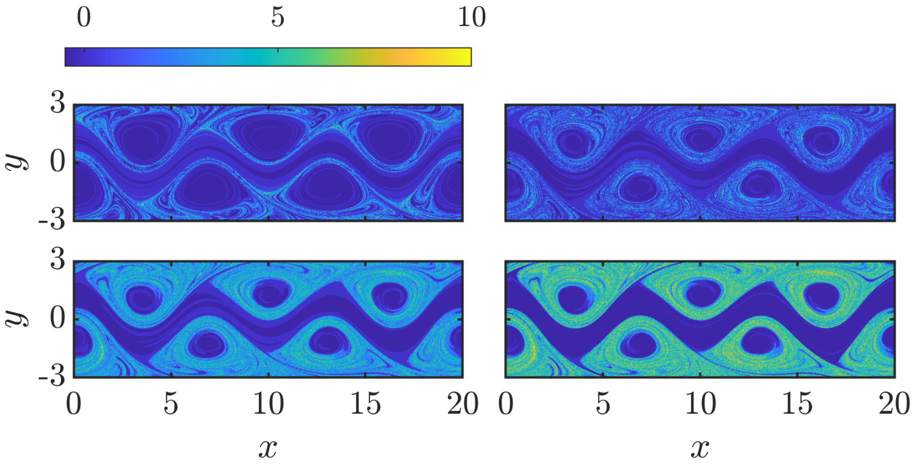

Figure 4.4: Scaled S^2 diagnostics for the Bickley Jet flow arranged in the same fashion as seen in Figure 4.1.

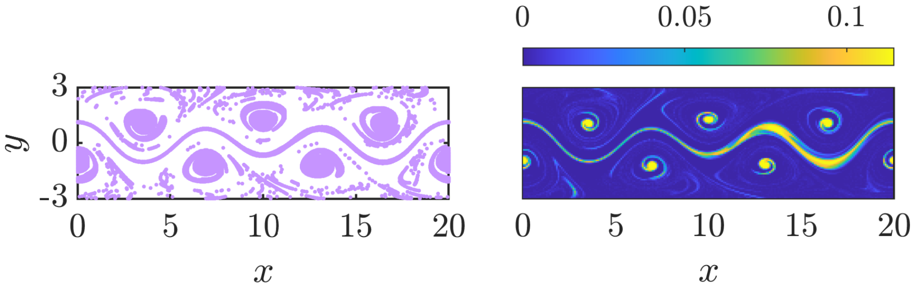

Figure 4.5: The robust set of lengthscale $L = 0.3$ Mm for the Bickley Jet flow (left) and the probability density function obtained from the stochastic realisations of this set (right).

The LAVD for the Bickley Jet is structurally robust against velocity uncertainty imposed on the system, as the results shown in Figure 4.3 attest. The stochastic LAVD illustrates the same LCSs (such as the vortices and the zonal jet), though the LAVD values between the vortices have changed and the vortical structures corresponding to maximal LAVD values appear to have shrunk in size. The mean field looks more like a smoothened out version of the stochastic LAVD field rather than the deterministic one, and the standard deviation is overall quite small and concentrated within the flow between the coherent vortices and along the edges of these vortices. This method is robust to velocity uncertainty in the sense that we detect similar LCSs amidst the presence of this uncertainty, however the LAVD quantity itself appears to be less robust due to fluctuations in its values and minor alterations to the shapes of the LCSs detected.

The results generated from the S^2 method show a considerable amount of suscepti-

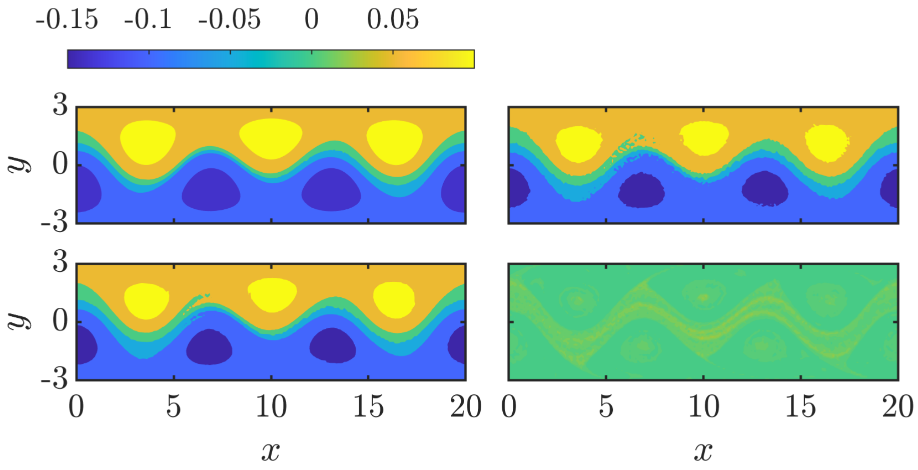

Figure 4.6: Transfer operator optimal vector diagnostics for the Bickley Jet flow arranged in the same fashion as seen in Figure 4.1.

bility to the noise imposed on the Bickley jet velocity, as the results shown in Figure 4.4 illustrate. The maximal ridges of the stochastic realisation of the Scaled S^2 change shape considerably, with noticeable wobbling in the ridges and maximal spikes of this quantity; along with the shrinkage of the Bickley jet vortices. The mean field generated shows that the largest Scaled S^2 values are concentrated within the flow between the vortices, while smaller Scaled S^2 values within the (smaller) vortices and the zonal jet remain much more consistent. The standard deviation of the Scaled S^2 is also quite large, particularly between the vortices where we observe these larger Scaled S^2 values. The large variation in this quantity can be attributed to the sharp increases in the values of S^2 which result in large gradients for this quantity even after we have scaled it (Balasuriya 2020*a*), which therefore causes more rapid changes to the values of S^2 amidst the presence of even mild amounts of velocity noise or uncertainty.

While the S^2 itself appears to be rather sensitive to this velocity uncertainty, the robust sets produced from this quantity are more self–consistent as we see from the results shown in Figure 4.5. We focus on the robust set relevant to a lengthscale $L = 0.3$ Mm, and produce a probability density function for this set from all of the stochastic realisations of S^2 produced. Large values of this function are observable within the centres of the Bickley jet vortices along with the zonal jet, which indicate that these structures are almost guaranteed to be included within this robust set even with this level of velocity uncertainty. The remainder of the domain corresponds to low values of the probability density function, indicating little to no probability of flow particles beginning within these regions belonging in this set at this velocity uncertainty level; with few regions of the domain corresponding to a medium level of this probability density function (aside from some thin light blue to green coloured streaks which surround the vortices).

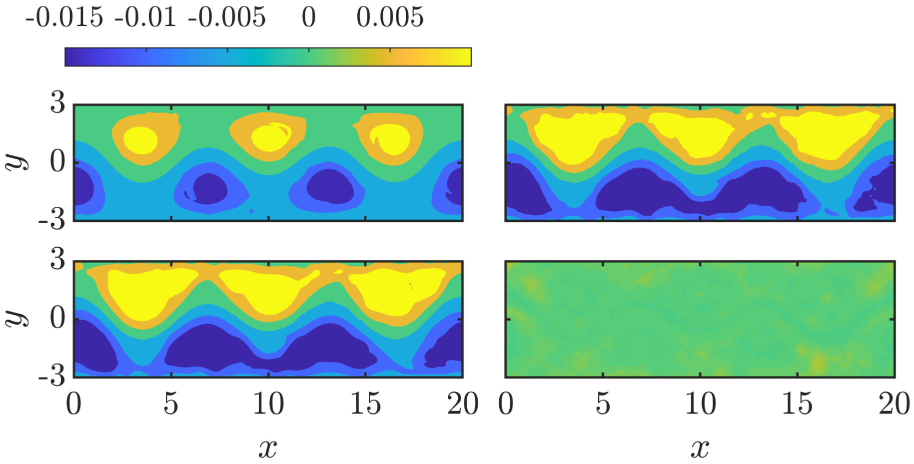

Figure 4.7: Dynamic Laplace operator second eigenvector diagnostics for the Bickley Jet flow arranged in the same fashion as seen in Figure 4.1.

The scalar field for the optimal vector of **P** shows considerable robustness to this velocity uncertainty, as we see from the results presented in Figure 4.6. The stochastic optimal vector scalar field does not differ too much from the deterministic result, save for the considerable amount of fuzziness observable around the edges of key coherent structures such as the zonal jet or the coherent vortices. These inconsistencies appear to have been smoothened out in the mean optimal vector field, with the structure of this field appearing more like that of the stochastic result rather than the deterministic result, owing mainly to the shrinkage of the Bickley jet vortices. The standard deviation field shows low variation in the optimal vector entries over the domain, with this variation concentrated more in or around the zonal jet and the vortices, accounting for the fuzziness observable in these structures under the influence of stochastic velocity noise.

The second eigenvector of the dynamic Laplace operator appears to have been slightly more compromised under the influence of stochastic velocity uncertainty, as we observe from the results displayed in Figure 4.7. The vortices produced in the stochastic realisation of this eigenvector have become larger in size and look considerably squashed or smudged. The flow between each vortex within both of these layers is highly volatile even without stochastic velocity uncertainty, and this combined with the numerical uncertainty triggered by the estimation of the Laplacian matrix operator is likely what is causing these anomalous results. As we will show in the next two Sections of this Chapter, this behaviour is not observable for the dynamic Laplace operator method in the other flow systems considered. It is worth noting that if the overall goal of this method is to divide our flow domain into two coherent sets, and the zonal jet (which acts as the separating boundary between these sets in this system) does not change shape or become too fuzzy under the influence of stochastic velocity noise, the division will be consistent with that

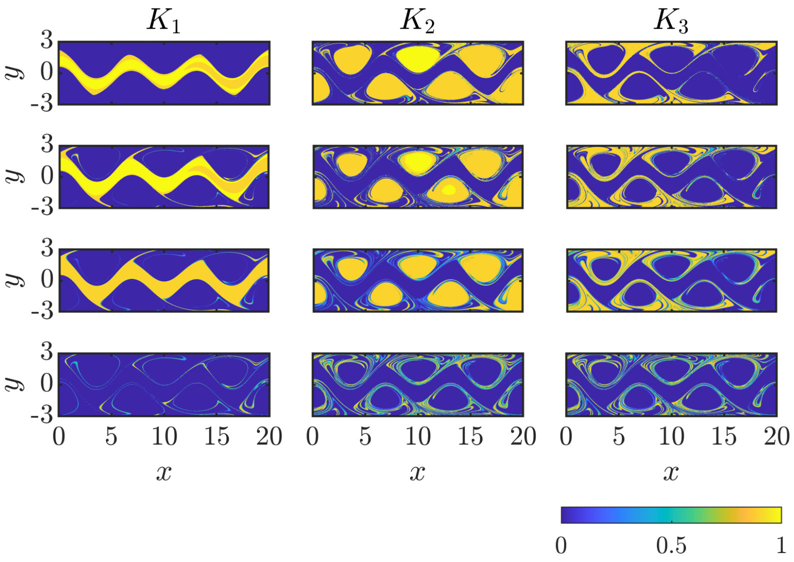

Figure 4.8: The deterministic membership probability fields for three FCM clusters relevant to the Bickley Jet flow (first row); one realisation of each of these fields with noise applied to the velocity (second row) and the mean (third row) and twice the standard deviation (fourth row) of 100 stochastic realisations of these probabilities.

achieved in the deterministic case. The mean eigenvector field looks much more like the stochastic realisation of this field, and the standard deviation is low and is concentrated around the edges of the vortices. Unlike with the transfer operator method (see Figure 4.6), we do not obtain high values of the standard deviation within the zonal jet or the vortices themselves.

The FCM clustering method also shows a sufficient degree of self–consistency under the influence of added velocity uncertainty, as observable by the results shown in Figure 4.8. We consider the $K = 3$ case for this system, and we undertake quantitative analysis on the membership probability fields relevant to each FCM cluster. In the fields pertaining to the stochastic realisations of these probabilities, we see only a minor deviation in the first cluster (representing the zonal jet), and a slight change in the sizes of the coherent vortices (the second cluster) as their edges become part of the chaotic sections of flow sitting between these vortices (the third cluster). The mean membership probability fields look like smoothened out versions of the stochastic realisations (a common observation amongst most of our methods so far), though the general formation of each cluster is more or less consistent in comparison to our deterministic result, bar the aforementioned alterations. The standard deviation is low in the case of the zonal jet cluster, but is slightly larger and more or less consistent between the vortex cluster and the rapid vortex layer flow cluster. In each case, the standard deviation is strongest around the edges of each Bickley jet vortex, where we see strongly chaotic particle flow behaviour and where

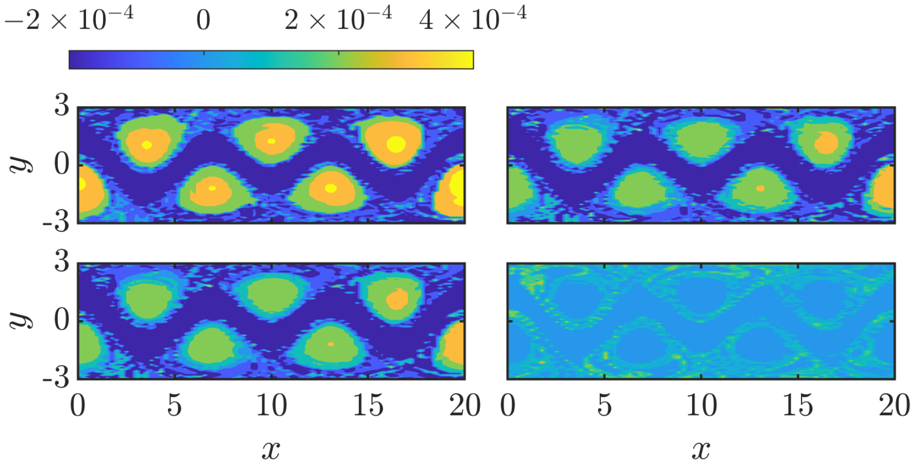

Figure 4.9: CSC diagnostics for the Bickley Jet flow arranged in the same fashion as seen in Figure 4.1.

the cluster classifications of each flow trajectory around these edges are most likely to switch with each stochastic realisation.

Finally we consider the CSC method, which is considerably robust to the presence of added stochastic velocity uncertainty as demonstrated by the results shown in Figure 4.9. The extremal values of the CSC coefficients in the stochastic realisation of this method considered have changed, though the same structures are visible and comprehensible from this scalar field and there is minimal fuzziness or lack of resolution observable within the field, in the same fashion as what we observed within stochastic realisations of the FTLE or transfer operator fields. In consistent fashion with most of the previously discussed methods, the mean field looks almost identical to the stochastic version of the CSC field instead of the deterministic one, due to the differences observable in the maximal values of the CSC coefficients between the mean field and the deterministic field, and the standard deviation is low with larger values of this more prominently observable in the chaotic flow regions lying between the Bickley jet vortices.

4.2 Kelvin–Helmholtz Version 2 (KH2)

We next turn our attention to the KH2 flow system, discussed earlier in Section 3.5. We consider the same time parameterisation used for the KH2 flow in Section 3.5, which ranges from $t_0 = 12$ seconds to $t_0 + T = 15$ seconds, and we do not change any of the values for the computational parameters pertinent to each LCS detection method when applied to this flow. As mentioned earlier, we wish to keep the velocity uncertainty scale v_r consistent amongst all three flow systems considered in this Chapter. This parameter now takes a value of 1.63×10^{-2} ms^{-1}, after converting v_r for the Bickley jet flow into the relevant units for this system. The mesh size for the data points within this numerically defined velocity system $h = 2\pi/1024 \approx 6.1 \times 10^{-3}\,\mathrm{m}$, which corresponds to a value of $\epsilon = 10^{-2}\ \mathrm{ms}^{-1/2}$, which is the value we insert into the SDE system (4.1) to produce the stochastic flow maps for this system.

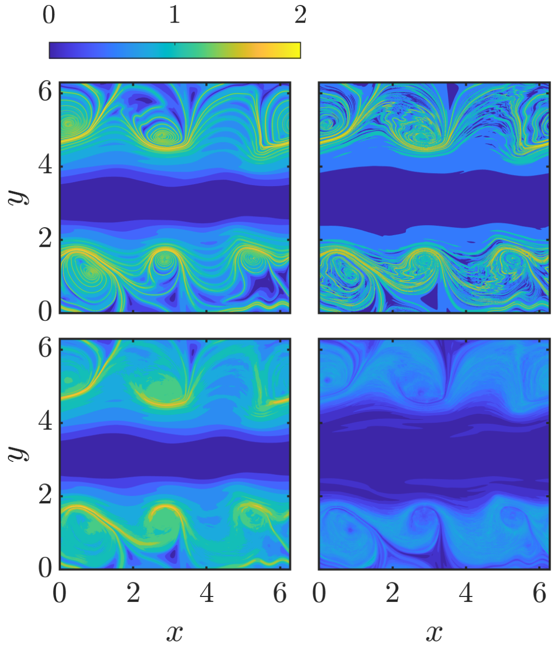

Figure 4.10: The deterministic FTLE of the KH2 flow (top-left), one realisation of this quantity with noise applied to the velocity (top-right) and the mean (bottom-left) and twice the standard deviation (bottom-right) of 100 realisations of the stochastic FTLE.

We begin again with the FTLE method, which is once again reasonably self–consistent under the influence of added velocity uncertainty as we can see from the results shown in Figure 4.10. We again plot twice the standard deviation or variation of a quantity in order to emphasise important features present within these fields which were not clearly observable when only 1× the standard deviation was plotted. As was the case with the Bickley jet system, the addition of stochastic noise to the KH2 velocity data has resulted in wobbling of the scalar field and the maximal ridges of the field becoming jagged and incoherent. Entities corresponding to low FTLE values, such as the central flow channel, have more or less maintained their shape. The mean field obtained looks like an ironed out version of our deterministic field which has lost the intricate detail of said field, including ridges corresponding to medium level values of the FTLE. That said, maximal ridges sitting between vortices and the flow channel are still identifiable, as is the flow channel itself. As expected, the standard deviation is largest within the Kelvin–Helmholtz vortex layers, between the vortices where chaotic flow particle advection is stronger.

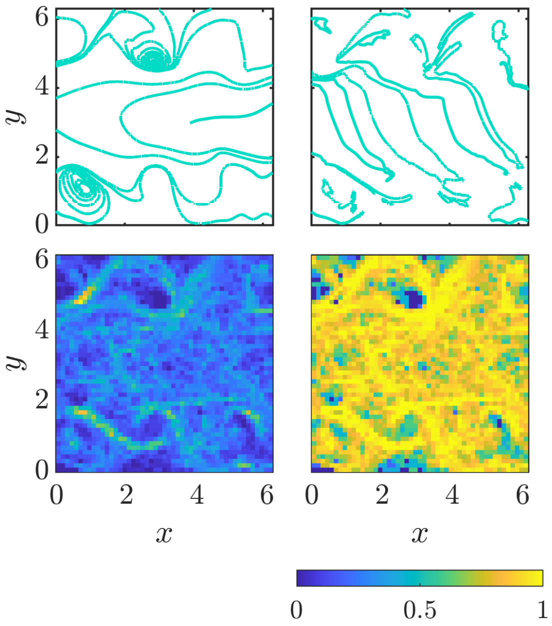

Figure 4.11: Deterministic hyperbolic VLCSs of the KH2 flow (top–left), one realisation of these VLCSs with noise applied to the velocity (top–right) and the likelihood (bottom–left) and twice the variation in this likelihood (bottom–right) of a VLCS passing through one of $50 \times 50 = 2500$ bins computed from 100 stochastic realisations of these VLCSs.

The self–consistency of the VLCS method has again been compromised under the influence of velocity uncertainty, as observed from the results shown in Figure 4.11. Applying noise to our velocity data has once again resulted in the total disfigurement of the hyperbolic VLCSs produced for this system in comparison to our deterministic results. To produce a likelihood field from 100 of these stochastic realisations, we this time divide the flow domain into $50 \times 50 = 2500$ equally sized bins so that the width of each of these boxes is approximately the same as the separation radius parameter ρ used to carry out the VLCS–B algorithm. From this field, despite the presence of some coherent hook shaped VLCSs which surround the coherent vortices within the vortex layers, we see the trace of these hyperbolic VLCSs more or less covering the flow domain, including the central flow channel which these structures should not traverse. The variation in this likelihood supports the poor self–consistency of this method under even mild velocity uncertainty, as this variation is considerably high across the whole domain, except for partitions of the flow lying closer to the centres of the coherent vortices where we detect almost no hyperbolic VLCSs at all.

The LAVD method again demonstrates structural, but not numerical, self–consistency under the influence of velocity noise as seen by the results shown in Figure 4.12. In this stochastic realisation of the LAVD, we detect the vortices in this system in approximately

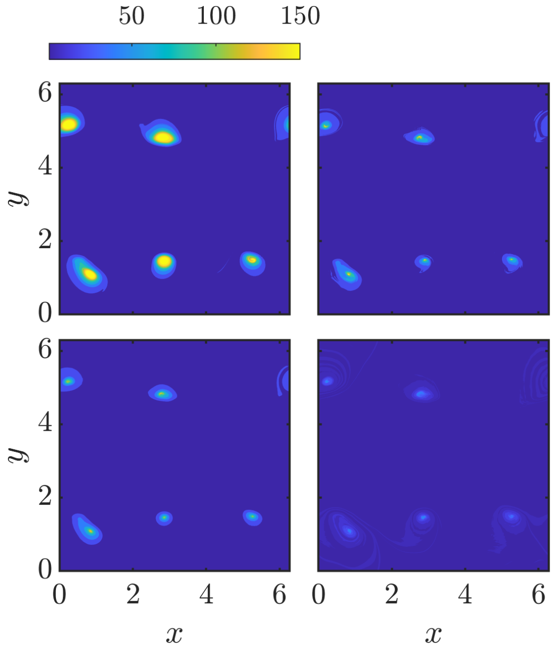

Figure 4.12: LAVD diagnostics for the KH2 flow arranged in the same fashion as seen in Figure 4.10.

the same position as the deterministic results. However, these vortices have diminished in size and the maximal values of the LAVD recorded have decreased. The mean LAVD scalar field appears to look more like an ironed out version of the stochastic realisation of the scalar field rather than the deterministic field, as the vortices in the mean field are not as large as those in the deterministic field and do not correspond to the same LAVD values recorded in the deterministic case. The standard deviation is considerably high and concentrated around the centres of our vortices. While the method was able to find the same vortices in the same locations of the domain, the shapes and the corresponding LAVD values of these structures have considerably changed, so there is some partial self–consistency observable for the LAVD in this system. We did not observe a strong change in the LAVD values in the centres of the Bickley jet vortices, which most likely owes to the fact that those vortices remain in one position and do not change shape or merge with other vortices as the flow proceeds.

We again observe some inconsistencies in the values of the S^2 for this system when stochastic noise is applied to the velocity, as we see from the results shown in Figure 4.13. The maximal ridges of the stochastic Scaled S^2 field become jagged and much less smooth, in similar circumstances to the FTLE results for this system. The mean field produced looks like a smoothened out version of the stochastic version of the S^2 field, not the deterministic field as the mean S^2 maximal ridges are not as smooth as the deterministic maximal ridges. The standard deviation is again considerably high closer to

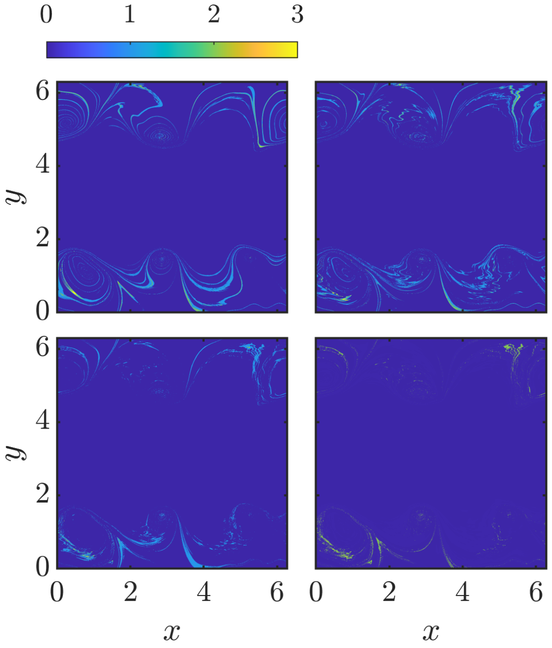

Figure 4.13: Scaled S^2 diagnostics for the KH2 flow arranged in the same fashion as seen in Figure 4.10.

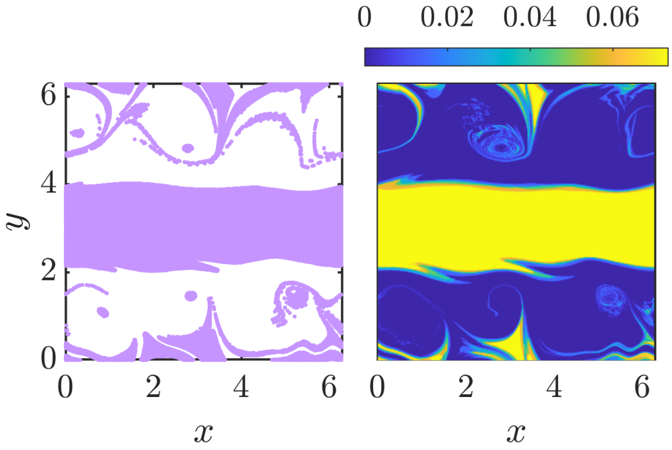

Figure 4.14: The robust set of lengthscale $L = 0.25$ m for the KH2 flow (left) and the probability density function obtained from the stochastic realisations of this set (right).

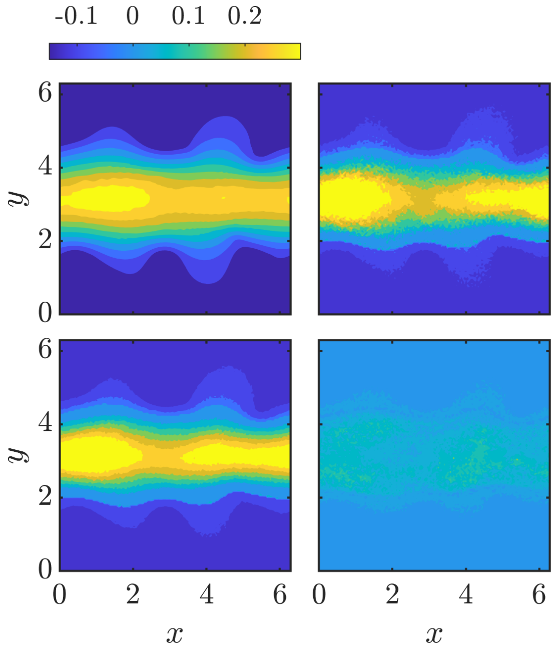

Figure 4.15: Transfer operator optimal vector diagnostics for the KH2 flow arranged in the same fashion as seen in Figure 4.10.

these maximal ridges, which again is to be anticipated given the large spikes observable in this quantity and the consequential high gradients which cause more rapid changes in this quantity when noise is applied to the velocity of this flow.

That said, as the results in Figure 4.14 attest, the robust set manages to withstand this added velocity uncertainty more considerably. We present a probability density function from the 100 $L = 0.25$ m robust sets produced from each stochastic realisation. Each of these realisations is almost guaranteed to contain the yellow regions of this function, including the central flow channel and certain flow structures which surround the centres of the vortices. Detecting the centres of the vortices themselves is not guaranteed when stochastic noise is added to our velocity data, however this is likely attributable to the fact that these are the products of merging the vortices and having these vortices move through the system as the flow progresses, which we do not see in a system such as the Bickley jet. The lack of a large number of structures corresponding to a medium level value of the probability density function indicates the overall self–consistency of this robust set once more.

The transfer operator optimal vector again is shown to be robust against the presence of velocity uncertainty, as we see from the results shown in Figure 4.15. While the stochastic optimal vector field becomes fuzzy as expected, we are still able to identify approximately the same two coherent sets for this system from this field. The features of the mean optimal vector field look similar to those identifiable in the stochastic field,

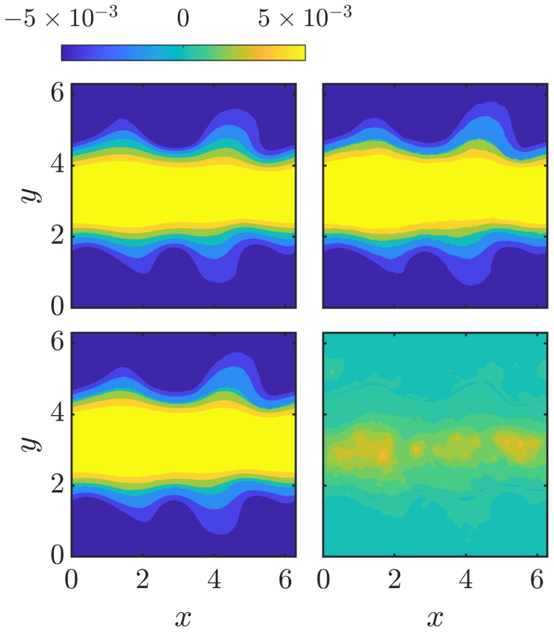

Figure 4.16: Dynamic Laplace operator second eigenvector diagnostics for the KH2 flow arranged in the same fashion as seen in Figure 4.10.

though the smoothness of the field is comparable to that of the deterministic result. The standard deviation is relatively low and can be observed more along the zero contour lines of the optimal vector field, with larger values also scattered throughout the central flow channel.

The second eigenvector field of the dynamic Laplace operator also manages to remain self–consistent under the influence of stochastic noise, as we can see from the results shown in Figure 4.16. There is almost no discernible impact on the eigenvector field side from some minor fuzziness observable within some of the contour lines. This owes to the fact that the Laplacian matrix operator diffuses and smoothens out variations which may exist in the transfer operator matrix.

We again observe robustness in the FCM clustering membership probability scalar fields from the results seen in Figure 4.17. We consider $K = 4$ clusters, and observe only minor changes to the membership probability fields for each of the clusters produced under the influence of stochastic noise. The mean irons out the variations observable within each scalar field, and the standard deviation generated is low and primarily concentrated along the boundaries of these clusters.

Finally we consider the CSC method, which has again shown to be robust under the influence of velocity uncertainty as demonstrated in Figure 4.18. The stochastic noise causes the CSC scalar field to become more jagged and lose its form in places; however the general shape of the field and the values recorded remain more or less consistent in

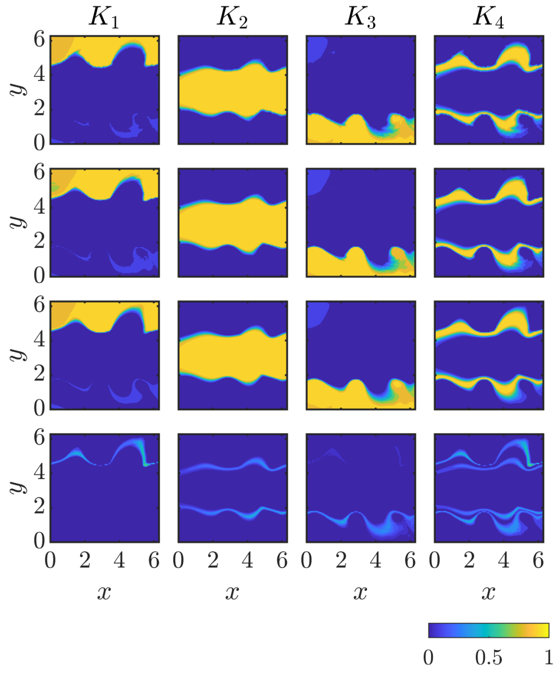

Figure 4.17: The deterministic membership probability fields for four FCM clusters relevant to the KH2 flow (first row); one realisation of each of these fields with noise applied to the velocity (second row) and the mean (third row) and twice the standard deviation (fourth row) of 100 stochastic realisations of these probabilities.

comparison to the deterministic results. The mean field appears to be a diffused, smooth version of the deterministic field while the standard deviation is overall relatively low and more concentrated within the vortex layers; particularly within the shear layer which separates the vortices from the flow channel.

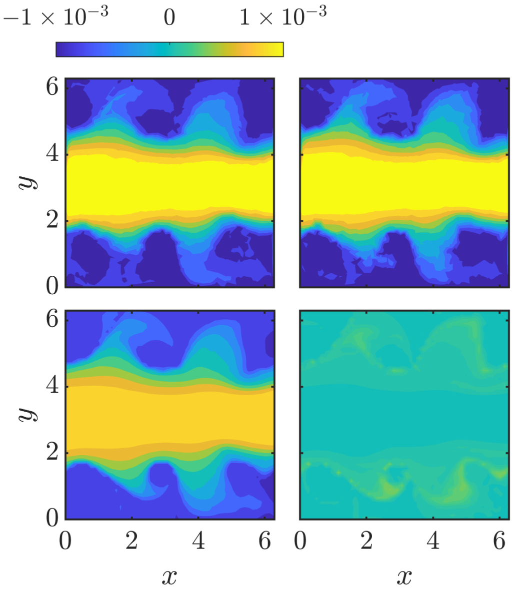

Figure 4.18: CSC diagnostics for the KH2 flow arranged in the same fashion as seen in Figure 4.10.

4.3 Gulf Stream

Finally, we consider our oceanographic velocity data set which represents the Gulf Stream flow, defined over the time interval ranging from t_0 = midnight, January 15, 2015 to $t_0 + T$ = midnight, April 15, 2015 (UTC) and using the same computational parameters for each LCS detection method as detailed earlier in Section 3.6. We again wish to keep the velocity uncertainty scale v_r consistent, and given the dimensions of this data set this parameter now takes a value of 1.27×10^{-2} degrees latitude/longitude per day. The velocity data is defined over a uniform initial grid with mesh size $h = 1/8$ degrees, therefore giving us a value of $\epsilon \approx 4 \times 10^{-2}$ degrees per square root day which we insert into (4.1) and then solve the equation to produce stochastic flow maps for our LCS analysis.

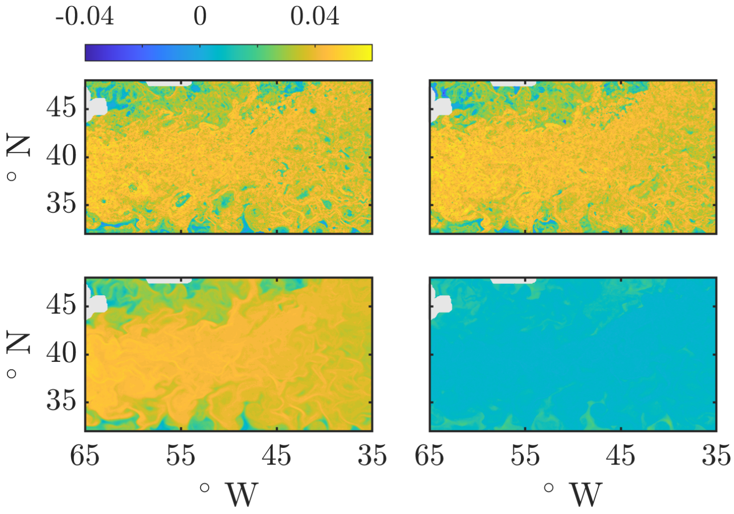

Figure 4.19: The deterministic FTLE of the Gulf Stream flow (top-left), one realisation of this quantity with noise applied to the velocity (top-right) and the mean (bottom-left) and standard deviation (bottom-right) of 100 realisations of the stochastic FTLE.

We begin once more with the FTLE, which has remained more or less robust as observable by the results shown in Figure 4.19. This time, there is no need to scale the standard deviation or variation fields in order to emphasise the structures present within these fields, so we only plot 1× the standard deviation instead of twice the standard deviation as was done with the previous two flow systems. We once again observe wobbling in the maximal ridges of the stochastic FTLE field. Because these ridges more or less cover the whole domain, and because the flow time of 90 days is considerably large, the wobbling of these ridges is considerably amplified, causing the small Gulf Stream eddies corresponding to low or negative FTLE values to vanish from the field altogether. Upon taking the mean of 100 of these realisations, we obtain a smooth field which has lost a lot of the detail observable in the deterministic field, and the standard deviation is relatively

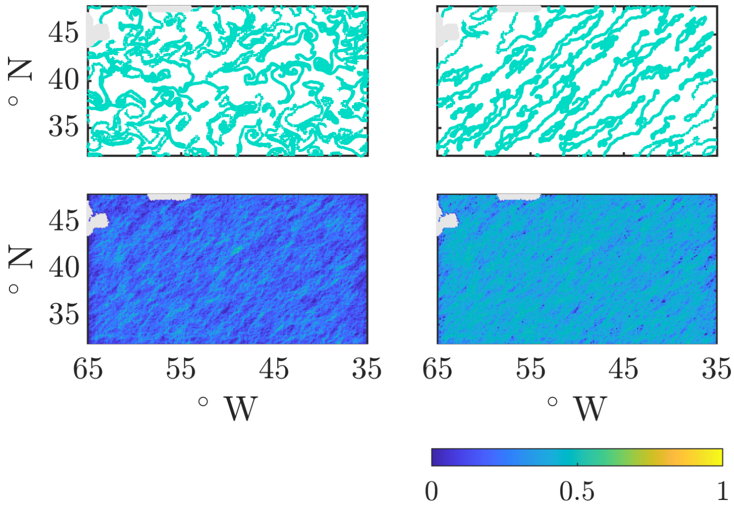

Figure 4.20: Deterministic hyperbolic VLCSs of the Gulf Stream flow (top–left), one realisation of these VLCSs with noise applied to the velocity (top–right) and the likelihood (bottom–left) and the variation in this likelihood (bottom–right) of a VLCS passing through one of 241 × 159 = 38319 bins computed from 100 stochastic realisations of these VLCSs.

low though more concentrated more along the Northern and Southern edges of the domain and within and around the Gulf of St Laurence at the top left hand corner of the domain.

Once again, the VLCS method has proven to be considerably susceptible to the presence of velocity uncertainty within our flow system, from the results in Figure 4.20. The VLCSs produced deform considerably once again, and all seem to become jagged diagonal lines which run across the area of the domain. We obtain these results most likely because over the course of our 90 day flow, many flow particles end up leaving the domain or washing up on land. We produce a likelihood field by dividing up the flow domain into 241 × 159 = 38319 bins, and we notice a large number of light blue or green coloured streaks forming similar shapes as observable in the stochastic realisations of this method. The variation in this likelihood is consistently large throughout the whole domain, indicating once again that the VLCSs are highly susceptible to uncertainty, and hence the reliability of the deterministic structures produced is in doubt.

We next consider the LAVD and observe structural robustness in the results produced for this quantity from the results shown in Figure 4.21. When we apply noise as with the previous two flow systems considered, we detect the same vortices and coherent structures with strong Lagrangian shear; however the values of the LAVD quantity have diminished in comparison to the deterministic results. The mean field looks like a diffused, smooth version of the stochastic results, not the deterministic result as the maximal LAVD values obtained in the deterministic field were not obtained in the mean LAVD field. The

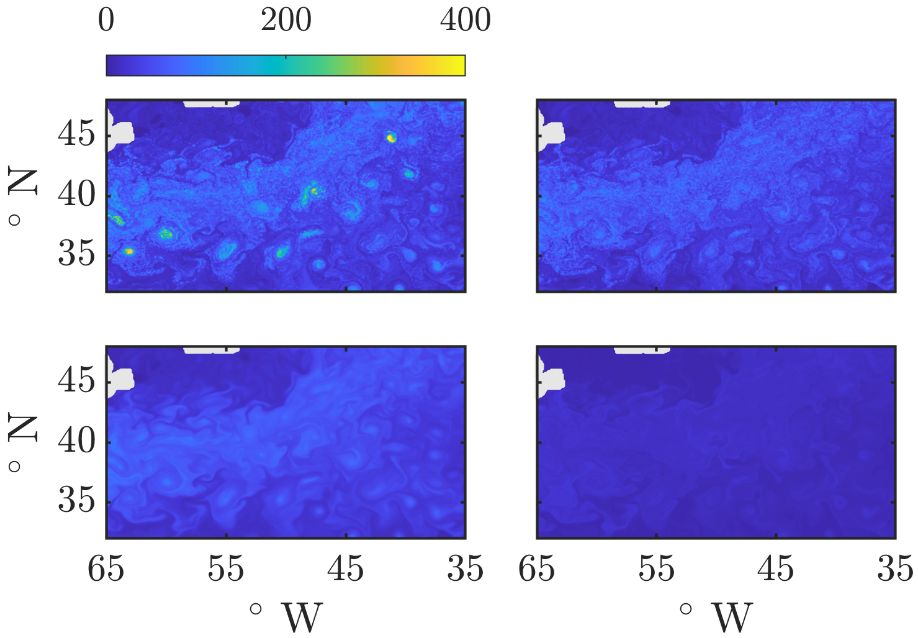

Figure 4.21: LAVD diagnostics for the Gulf Stream flow arranged in the same fashion as seen in Figure 4.19.

standard deviation is relatively small but concentrated more within the vortices and the Gulf Stream (i.e. where the LAVD is larger). Over the 90 days of flow, the Gulf Stream vortices move through the system and in some cases may break apart altogether, which is why we see such fluctuations in the LAVD values within the vortices, which was also observed in the KH2 flow but not in the Bickley Jet.

The S^2 once again appears to have been compromised to some degree by the addition of noise, as we can see from Figure 4.22. The stochastic realisation of the Scaled S^2 field produces more or less the same structures; however we observe more extremal values once again attributable to the large gradients and rapid spikes we normally observe for this quantity. We obtain once again a smoothened out version of the stochastic version of this field after taking the mean, and the standard deviation is again noticeably large where S^2 is large, particularly in the Western half of the domain due to large fluctuations in the quantity and the changing positions of the large spikes.

Once again, however, the robust set produced using S^2 remains self–consistent as we observe in Figure 4.23. We produce another probability density function from 100 stochastic versions of the robust set, this time considering the lengthscale $L = 6$ degrees latitude/longitude. The largest values are observable within the Gulf of St Laurence, along with flame shaped structures which lie along the Southern edge of the domain. This "coherence" we observe is likely attributed to the "washing up" of particles along the Southern boundary of the domain over the 90 day period of flow. Throughout the

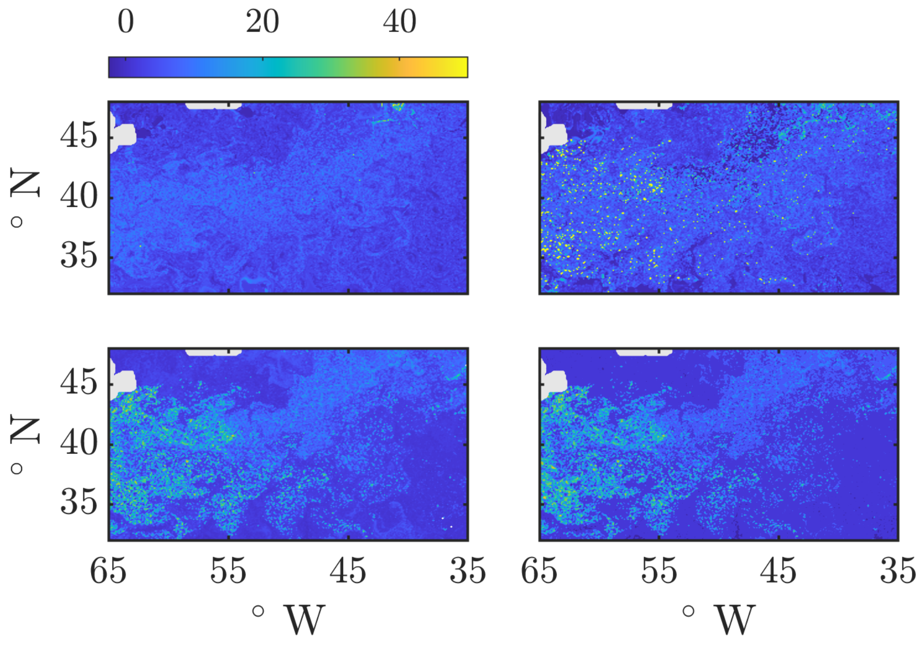

Figure 4.22: Scaled S^2 diagnostics for the Gulf Stream flow arranged in the same fashion as seen in Figure 4.19.

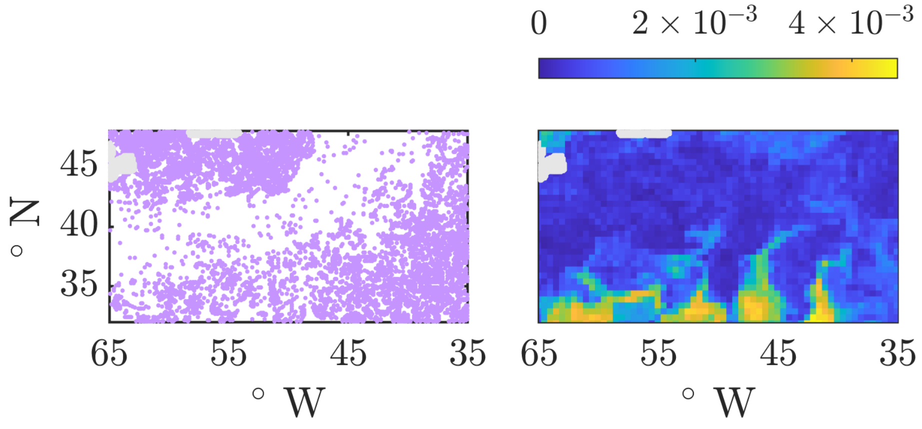

Figure 4.23: The robust set of lengthscale $L = 6$ degrees for the Gulf Stream flow (left) and the probability density function obtained from the stochastic realisations of this set (right).

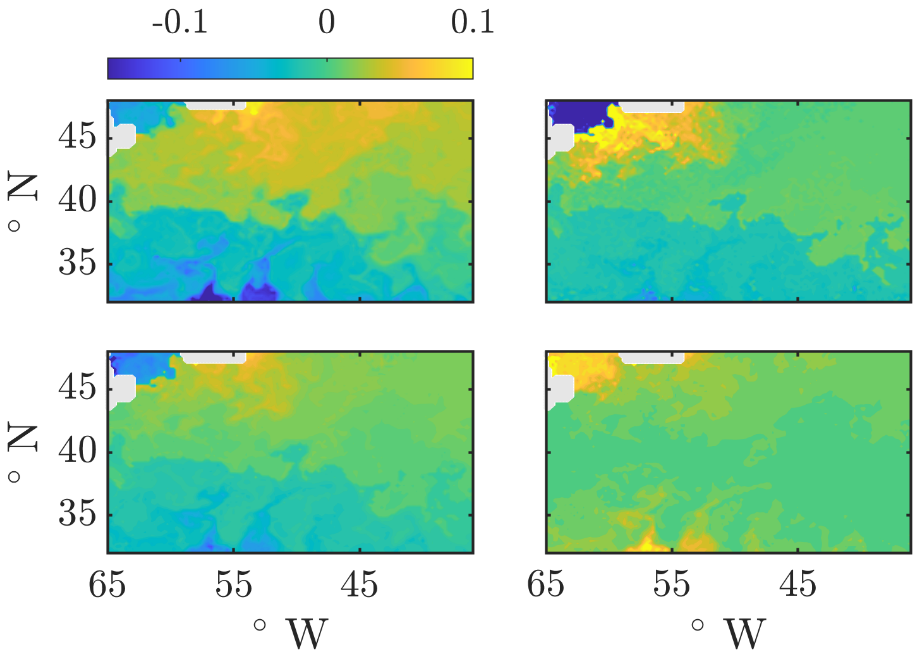

Figure 4.24: Transfer operator optimal vector diagnostics for the Gulf Stream flow arranged in the same fashion as seen in Figure 4.19.

Gulf Stream itself, the probability density function is low, which is consistent with the results achieved from the deterministic robust set and indicates that over 90 days of flow, the Gulf Stream moves and evolves considerably to the extent that it cannot be considered robust or coherent in the sense of predictability of eventual particle locations.

Next, we consider the transfer operator optimal vector, which is also reasonably robust to this velocity uncertainty as indicated by Figure 4.24. Consistent with the study undertaken by Froyland et al. (2019), we consider the third right singular vector of the transfer operator to produce the optimal vector rather than the second singular vector. When stochastic noise is added to the velocity, we obtain a scalar field that somewhat looks similar to the deterministic result, except that we see more extremal values in the transfer operator vector coefficients around land and the Gulf of St Laurence, while extremal negative values along the Southern boundary of the domain iron out and decrease in value. The mean field looks like a smoothened version of the deterministic field, bar some differences in the values of the transfer operator entries, particularly those along the Northern and Southern edges of the domain. The standard deviation is mostly low, especially around the Gulf Stream and the zero contour lines of the field, which is important in considering the reliability of the method in producing two coherent sets for the domain. Large values of the standard deviation are concentrated more within the Gulf of St Laurence, closer to land, and along the Northern and Southern edges of the domain.

The third eigenvector of the dynamic Laplace operator is also quite self–consistent

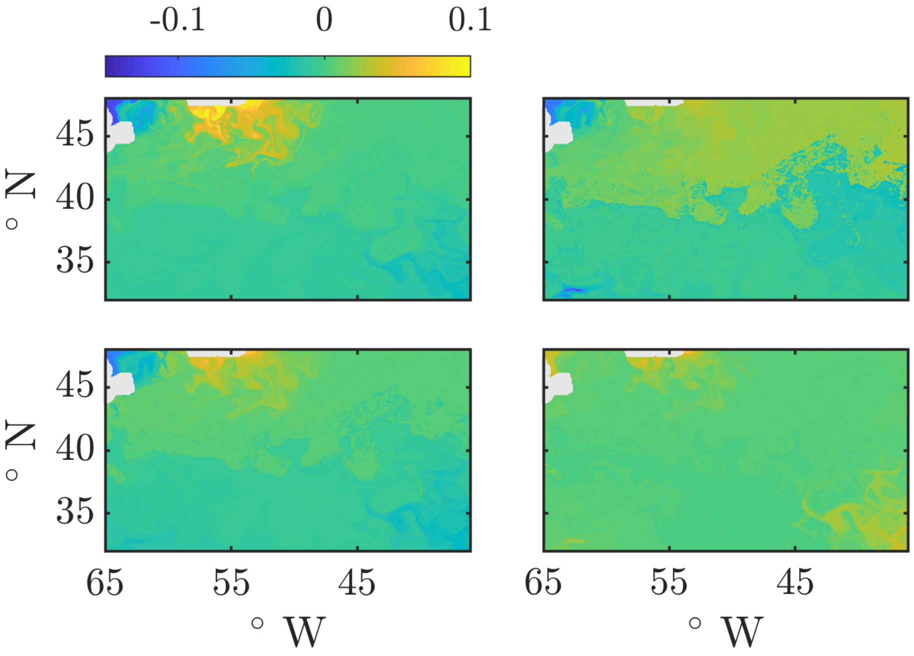

Figure 4.25: Dynamic Laplace operator third eigenvector diagnostics for the Gulf Stream flow arranged in the same fashion as seen in Figure 4.19.

as we can observe from Figure 4.25. Again, we consider the third eigenvector instead of the second, in line with the work undertaken by Froyland et al. (2019). Adding noise does not change the eigenvector field a great deal, except for the reduction in extremal values of the field and considerable fuzziness along the zero contours within the Eastern half of the domain (the contours within the Western half of the domain remain more or less consistent). The mean field looks like the deterministic field bar some differences, and the standard deviation is low and concentrated more close to land and in the South Eastern corner of the domain. Like with the transfer operator, the standard deviation is low around the zero contours of the field, which is promising when assessing the reliability of this method.

The FCM Clustering method again shows to be robust as we see from Figure 4.26. We consider the $K = 5$ case assessed previously in section 3.6, and upon adding noise to our velocity data, we observe minimal changes in the formation of the membership probability scalar fields for each of these clusters. The mean fields once again look like smoothened versions of the deterministic scalar fields, and the standard deviation is relatively low and concentrated around the borders of each cluster. While the results appear to be robust, it must be stressed that the results of this method only divide our flow domain into five roughly equal portions, and do not give a greater picture of the finer flow patterns.

Finally, we consider the coherent structure colouring method, which again shows self–consistency amidst velocity noise as demonstrated in Figure 4.27. Upon adding noise

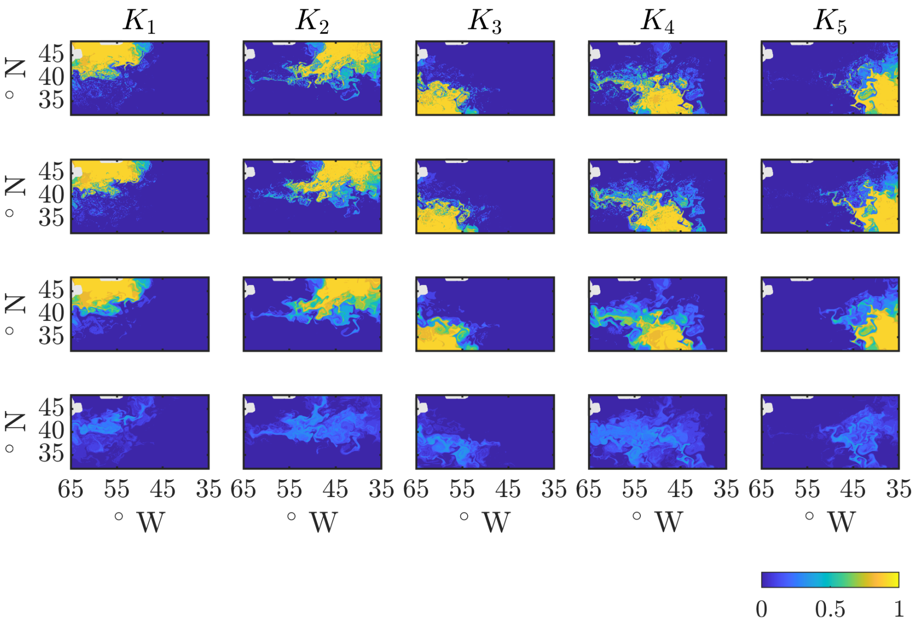

Figure 4.26: The deterministic membership probability fields for five FCM clusters relevant to the Gulf Stream flow (first row); one realisation of each of these fields with noise applied to the velocity (second row) and the mean (third row) and standard deviation (fourth row) of 100 stochastic realisations of these probabilities.

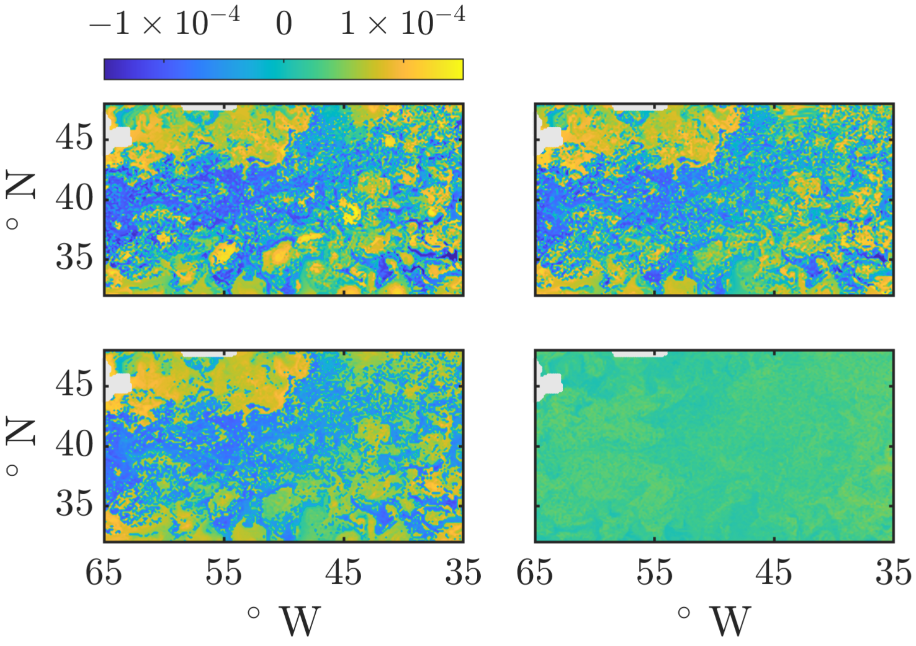

Figure 4.27: CSC diagnostics for the Gulf Stream flow arranged in the same fashion as seen in Figure 4.19.

to this system, we obtain a similar picture of the Gulf Stream and the various coherent objects (such as the eddies) observable within the flow, with the only noticeable change being the extremal positive and negative CSC coefficients softening in comparison to the deterministic results. The mean field looks like a smooth version of the stochastic version of the CSC field, and differs from the deterministic field in the maximal and minimal values of the CSC coefficients. The standard deviation is relatively low and looks to be spread more or less uniformly across the whole flow domain, bar some small exceptions.

Chapter 5

Sensitivity of Lagrangian Coherent Structures to Spatial Velocity Resolution

In this Chapter, we examine the functionality of each of our LCS detection methods against changes in the spatial resolution of a velocity data set, that is, how much velocity data is available and how much interpolation of this data will be required to "fill in the gaps" in order to produce Lagrangian flow trajectories. When considering a flow system where the velocity is analytically defined through continuous functions, obtaining the velocity necessary to calculate Lagrangian flow trajectories will be relatively straightforward. Rarely, if ever, will *observational* flow data be defined by analytic velocity equations, and instead will be defined by instances of velocity data available over a finite grid of points and at a finite number of time steps. While every effort can and often is taken to reduce the grid mesh size of the spatial velocity data to obtain the maximum velocity resolution possible, there will always be gaps in the velocity data of these observational data sets which need to be filled by interpolation (estimation) of the velocity between these data points. If the spacing between data grid points is too large and therefore the velocity resolution is not strong enough, the construction of flow trajectories will be inaccurate, thereby producing unreliable LCSs and an unreliable visualisation of the flow system overall.

We therefore seek to examine what sort of an impact a reduction in the spatial resolution of velocity has on the reliability of our LCS detection methods. In similar fashion to our stochastic noise analysis in the previous Chapter, we consider three of the flow systems defined in Chapter 3, namely the Bickley Jet, the KH1 flow (instead of KH2), and the Gulf Stream. We perform LCS analysis on each of these systems using analytically defined velocity (where possible), or numerically defined velocity of the largest resolution available, and compare these "reference case" results to the same analysis undertaken at lower resolutions. We then qualitatively compare the structures produced for each resolution level to see how similar the results are in each case, and to try and estimate "how low is too low" when it comes to the available resolution levels of a data set. In doing this, we will be able to identify a rule of thumb for the number of velocity data points which must be contained within a coherent structure in order for it to be identified by a detection method.

We will also consider more quantitative methods of assessing these results, specifically relating to the errors between certain LCS quantities computed with the best spatial resolution available and with reduced resolution. We achieve this by producing histograms of the absolute errors between an LCS quantity computed using the maximal resolution and at a reduced resolution level over all initial grid points $\mathbf{x}_0 \in \Omega$; assessing the features of this histogram (such as the distribution of the errors, the mode of this distribution, the spread of the errors etc.), and comparing this with similar results for the same LCS quantity generated with other resolution levels. We will also feature log–log plots of the medians of these absolute errors against the velocity data grid mesh size h to see if there is any noteworthy correlation between these quantities. We will perform this quantitative analysis on the FTLE, LAVD, S^2 and entries of the transfer operator optimal vector, the dynamic Laplace operator eigenvector and the CSC coefficient vector. We will not perform this analysis on the VLCS method as this method produces material curves that are difficult to quantitatively compare in this manner, or the membership probabilities of the FCM clustering method because, as will be demonstrated later on in this Chapter, the sets of coherent clusters produced using this method do not remain consistent with each change in the spatial resolution of the velocity.

5.1 Bickley Jet

We begin with the Bickley Jet flow, considered earlier in Sections 3.3 and 4.1. Since this flow is analytically defined, one could argue that it makes no sense to perform resolution analysis on this system as the velocity of the flow can be obtained at any point in Ω and at any time step using the system of equations (3.3). However, we can artificially transform this system into a numerically defined velocity system by sampling a grid of points within Ω at various time steps, recording the velocity at these points using the provided equations, and producing an interpolant for the velocity based on this data like we would for a numerically defined velocity system like the simulated Kelvin–Helmholtz systems or the Gulf Stream flow. We produce an interpolant for the Bickley Jet system over time steps ranging from $t_0 = 0$ to $t_0 + T = 40$ days separated by a step of 0.01 days, and over spatial grids with points separated by grid mesh sizes h of 0.083 Mm, 0.167 Mm, 0.333 Mm, 0.667 Mm, 0.952 Mm, 1.333 Mm and 2.222 Mm. To execute each of the LCS detection methods, we again use the same computational parameters for each method as defined in Section 3.3. For the sake of brevity, we will not consider all of these resolution levels when qualitatively illustrating the resultant LCSs in the following Figures, rather

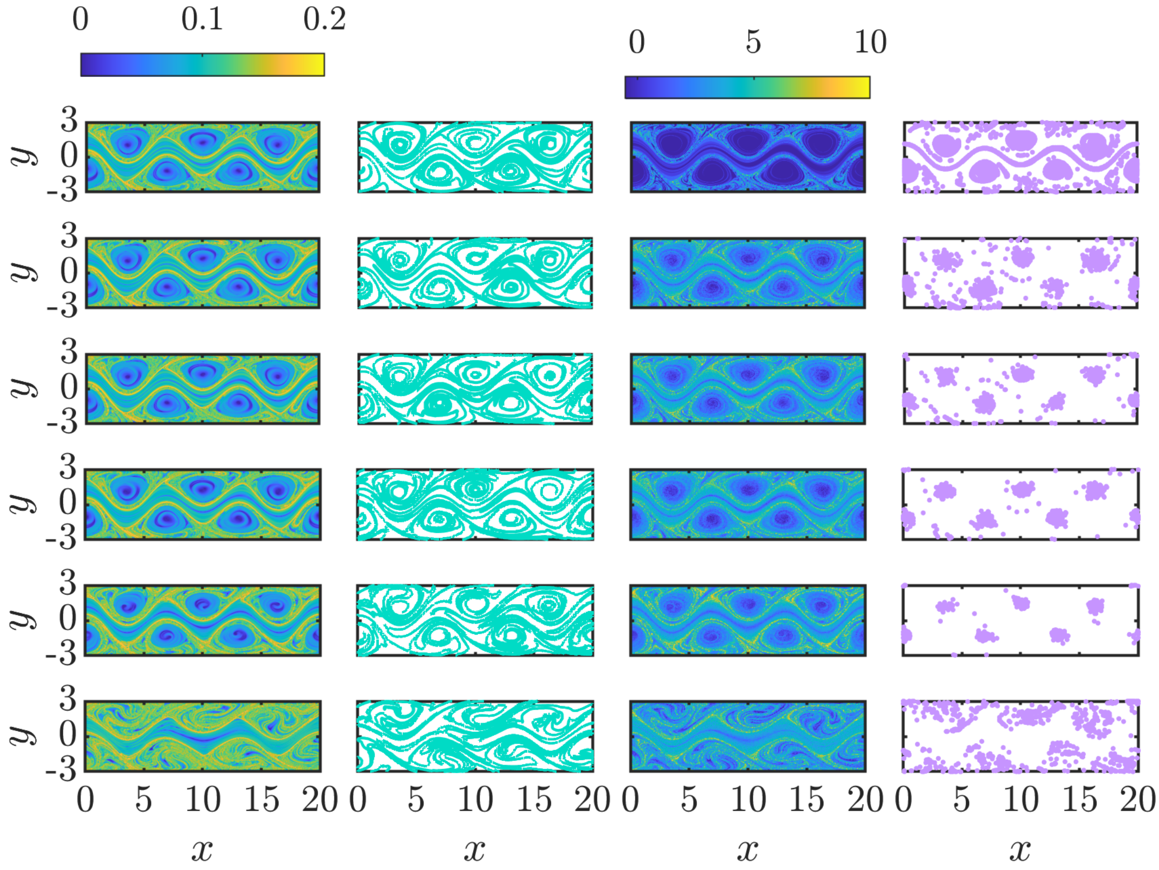

Figure 5.1: The FTLE (Column 1), VLCSs (Column 2), Scaled S^2 (Column 3) and $L = 0.3$ Mm robust sets (Column 4) for the Bickley Jet flow computed from analytically defined velocity (Row 1), and numerically defined velocity over a grid with spacing $h \approx 0.167$ Mm (Row 2), 0.667 Mm (Row 3), 0.952 Mm (Row 4), 1.33 Mm (Row 5) and 2.22 Mm (Row 6).

we will only show instances of resolution levels where significant changes in the LCSs produced from each method are observable.

In Figure 5.1, we display LCSs produced from the FTLE, VLCS (VLCS–B only, as done in the previous Chapter) and S^2 (including the robust set with $L = 0.3$ Mm) methods produced from varying spatial resolution levels. A common thread observable amongst all of these results is that while some of the structures identified vary in shape as the resolution reduces, key robust structures (such as the Bickley jet vortices and the zonal jet) appear to maintain their shape until the resolution of the velocity becomes unworkably small. In addition, the deterioration of the results as the resolution becomes worse appears to be consistent across all of the methods, due to consistencies in the inaccuracies of the Lagrangian flow trajectories generated. In the case of the FTLE method, most of the maximal ridges and key structures maintain their shape (bar some minor differences), until we reach the $h \approx 0.952$ Mm case where we notice changes in the interior detail of the vortices, though they still maintain their shape along with most of the other structures. The same can be observed for the 1.33 Mm case, however when the grid spacing for the velocity data increases to 2.22 Mm, the vortices have smeared out and are no longer identifiable, while the zonal jet can somehow still be identified. A possible theory explaining this is that the grid spacing of the velocity data is too large in comparison to the size of the coherent structures in the flow, and without a sufficient number of data points covering the structure (say, four or five points), the structure will no longer be identifiable. A similar set of observations are identifiable for the VLCS method, which appears to be much less sensitive to changes in resolution that it has been to changes in velocity uncertainty. The composition of VLCSs changes with each resolution level, giving a slightly different picture of hyperbolic flow barriers with each iteration produced. This can have a lot to do with the algorithms of this method, and how the candidate points for eigenvector field integration subsequently change with each new iteration and computation of Lagrangian flow trajectories. The S^2 method, however, appears to be more acutely sensitive to changes in data resolution, as the ridges are considerably more fuzzy and disoriented; resulting in the robust set identifying fewer and more incoherent flow objects with each decrease in resolution. This is an expected result due to the large gradients and rapid spikes in S^2, which cannot be resolved as the spatial resolution decreases. This is likely to have been further amplified by the fact that particle flow between the Bickley jet vortices (as mentioned earlier) is highly chaotic and unpredictable, with this unpredictability only amplified by the decrease in fidelity of the velocity data.

Similarly, the results shown in Figure 5.2 demonstrate that the LAVD, transfer operator, dynamic Laplace operator and CSC methods behave in similar general fashion to the methods shown in Figure 5.1. Consistently across all of these methods, a reduction in the spatial resolution of the velocity does not immediately cause the resultant LCSs identified to change too drastically, however in the $h \approx 1.33$ Mm case, the Bickley jet vortices begin to fade from view, completely disappearing in the $h \approx 2.22$ Mm case with similarly meaningless flow features as in Figure 5.1. In similar fashion to the velocity uncertainty analysis from the previous Chapter, the LAVD typically decreases in value as the resolution decreases even though the same structures are identifiable, with the values increasing in the $h \approx 2.22$ Mm case as more meaningless flow structures are produced. In all of the vector based methods (transfer operator, dynamic Laplace operator and CSC), it can be observed that a decrease in resolution facilitates the production of discontinuities within the scalar fields. This is likely a direct result of the lack of velocity resolution available and the associated difficulty in assigning an eigenvector/singular vector coefficient to each initial box/point within the grid. Despite this, all of these methods are able to identify the zonal jet, the transfer operator and dynamic Laplace operator methods can identify

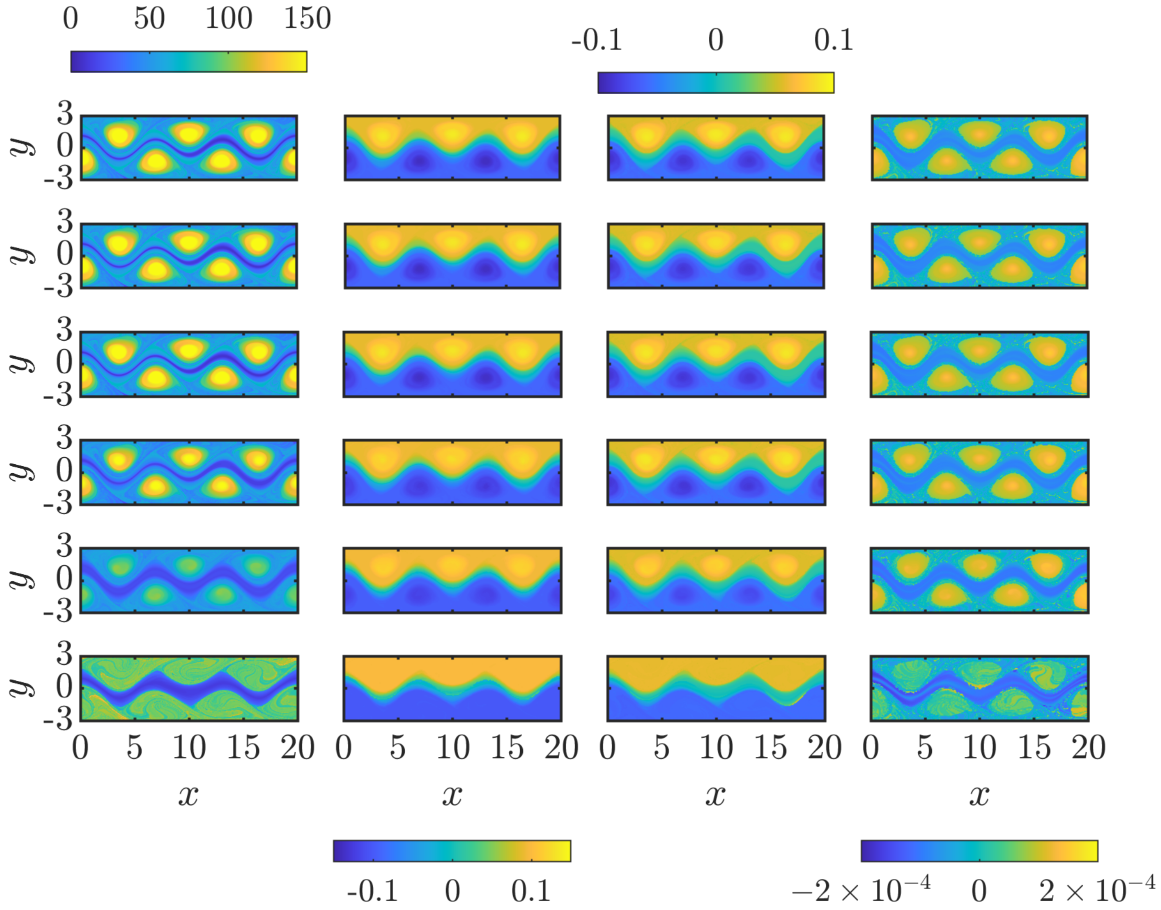

Figure 5.2: The LAVD (Column 1), second $\mathbf{P}$ singular vector (Column 2), second $\boldsymbol{\Delta}^D$ eigenvector (Column 3) and CSC scalar field (Column 4) for the Bickley Jet flow computed from analytically defined velocity (Row 1), and numerically defined velocity over a grid with spacing $h \approx 0.167$ Mm (Row 2), 0.667 Mm (Row 3), 0.952 Mm (Row 4), 1.33 Mm (Row 5) and 2.22 Mm (Row 6).

the upper and lower vortex layers as coherent sets (even though the vortices can no longer be identified when $h \approx 2.22$ Mm); and the CSC method can still outline the shapes of the Bickley jet vortices even in the worst resolution case, despite increased discontinuities.

The results for the FCM clustering method are generally quite inconsistent across all of the resolution levels considered, as can be seen from the membership probability fields shown in Figure 5.3. We again consider $K = 3$ clusters, and while for most of the resolution levels the three clusters are consistent with the results obtained from the analytically defined velocity, in the $h \approx 0.667$ Mm and $h \approx 0.952$ Mm cases three completely different clusters have been isolated by this method. The zonal jet cluster has become more thin, a second cluster consists of the two disjoint full vortex layers, and a third cluster has been

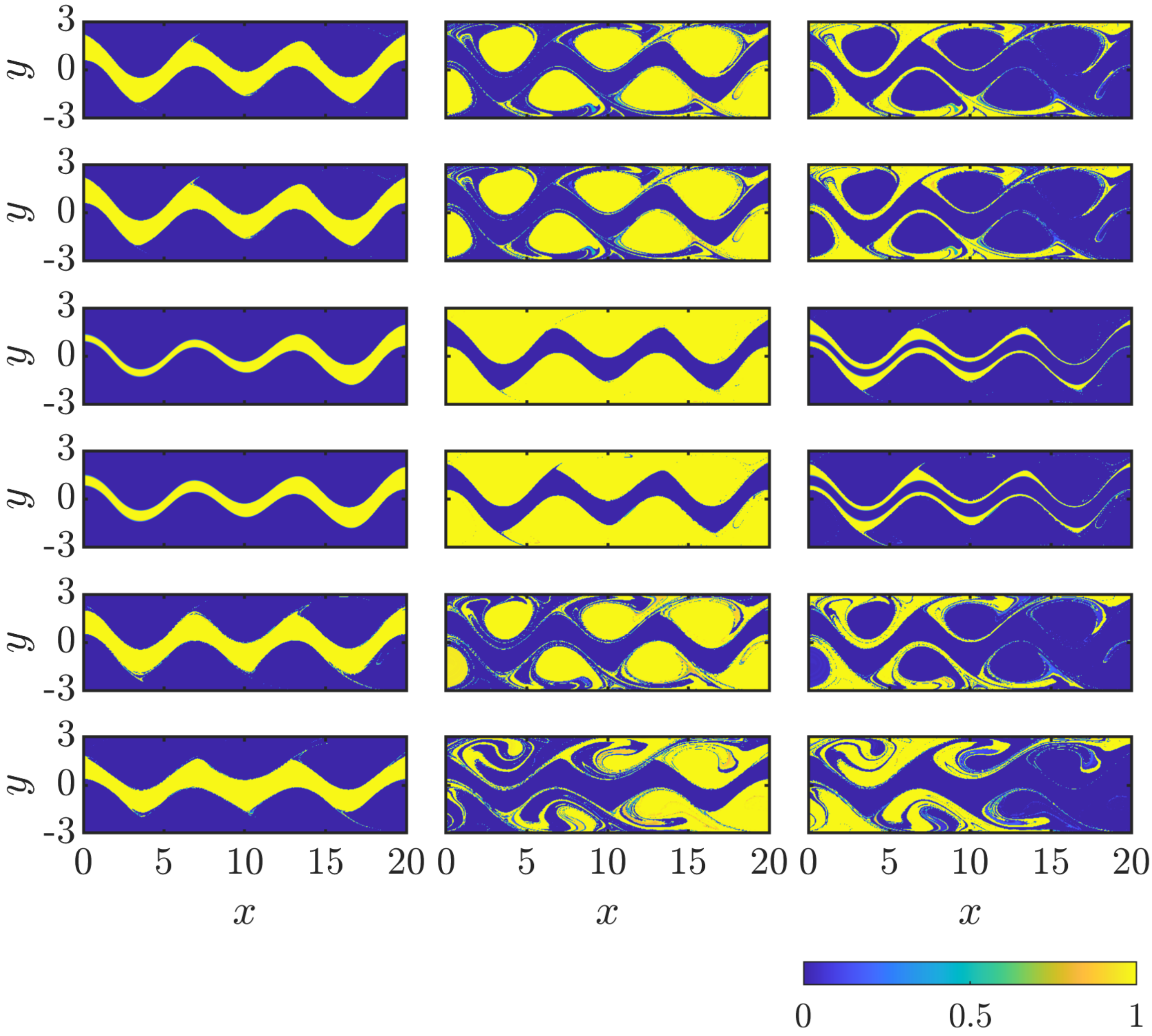

Figure 5.3: Scalar fields of the membership probabilities for three FCM clusters for the Bickley Jet flow computed from analytically defined velocity (Row 1), and numerically defined velocity over a grid with spacing $h \approx 0.167$ Mm (Row 2), 0.667 Mm (Row 3), 0.952 Mm (Row 4), 1.33 Mm (Row 5) and 2.22 Mm (Row 6). Each column represents one of the FCM clusters identified for each resolution level, with FCM clusters that feature similar coherent patterns grouped in the same column.

formed from two jet shaped objects above and below the central jet cluster detected. The most likely explanation for the inconsistencies is that due to the variation in flow trajectories produced, the algorithm will group these together differently. Strangely enough, we return to the original group of three clusters when $h \approx 1.33$ Mm, and when $h \approx 2.22$ Mm we obtain the zonal jet and two muddled clusters with flow objects which look similar to those obtained from the methods discussed earlier. The inconsistent FCM Clusters obtained can possibly be rectified by altering the cluster sharpness parameter m, or by changing the number of initial points seeded within the domain. We will not perform quantitative analysis on this LCS detection method as it makes little sense to given the inconsistencies observed.

Now that we have qualitatively assessed how the results of these LCS detection methods change, we turn our attention to more quantitative resolution analysis focusing on the distribution of the absolute error in relevant LCS fields as the resolution changes. We produce probability density functions of the absolute errors, scaling the LCS quantities by their largest values computed under analytically defined velocity and removing outliers from the sample. For ease of illustration, the histograms of these probability density functions have been split into two Figures - Figure 5.4 shows the functions for the FTLE, LAVD and S^2 (without scaling the latter quantity first); while Figure 5.5 shows similar functions for the entries obtained from the transfer operator second singular vector, the dynamic Laplace operator second eigenvector, and the CSC coefficient vector. A key feature across all of these density functions is that for larger resolution levels, the distribution of these functions for the errors is concentrated around zero before decaying in shape, with the maximal height of the functions decreasing. From these results, one can speculate that the distribution of these errors will take the form of a Dirac delta distribution at infinite velocity resolution, and will approach a uniform distribution when the resolution is unreasonably coarse. In terms of the spread of these errors, the largest spread is observable within the FTLE and LAVD errors where the largest errors recorded are of order 10^{-1} when the velocity data grid spacing is especially large. The transfer operator and dynamic Laplace operator coefficients can have error up to the order of 10^{-3}, with the mode of the errors shifting closer towards the median of these values with decreasing resolution. The CSC coefficients also have an error spread reaching an order of 10^{-3}, though at the coarsest resolution the error density function looks much more like a uniform distribution. The S^2 errors appear to be extremely low even at a coarse resolution (order 10^{-8}), however we have removed all outliers from the error distribution of which most of these can be extremely large (the largest of order $\sim 10^{20}$), again due to large spikes and gradients in the S^2 field. Where these spikes are low and more smooth, the values appear to remain much more consistent, however this large disparity in the range of the error prevents the absolute errors of this quantity from assuming a uniform distribution, even for extremely coarse data resolution.

As an added quantitative experiment for these absolute errors, we have recorded the

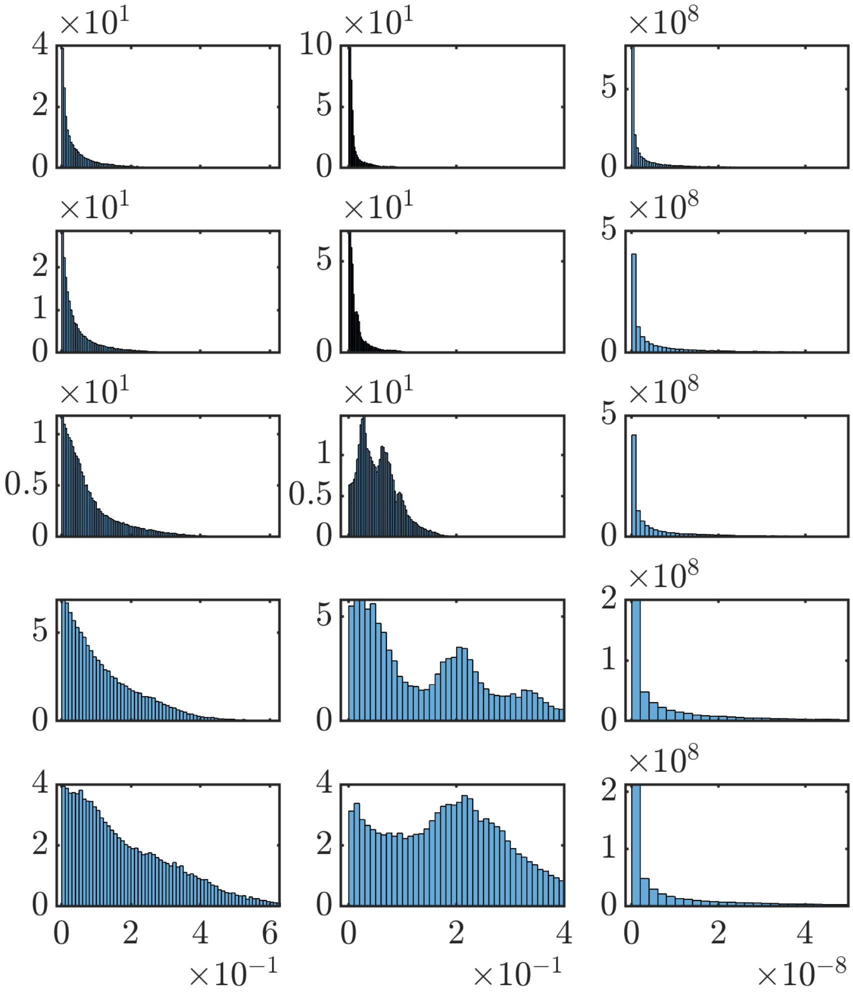

Figure 5.4: Probability density function histograms for the absolute differences between various Bickley Jet LCS quantities computed from analytically defined velocity and numerically defined velocity over a grid with spacing $h \approx 0.167$ Mm (Row 1), 0.667 Mm (Row 2), 0.952 Mm (Row 3), 1.33 Mm (Row 4) and 2.22 Mm (Row 5). The quantities in question are the FTLE (Column 1), LAVD (Column 2) and S^2 (Column 3).

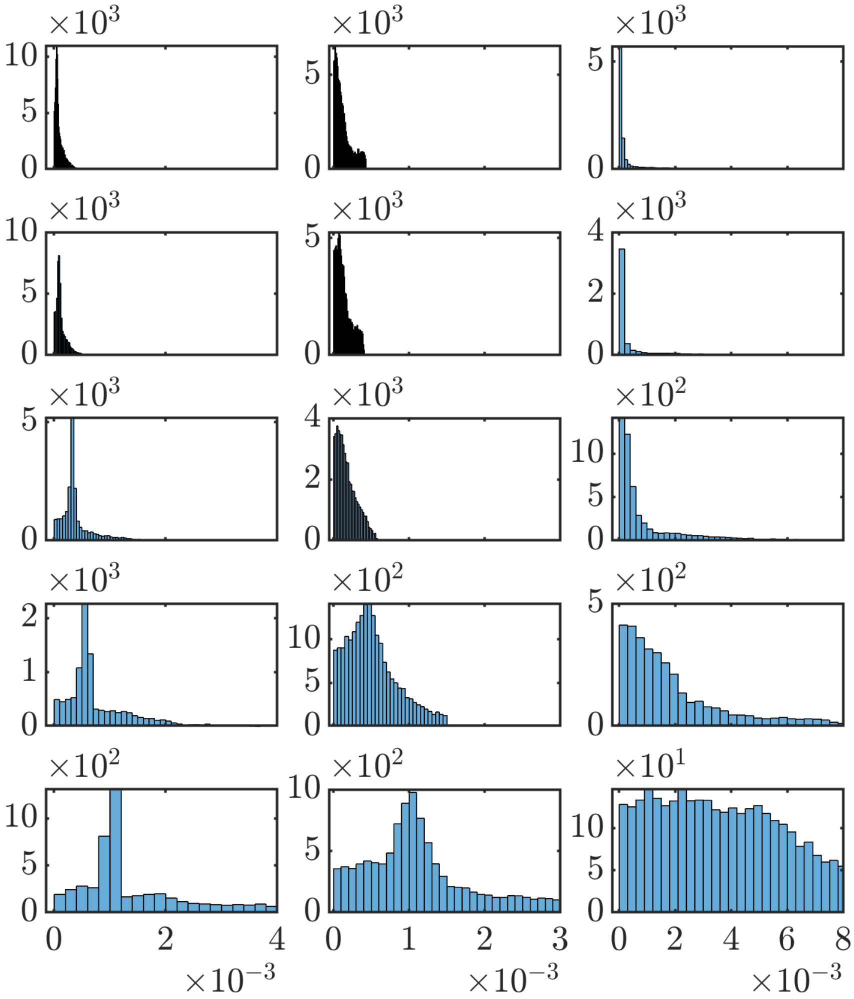

Figure 5.5: Probability density function histograms for the absolute differences between various Bickley Jet LCS quantities computed from analytically defined velocity and numerically defined velocity over a grid with spacing $h \approx 0.167$ Mm (Row 1), 0.667 Mm (Row 2), 0.952 Mm (Row 3), 1.33 Mm (Row 4) and 2.22 Mm (Row 5). The quantities in question are the second singular vector of $\mathbf{P}$ (Column 1), the second eigenvector of $\mathbf{\Delta}^D$ (Column 2) and the CSC (Column 3).

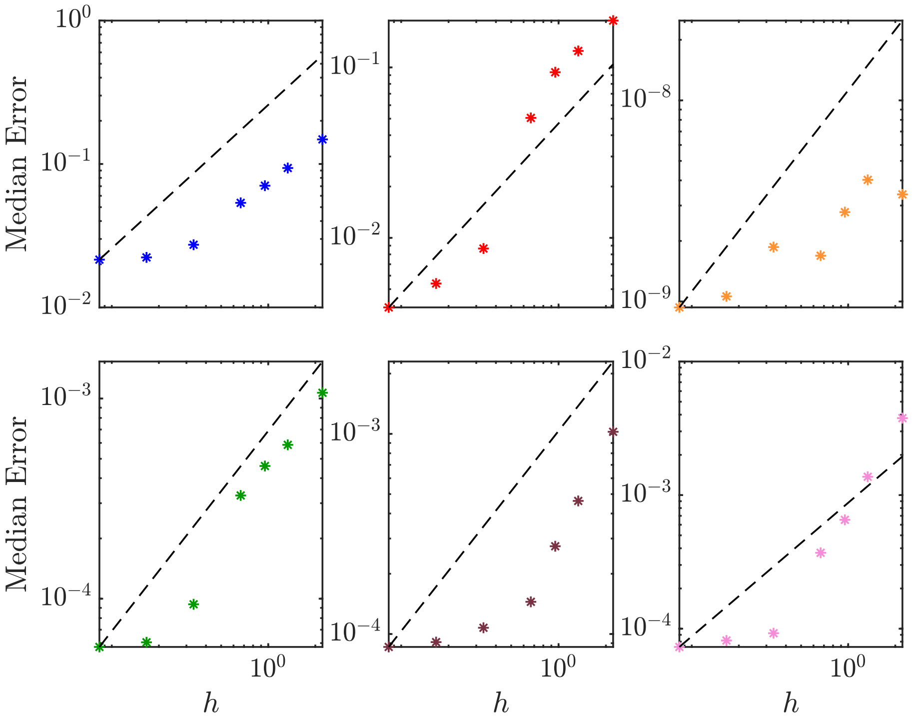

Figure 5.6: Log–log plots of the median absolute differences between various LCS quantities for the Bickley Jet flow computed from analytically defined velocity and numerically defined velocity defined over data grids of different levels of spacing. The quantities in question are (from left to right): the FTLE, the LAVD, S^2 (row 1), the transfer operator second singular vector, the dynamic Laplace operator second eigenvector and the CSC (row 2). For comparison against this data, a dashed line of slope 1 (in log–log space) has been included within each plot.

median values of all of the reduced resolution errors listed above and produced log–log plots of these median values against the grid spacing of the velocity data. We have experimented with other average measures such as the mean and root mean square error, but have settled with the median as the plots produced for these other measures have similar shape and form to those produced using the median with no greatly distinguishable features making these more viable of a measure for the average of these error values. In addition, the median is a more accommodating average measure for these errors as the absolute error data for each of the relevant LCS quantities is skewed and contains a considerable number of outliers. The log–log plots themselves can be seen in Figure 5.6, with the addition of a line of slope 1 in log–log space as guide. Some positive correlation between the velocity grid spacing and the median absolute error of these LCS quantities may appear to exist, with the slopes of these relationships lying somewhere between 0.5 and 2. This relationship can be further verified by computing the median absolute errors for more values of the data grid mesh size h, which was not done in this case because the computation time required to complete these extra calculations was rather extensive.

5.2 Kelvin–Helmholtz Version 1 (KH1)

Next, we consider version 1 of the numerically simulated Kelvin–Helmholtz flow (the KH1 flow), which we have previously looked at in detail in Section (3.4). While we have considered the KH2 flow in Section 4.2 for analysis into the impact of velocity uncertainty on our LCS detection methods, we instead turn our attention to the KH1 flow this time as it contains more robust vortices which can change shape or disappear as the resolution of the system changes. Unlike the previously studied Bickley Jet system, this flow is not analytically defined, hence the "reference case" will be derived from the best data resolution available, which was derived from the 1024×1024 grid of initial points within the domain Ω for this system, corresponding to a grid mesh size $h = 2\pi/1024 \approx 6.1 \times 10^{-3}$ m. We obtain velocity data of reduced spatial resolution by sub–sampling on our originally simulated velocity data over grids with spacing $h = 2\pi/256 \approx 0.0245$ m, $h = 2\pi/128 \approx 0.0491$ m, $h = 2\pi/64 \approx 0.0982$ m, $h = 2\pi/32 \approx 0.196$ m, $h = 2\pi/16 \approx 0.393$ m and $h = 2\pi/8 \approx 0.785$ m. Again, for the sake of brevity we will not display the results for all of these resolution levels in the below Figures. All other computational parameters pertinent to each method remain consistent with those selected in Section (3.4).

We begin with Figure 5.7, where we show the results for the FTLE, VLCS and S^2 (with robust set, $L = 0.2$ m) methods given gradual reductions in the resolution of our velocity data. Like with similar results displayed for the Bickley jet system (see Figure 5.1), the FTLE and S^2 methods are initially resistant to changes in the resolution, but the Kelvin–Helmholtz vortices begin to warp when $h = 2\pi/16$ m, before vanishing completely when $h = 2\pi/8$ m. The maximal ridges of the scalar fields of these quantities display similar behaviour. The central flow channel hardly changes shape for any of these realisations, owing to the size of the structure and that there are enough velocity data points within the structure to enable it to be detected sufficiently. Unlike for the Bickley Jet, the S^2 field and robust set are now much more robust against these changes in resolution, as we do not observe any fuzziness in the scalar field or grainy robust flow objects within the robust set plot. The foliations of VLCS curves appear to noticeably change with each new resolution level, as observed within the Bickley jet system earlier on. This is likely due to a variation in the candidate points for eigenvector field integration selected due to the change in flow trajectory behaviour brought on by the change in data resolution. By the time the spatial resolution decreases to $h = 2\pi/8$ m, the VLCS method ends up producing a foliation of almost straight lines, with most of these horizontal particularly where the vortex layers should exist.

Next, we turn our attention to the LAVD, transfer operator, dynamic Laplace operator and CSC results for this system; displayed in Figure 5.8. The patterns we observe here are

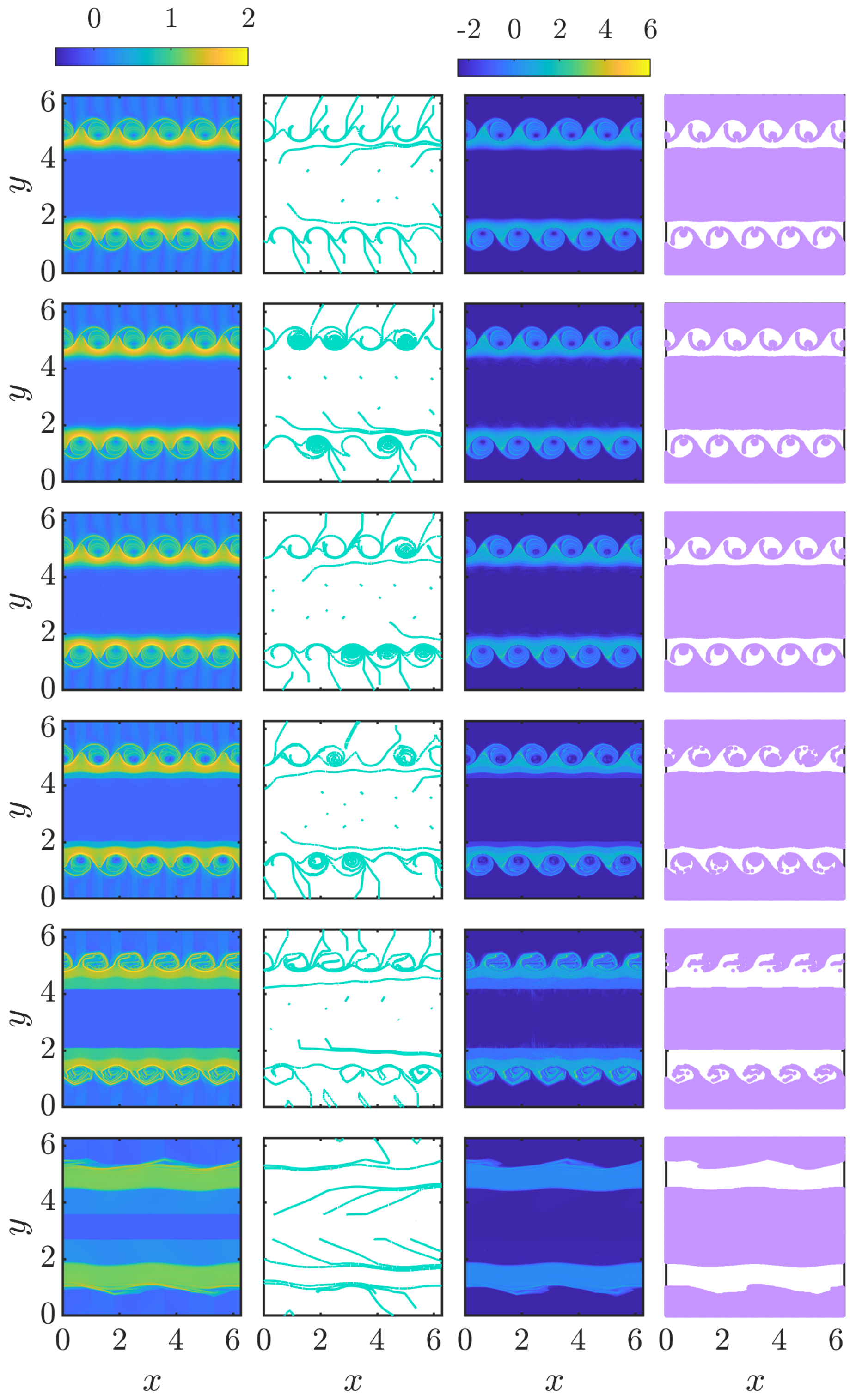

Figure 5.7: The FTLE (Column 1), VLCSs (Column 2), Scaled S^2 (Column 3) and $L = 0.2$ m robust sets (Column 4) for the KH1 flow computed from numerically defined velocity of the best available resolution (grid spacing $h = 2\pi/1024$ m) (Row 1), and of reduced resolution levels over grids with spacing $h = 2\pi/128$ m (Row 2), $2\pi/64$ m (Row 3), $2\pi/32$ m (Row 4), $2\pi/16$ m (Row 5) and $2\pi/8$ m (Row 6).

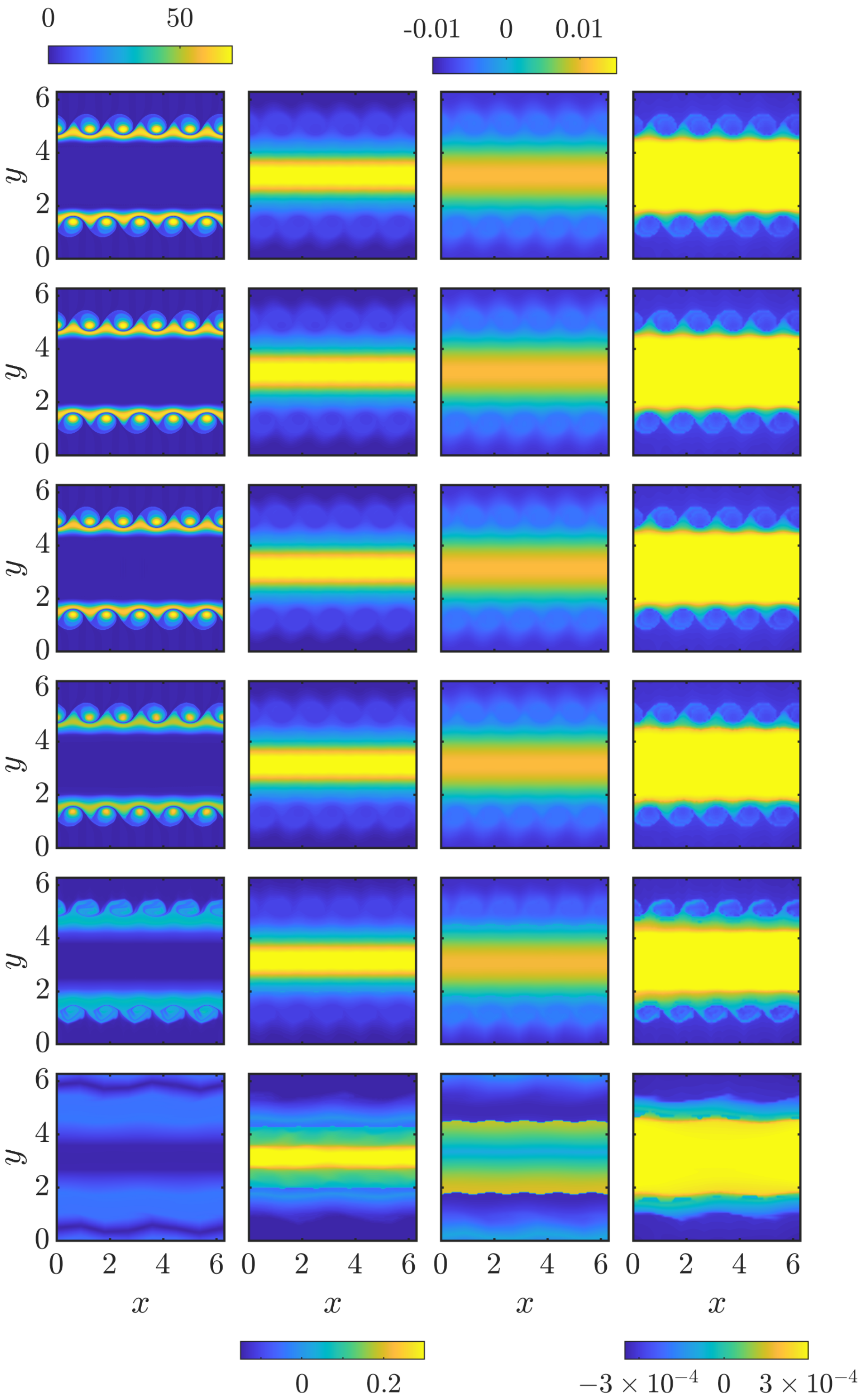

Figure 5.8: The LAVD (Column 1), third $\mathbf{P}$ singular vector (Column 2), third $\boldsymbol{\Delta}^D$ eigenvector (Column 3) and CSC scalar field (Column 4) for the KH1 flow computed from numerically defined velocity of the best available resolution (grid spacing $h = 2\pi/1024$ m) (Row 1), and of reduced resolution levels over grids with spacing $h = 2\pi/128$ m (Row 2), $2\pi/64$ m (Row 3), $2\pi/32$ m (Row 4), $2\pi/16$ m (Row 5) and $2\pi/8$ m (Row 6).

comparable to those observed for the Bickley jet flow (see 5.2), with the majority of flow structures beginning to fade only at extremely low levels of resolution (except for the large central flow channel which is always observable to some degree). The vortices and shear layers detected by the LAVD method retain their shape until $h = 2\pi/16$ m, after which point the vortices warp and combine with the shear layers to produce two incoherent flow channels above and below the central flow channel when $h = 2\pi/8$ m. Similar behaviour is observable for the other three methods and, like with the Bickley jet flow system, as the spatial resolution decreases more discontinuities appear in the scalar fields of the $\mathbf{P}$ third singular vector, the $\mathbf{\Delta}^D$ third eigenvector and the CSC; as resolving these coherent structures becomes harder. These discontinuities are more greatly observable in the transfer operator and dynamic Laplace operator methods, though considerably less so in the case of the CSC method.

In Figure 5.9, we consider the results for the FCM clustering method, with $K = 4$ clusters being sought. While the results for this flow are overall more consistent with the other LCS detection methods than the results for the Bickley jet flow (see Figure 5.3), we still obtain one unusual collection of clusters from one realisation of this method. When $h = 2\pi/64$ m, we obtain consistent clusters containing the top and bottom vortex layers which extend to the closest vertical boundary; however rather than obtaining the full central flow channel as one cluster, only the right hand end of this channel forms one cluster while the left hand end is grouped up with the shear layers on either vertical end of the channel.

We next turn our attention to the probability density functions of the absolute errors recorded for our LCS results for instances of lower resolution against the reference cases generated from the best available resolution. Figure 5.10 shows the sample probability density function histograms for the FTLE, LAVD and raw S^2; while in Figure 5.11 we consider the transfer operator third singular vector and the dynamic Laplace operator third eigenvector; along with the CSC coefficients. While the histograms exhibit somewhat similar behaviour to the results obtained from the Bickley jet flow (see Figures 5.4 and 5.5), for most of the methods (aside from the transfer operator and dynamic Laplace operator methods), by the time the data spacing reaches a level of $2\pi/8$ m, there is still a considerable skew in the distribution in the positive direction, caused by a large amount of small error values. These small values are likely located within the central flow channel which, as seen earlier, retains more or less a consistent shape for each of these methods even when the spatial resolution is extremely coarse. The FTLE and LAVD again exhibit large spreads in their collections of error values (with said values extending to an order of 10^{-1}), while the spread extends to an order of 10^{-3} for the transfer operator, dynamic Laplace operator and CSC methods. In the S^2 case, the spread extends to an order of 10^{-7}; however we have once again excluded extreme outlying error values which were found to exist, though were not as numerous as the errors recorded from the more volatile Bickley jet system.

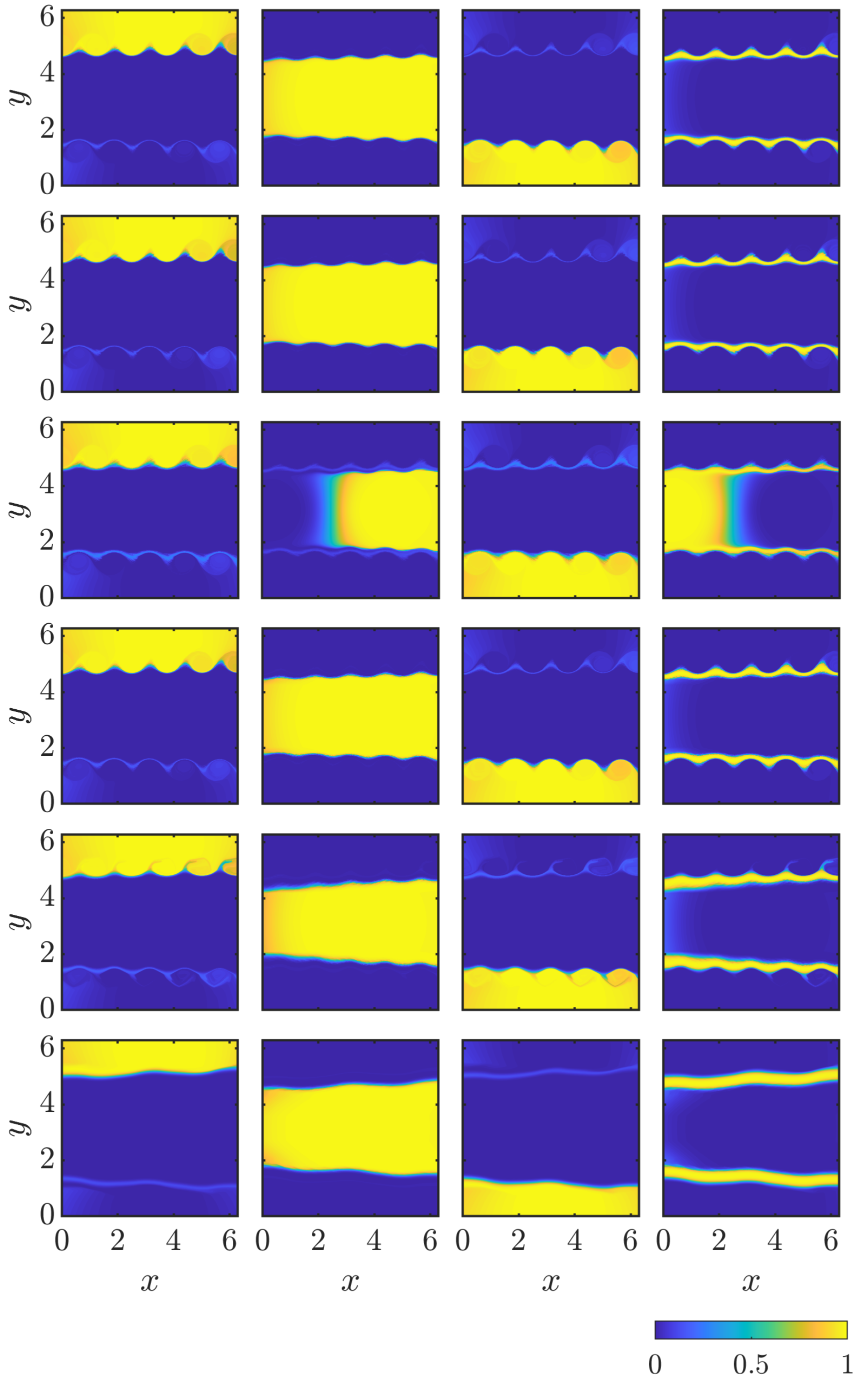

Figure 5.9: Scalar fields of the membership probabilities for four FCM clusters for the KH1 flow computed from numerically defined velocity of the best available resolution (grid spacing $h = 2\pi/1024$ m) (Row 1), and of reduced resolution levels over grids with spacing $h = 2\pi/128$ m (Row 2), $2\pi/64$ m (Row 3), $2\pi/32$ m (Row 4), $2\pi/16$ m (Row 5) and $2\pi/8$ m (Row 6). Each column represents one of the FCM clusters identified for each resolution level, with FCM clusters that feature similar coherent patterns grouped in the same column.

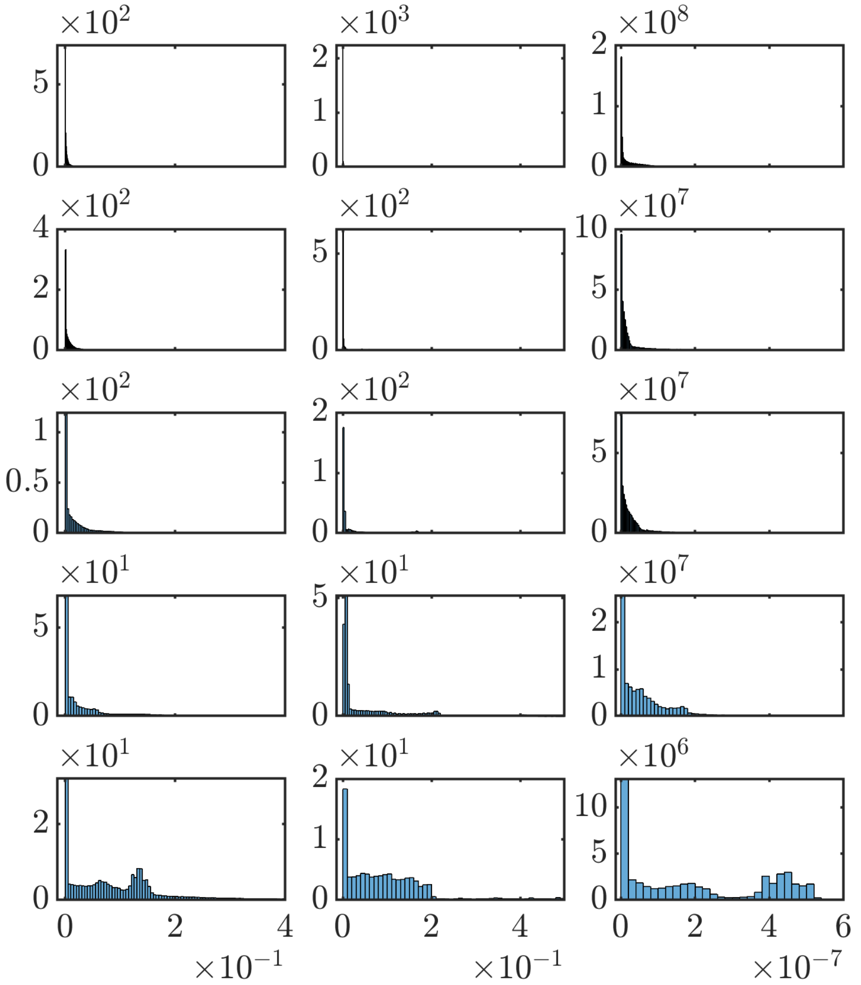

Figure 5.10: Probability density function histograms for the absolute differences between various KH1 LCS quantities computed from numerically defined velocity of the best available resolution (grid spacing $h = 2\pi/1024$ m) and of reduced resolution levels over grids with spacing $h = 2\pi/128$ m (Row 1), $2\pi/64$ m (Row 2), $2\pi/32$ m (Row 3), $2\pi/16$ m (Row 4) and $2\pi/8$ m (Row 5). The quantities in question are the FTLE (Column 1), LAVD (Column 2) and S^2 (Column 3).

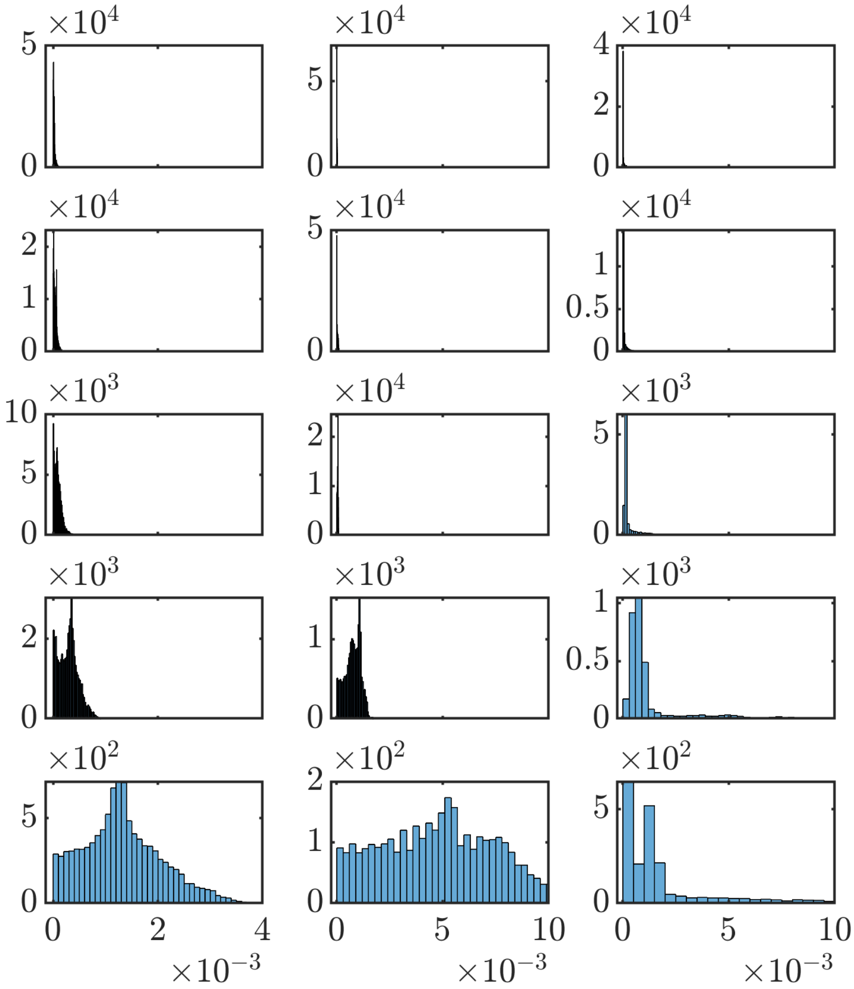

Figure 5.11: Probability density function histograms for the absolute differences between various KH1 LCS quantities computed from numerically defined velocity of the best available resolution (grid spacing $h = 2\pi/1024$ m) and of reduced resolution levels over grids with spacing $h = 2\pi/128$ m (Row 1), $2\pi/64$ m (Row 2), $2\pi/32$ m (Row 3), $2\pi/16$ m (Row 4) and $2\pi/8$ m (Row 5). The quantities in question are the third singular vector of $\mathbf{P}$ (Column 1), the third eigenvector of $\mathbf{\Delta}^D$ (Column 2) and the CSC (Column 3).

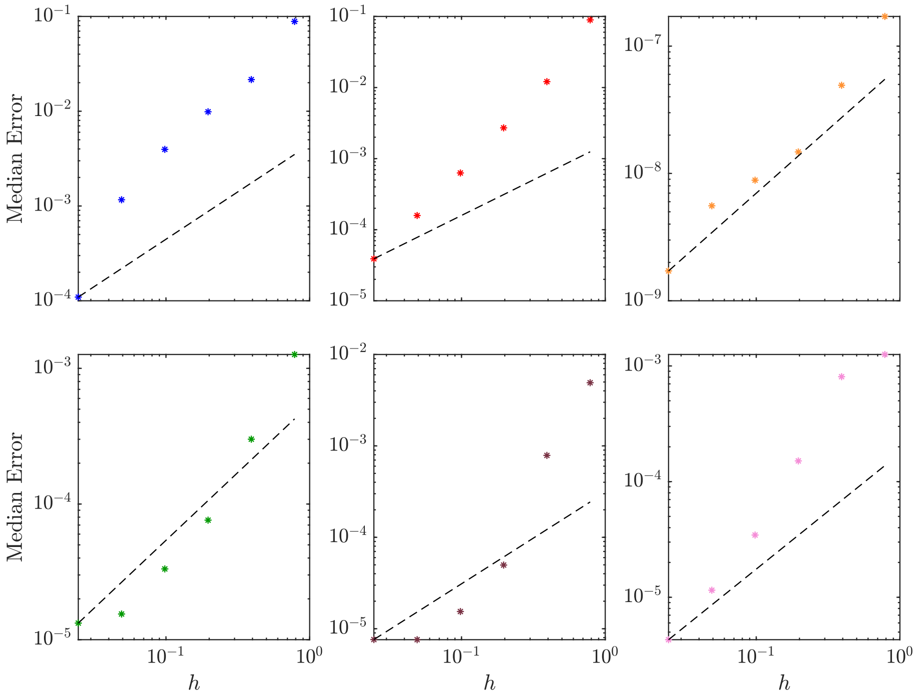

Figure 5.12: Log–log plots of the median absolute differences between various LCS quantities for the KH1 flow computed from numerically defined velocity of the best available resolution and of reduced resolution defined over different grid mesh sizes h. The quantities in question are (from left to right): the FTLE, the LAVD, S^2 (row 1), the transfer operator third singular vector, the dynamic Laplace operator third eigenvector and the CSC (row 2). For comparison against this data, a dashed line of slope 1 (in log–log space) has been included within each plot.

Finally, we pinpoint a positive (but not a linear) correlation between the median values of these absolute errors and the grid spacing h. These plots can be seen in Figure 5.12, again with a line of slope 1 included to compare the scatter plot against. Like with the results obtained for the Bickley jet flow, a positive correlation between h and the median absolute error values can be seen for each of these detection methods, however it does not look as though the slope of the correlation is 1. In the case of the FTLE, LAVD, S^2 and CSC, we may be able to identify some correlation of slope greater than 1, with the correlation in the LAVD case looking particularly strong for this system. In the case of the transfer operator and dynamic Laplace operator, the correlation appears to begin flat for smaller values of h, before a more positive correlation emerges for larger values of h (greater than 0.1).

5.3 Gulf Stream

Finally, we consider our oceanographic data set representing the Gulf Stream, which we have previously considered in Section 3.6, and in Section 4.3 for analysis related to velocity uncertainty. The best available resolution that we have for this data set corresponds to a grid spacing $h = 0.125°$ latitude/longitude, which will serve as our "reference" situation. We will reduce the resolution of this data set in similar fashion to the KH1 flow by sub–sampling the reference data. The resolution levels which we have considered correspond to grid spacing levels h of $0.167°$, $0.2°$, $0.25°$, $0.333°$, $0.5°$, $1°$, $2°$ and $3.33°$. Again, we won't display the results for all of these resolution levels below, particularly when two or more realisations are too similar with no notable or distinguishable differences. All other computational parameters considered for each of our LCS detection methods are the same as those used in Sections 3.6 and 4.3.

The plots shown in Figure 5.13 detail how the results obtained from the FTLE, VLCS and S^2 methods (including the robust set, $L = 6°$) change as the resolution is reduced in fidelity. Once again, we see that all of these methods are robust to these changes in resolution up until a certain point where the distinguishing features of the flow system change. As the grid spacing h reaches 1 degree, the vortices surrounding the Gulf Stream vanish or change shape, the Gulf Stream itself becomes less distinguishable, and patches of "robust" particle flow (characterised by low values of the FTLE and S^2) begin forming around the Northern edge of the domain, likely caused by the exodus of flow particles from the domain and the lack of refinement of the underlying velocity grid. These patches become larger when $h = 2$ degrees, but then oddly disappear when $h \approx 3.33$ degrees. Interestingly, the S^2 field becomes less fuzzy as the resolution of the velocity becomes more coarse, as a result of this velocity undersampling which irons out large variations. This was not observed for the previous two systems, with the fuzziness of this field remaining more or less consistent as the resolution changed. The VLCSs, while (consistently with other methods) incorrect for highly coarse resolution levels, still manage to remain quite

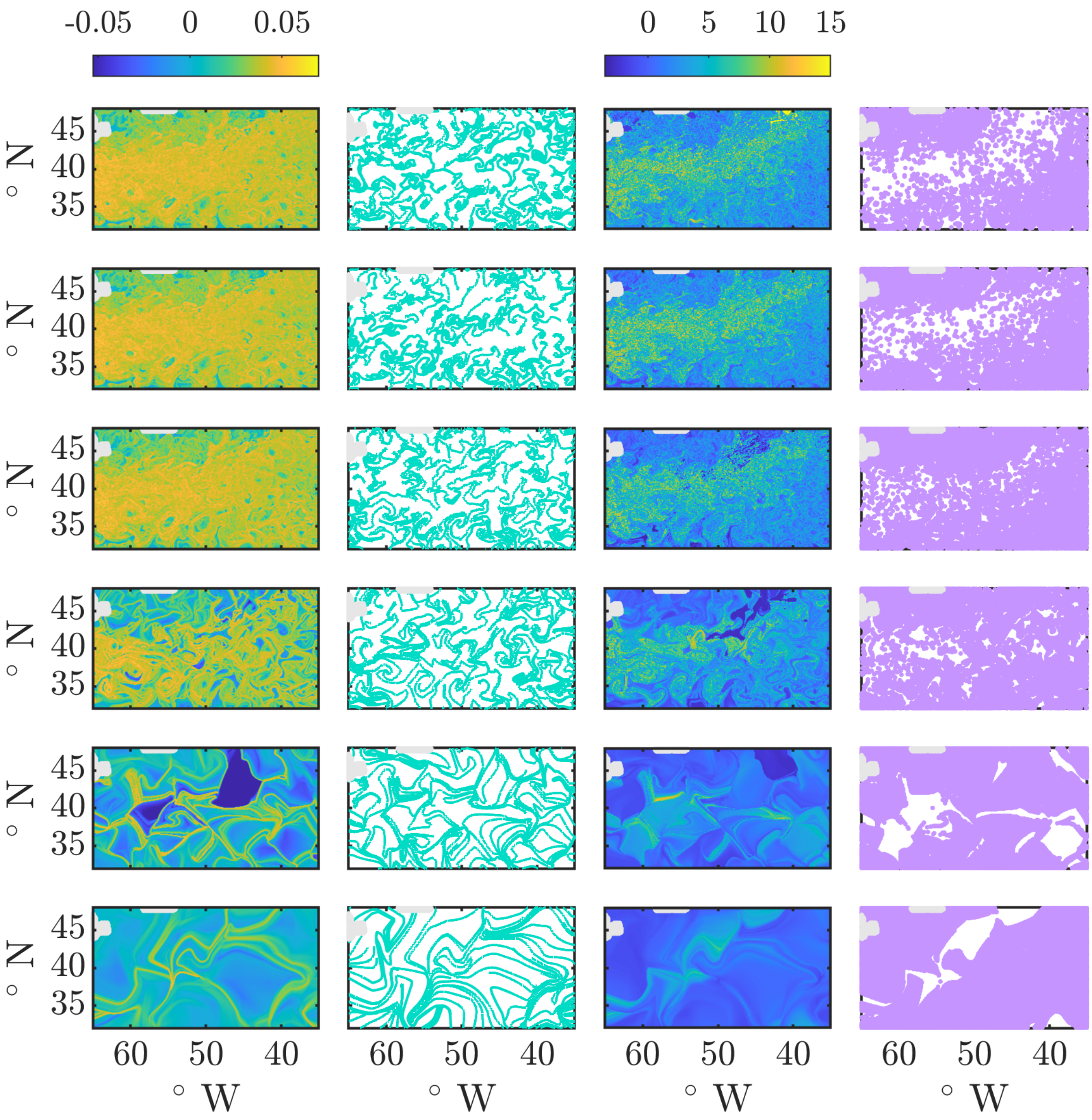

Figure 5.13: The FTLE (Column 1), VLCSs (Column 2), Scaled S^2 (Column 3) and $L = 6°$ robust sets (Column 4) for the Gulf Stream flow computed from numerically defined velocity of the best available resolution (grid spacing 0.125° latitude/longitude) (Row 1), and of reduced resolution levels over grids with spacing 0.2° (Row 2), 0.5° (Row 3), 1° (Row 4), 2° (Row 5) and 3.33° (Row 6).

smooth in these instances, in contrast to its high susceptibility to velocity uncertainty

Next, in Figure 5.14, we display the results for the LAVD, transfer operator, dynamic Laplace operator and CSC methods. In consistent fashion with the other two flows considered in this Chapter, as the resolution of the velocity decreases the peak values recorded for the LAVD method decrease. When $h = 1$ degree, we can still roughly pinpoint the coherent vortices and the shear within the Gulf Stream for this system, however after that point the coherent structures obtained begin to smear out and by the time $h \approx 3.33$ degrees, obtaining coherent vortices along with the Gulf Stream itself becomes virtually impossible. The dilapidation of the structures is consistent with the results observed in Figure 5.13. The transfer operator and dynamic Laplace operator are still able to identify the same coherent sets isolated for this system earlier when the resolution is relatively coarse, however as we have seen with the other two flow systems, discontinuities in the eigenvector/singular vector entries begin to appear by $h = 1$ degree. After that point, the coherent sets identified begin to significantly change shape, particularly in the worst resolution case ($h \approx 3.33$ degrees). This is not something that we observed in the other two systems, as both of these methods were still able to more or less identify the same coherent sets when the resolution of the data was at its worst (such as the Bickley jet vortex layers or the Kelvin–Helmholtz flow channel). The reason for this observation is that the coherent structures in this system are considerably smaller than the grid mesh size h in these instances of highly coarse resolution. The CSC method is still able to identify the wide array of coherent objects present within this flow system up until the grid spacing h increases beyond 1 degree, by which point more discontinuities begin to form within the CSC coefficients and coherent objects previously detected have now been lost. This occurs at roughly the same resolution threshold for all of our LCS methods.

In Figure 5.15, we display the results for the FCM clustering method with $K = 5$ coherent clusters sought. The five clusters remain consistent up until the same resolution level as the other methods began to produce meaningless results ($h = 1$ degree), after which they change shape and location.

Next, we consider the probability density function histograms for the absolute errors. In Figure 5.16, we show the histograms for the FTLE, LAVD and raw S^2, while in Figure 5.17 we focus on the transfer operator third singular vector, the dynamic Laplace operator third eigenvector and the CSC. In consistent manner to our observations from the previous two methods, it would appear that the distributions of each of these errors are mostly concentrated near zero, approaching more uniformity as h gets larger. The spread of the errors is considerably higher in comparison to the previous two flows, with the FTLE error spread reaching an order of 1, 10^{-1} for the LAVD, 10^{-2} for the CSC and transfer operator, and 10^{-3} for the dynamic Laplace operator. The biggest change is observable in S^2, where the error spread, even with the extreme outliers removed, sits at a value of around 7 or 8, rather than an extremely small value of order 10^{-7}. This most likely can be explained by the reduction in fuzziness as the resolution decreases, as this was not

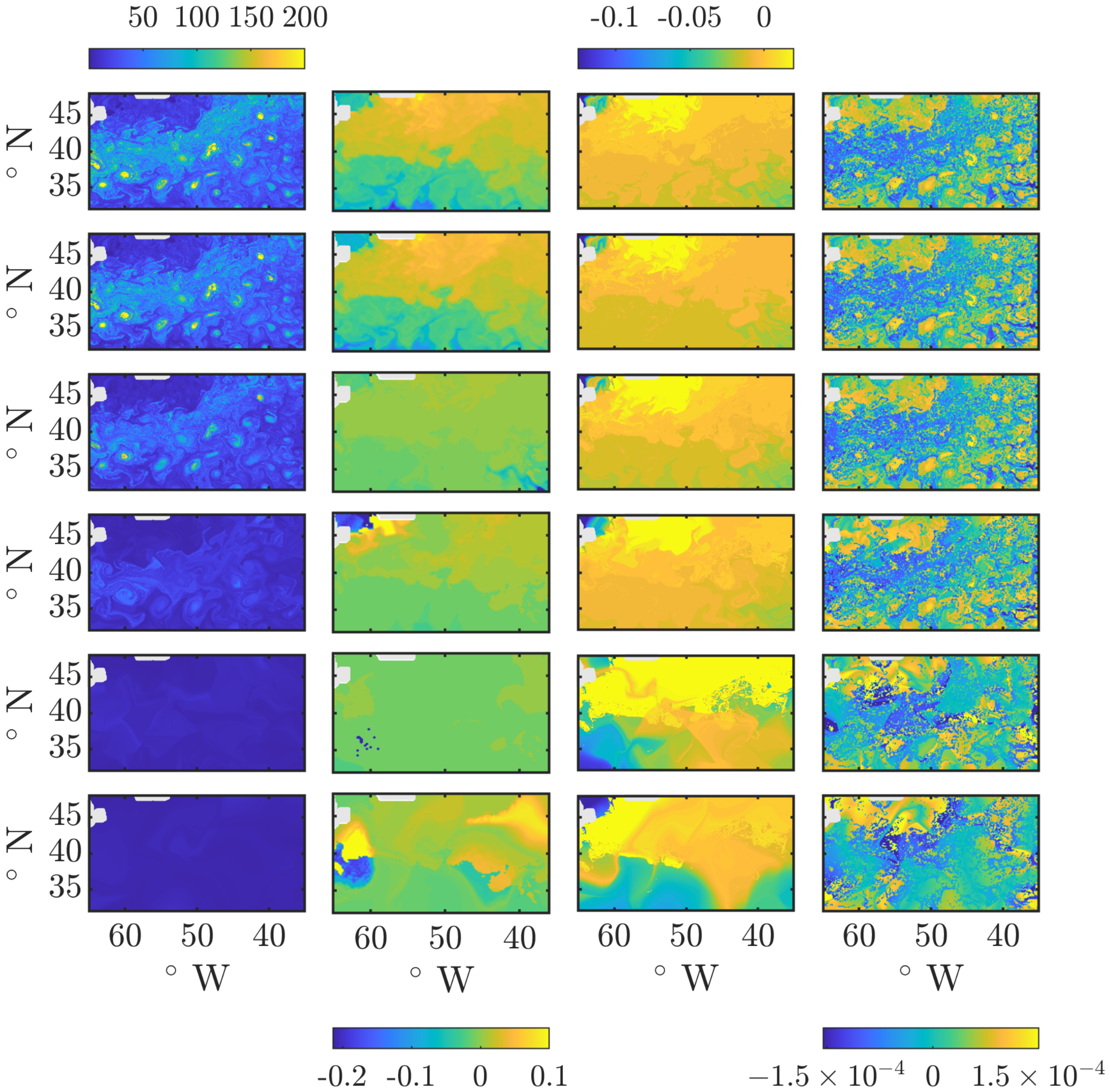

Figure 5.14: The LAVD (Column 1), third $\mathbf{P}$ singular vector (Column 2), third $\mathbf{\Delta}^D$ eigenvector (Column 3) and CSC scalar field (Column 4) for the Gulf Stream flow computed from numerically defined velocity of the best available resolution (grid spacing 0.125° latitude/longitude) (Row 1), and of reduced resolution levels over grids with spacing 0.2° (Row 2), 0.5° (Row 3), 1° (Row 4), 2° (Row 5) and 3.33° (Row 6).

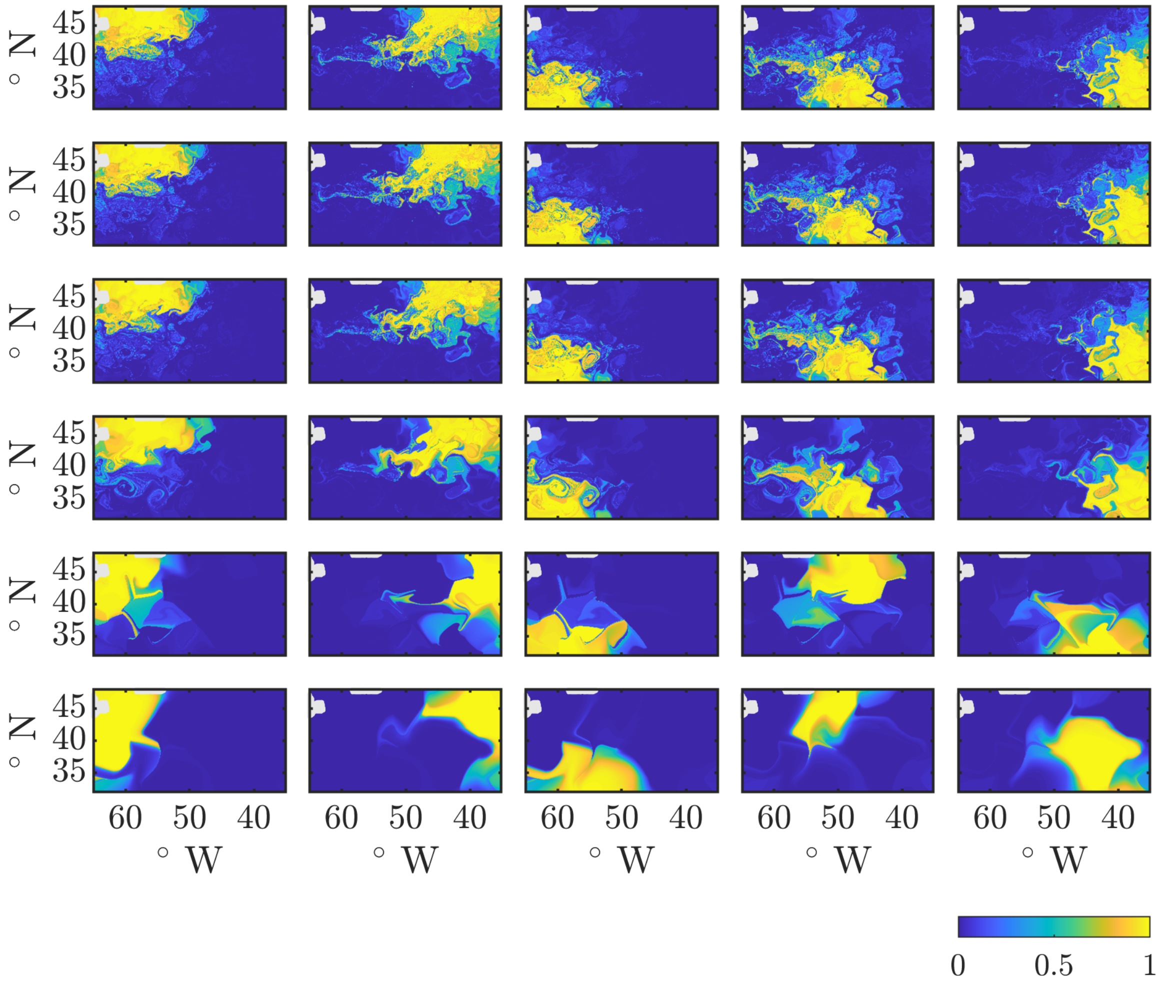

Figure 5.15: Scalar fields of the membership probabilities for five FCM clusters for the Gulf Stream flow computed from numerically defined velocity of the best available resolution (grid spacing 0.125° latitude/longitude) (Row 1), and of reduced resolution levels over grids with spacing 0.2° (Row 2), 0.5° (Row 3), 1° (Row 4), 2° (Row 5) and 3.33° (Row 6). Each column represents one of the FCM clusters identified for each resolution level, with FCM clusters that feature similar coherent patterns grouped in the same column.

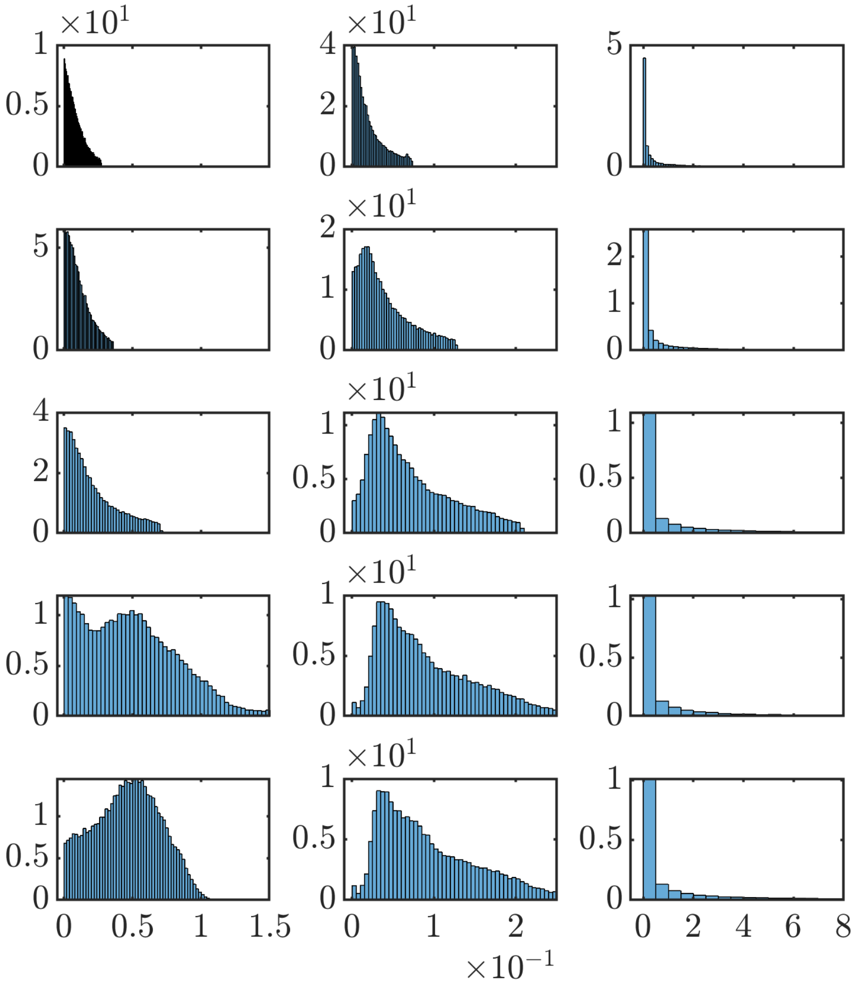

Figure 5.16: Probability density function histograms for the absolute differences between various Gulf Stream LCS quantities computed from numerically defined velocity of the best available resolution (grid spacing 0.125° latitude/longitude) and of reduced resolution levels over grids with spacing 0.2° (Row 1), 0.5° (Row 2), 1° (Row 3), 2° (Row 4) and 3.33° (Row 5). The quantities in question are the FTLE (Column 1), LAVD (Column 2) and S^2 (Column 3).

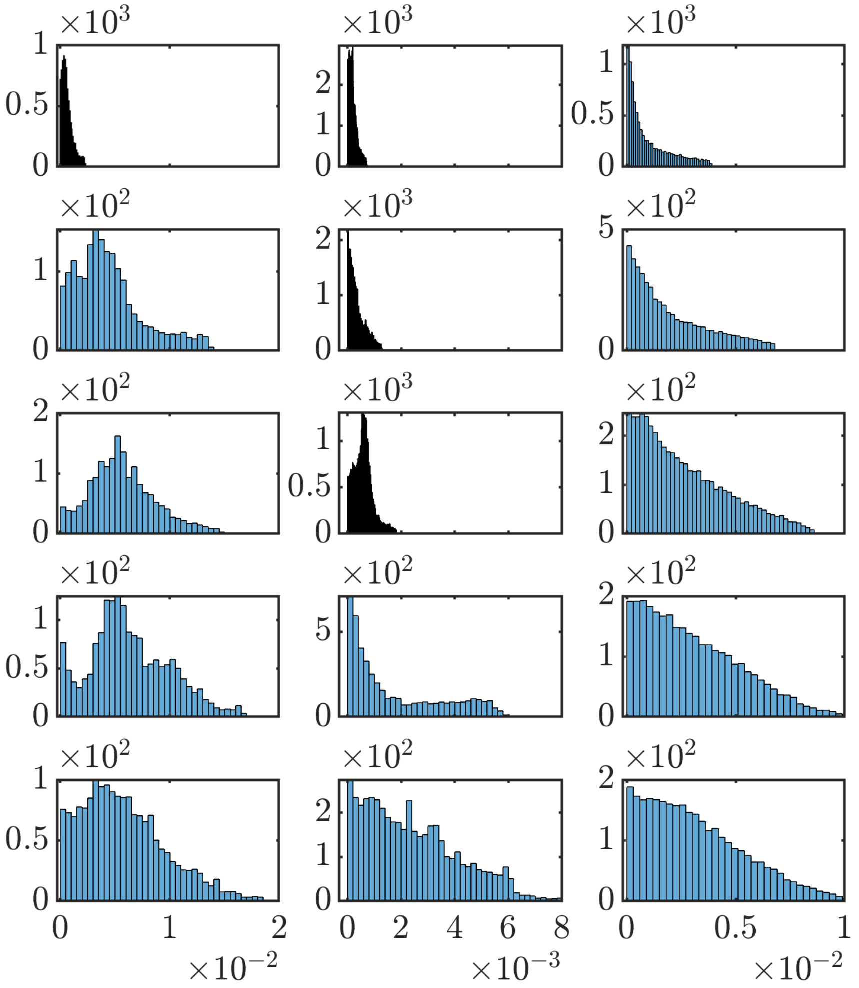

Figure 5.17: Probability density function histograms for the absolute differences between various Gulf Stream LCS quantities computed from numerically defined velocity of the best available resolution (grid spacing 0.125° latitude/longitude) and of reduced resolution levels over grids with spacing 0.2° (Row 1), 0.5° (Row 2), 1° (Row 3), 2° (Row 4) and 3.33° (Row 5). The quantities in question are the third singular vector of $\mathbf{P}$ (Column 1), the third eigenvector of $\mathbf{\Delta}^D$ (Column 2) and the CSC (Column 3).

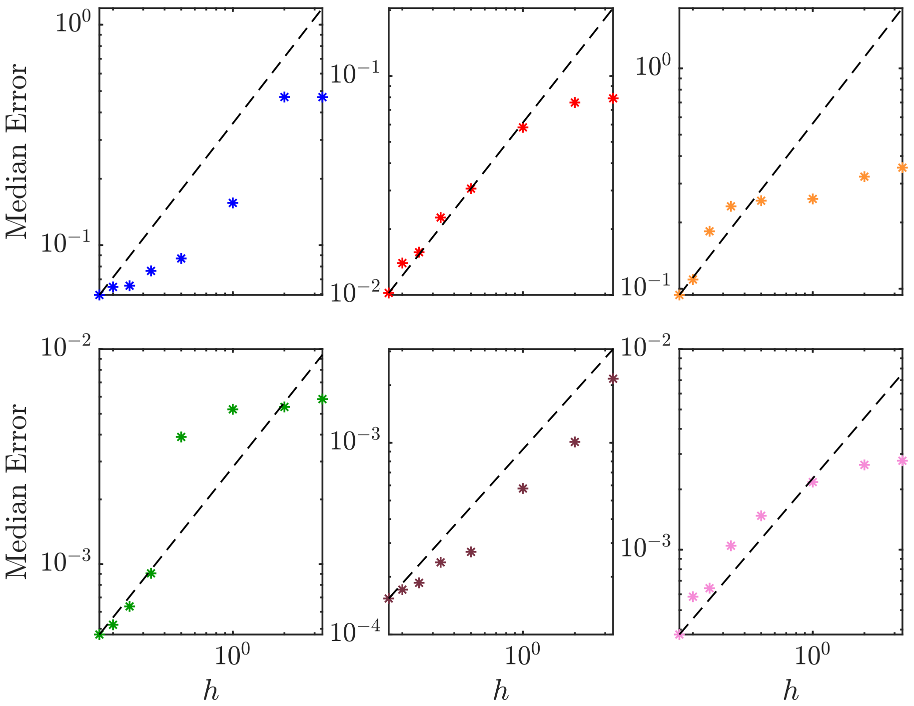

Figure 5.18: Log–log plots of the median absolute differences between various LCS quantities for the Gulf Stream flow computed from numerically defined velocity of the best available resolution and of reduced resolution defined over different grid mesh sizes h. The quantities in question are (from left to right): the FTLE, the LAVD, S^2 (row 1), the transfer operator third singular vector, the dynamic Laplace operator third eigenvector and the CSC (row 2). For comparison against this data, a dashed line of slope 1 (in log–log space) has been included within each plot.

observed in the results for the previous two flows.

Finally, we look to see if there is a correlation between the median errors recorded and the grid spacing of our velocity data h using the log–log plots featured in Figure 5.18. Like with the other two flows, we observe a positive correlation between these two variables for each of these methods. The results for the FTLE and dynamic Laplace operator methods both appear to have reasonably strong positive correlation, while the LAVD, S^2 and CSC errors seem to have a positive correlation for smaller h before this is seen to plateau as h gets larger. The transfer operator follows a line almost of slope 1 in log–log space, before suddenly spiking in value and assuming a different slope for larger h.

5.4 Chapter Summary

From our qualitative assessment of the impact of spatial resolution on the detection of LCSs, it is evident that the robustness of an LCS detection method is dependent on the size of the coherent structures present within a flow. For most of our detection methods, making small reductions in the resolution does not have a highly significant impact on the results obtained. We only begin to notice a change in these results when the resolution becomes so low that an LCS can no longer be detected as not enough velocity grid points cover the lengthscale or area of this structure. The FCM clustering method is a notable exception in this case, as the sets of coherent clusters generated by this method have been found to be inconsistent between different resolution levels (see in particular Figures 5.3 and 5.9). These inconsistencies are observed as a result of how the FCM clustering algorithm groups together Lagrangian flow trajectories with each resolution level, and as mentioned earlier can be repaired by altering the computational parameters pertinent to this method; such as the number of flow trajectories used or the cluster sharpness parameter m.

From our quantitative analysis of the LCS error histograms, we estimate that the distribution of these errors will take the form of a Dirac delta distribution when our velocity has infinite spatial resolution. As the resolution decreases, the errors have a decaying distribution with the mode being 0; though for some of our methods the mode can increase with smaller resolutions. This decay continues until the resolution becomes so unworkably small that the errors become completely uncorrelated and assume a uniform distribution. We have also been able to show that the median values of these errors have some positive correlation with h, the mesh size of the grid over which our velocity has been defined.

Chapter 6

Conclusion

The core focus of this research was to investigate the *self–consistency* of a range of Lagrangian coherent structure (LCS) detection methods, in particular how these methods respond to velocity uncertainty and how dependent these methods are on spatial velocity resolution. We have implemented these detection methods on analytically defined flow models initially to gain some intuition into how each method works and what kinds of coherent structures are detected with each method. However, more importantly we also consider three flavours of realistic velocity data to gain some idea of which of these detection methods would be the most suitable for use in real world analysis of fluids.

The eight Lagrangian coherent structure detection methods considered in this research form the tip of the iceberg as far as this concept is concerned, with many other detection methods available and each one detecting different types of coherent structures through various computational means. That said, out of these eight methods some clear "winners" and "losers" have emerged from the group, based on the structures these methods are able to identify, how user–friendly their computational algorithms are, and how the functionality of these methods is impacted by the quality of the velocity data representing a flow; specifically how much data is available and how accurate this data is. Ultimately, no Lagrangian coherent structure detection method is perfect and which method is considered for which particular data set will have to be determined on a case–by–case basis. While there is no one–size–fits–all method for detecting these coherent structures, there are some which may be considered a more go–to option to consider over others.

The finite time Lyapunov exponent (FTLE) method is one of the most frequently used Lagrangian coherent structure detection methods in the literature, owing primarily to its relative ease in computation in comparison to most other methods. Despite this method having been shown to be unreliable for Lagrangian coherent structure detection in some flow systems (Haller 2011, Balasuriya et al. 2018), our results indicate that this method is more than capable of detecting viable coherent flow objects that are consistent with those produced by other detection methods. Despite some minor differences and discrepancies, the results for this method featured in Chapter 3 are comparable with the results of other methods, such as maximal ridges of the stochastic sensitivity field and material curves produced by the variational LCS method. While this method is primarily used to detect these flow barriers, the FTLE scalar field can also be used to identify more robust particle flow from its minimal values. Structures such as the rotating gyres in the Double gyre flow, the zonal jet within the Bickley jet flow, the Kelvin–Helmholtz central flow channel and vortices within the Bickley jet, Kelvin–Helmholtz and Gulf Stream flows were all identifiable to a considerable degree and shared similar shapes, sizes and features to structures of similar form identified from other methods such as the LAVD, CSC and S^2 robust sets. The FTLE method is robust and workably reliable when the resolution of velocity data is reduced to a reasonable level, however the presence of velocity uncertainty within said data will cause the maximal ridges of the FTLE scalar field to considerably change in shape and smoothness. That said, regions of a flow domain corresponding to lower levels of FTLE are more robust against this velocity uncertainty, so if this method was to be used to detect more robust coherent structures instead of hyperbolic flow barriers, the results obtained would most certainly be more reliable.

Out of all of the methods considered in this study, the variational LCS detection method is by far the least recommended detection method based on our results. While this method has been sold as a viable alternative to the FTLE method for the detection of flow segregating material curves, executing this method numerically has proven particularly arduous. In our research, three different algorithms were tested to execute this method (with possibly more being available in the literature), with each one relying on a different set of computational parameters and different numerical criteria to regulate integration of the Cauchy–Green eigenvector fields and choose the most influential flow barriers contained within a flow system. Ideally, each algorithm should produce similar coherent structures and while for most of the flows detailed above the general picture remains the same, the individual foliations of hyperbolic variational structures changes with each algorithm. This is owed to the fact that each of these algorithms give very ad hoc ways of selecting coherent structures from collections of hyperbolic strainlines. It is therefore no surprise that this method is highly sensitive to uncertainty within velocity data, so much so that the collection of structures produced (regardless of the flow system considered) can completely deform to the extent that they bear no resemblance to their deterministic counterparts. The structures do hold their form under reduced velocity resolution, but the overall foliation of structures changes as the (highly sensitive) acceptable set of candidate points change. Unlike the FTLE method, where low values of this quantity can be used to highlight more robust particle flow, such inferences cannot be made from the hyperbolic variational LCSs as these are merely material curves representing the structures from which flow trajectories repel at a locally maximal rate.

The Lagrangian averaged vorticity deviation (LAVD) method has proven to be a useful tool for detecting elliptic or vortical structures present within a flow system. The method is easy to program, with the LAVD computable using a system of ODEs or an integral, and the quantity itself can be used to isolate vortical structures corresponding to strong rotational coherence, or even stream like objects with strong Lagrangian shear (such as the Gulf Stream or the shear layers observable in version 1 of the Kelvin–Helmholtz flow). The method is slightly sensitive to noise in that the value of the LAVD changes when additional uncertainty is added to the velocity, though the overall shapes of the coherent structures detected remain consistent, and this method is resistant to changes in spatial resolution so long as this resolution is not too coarse (a consistent observation for most of our LCS detection methods). The only downside to this method is that extensionally coherent structures, such as the central flow channels in the Kelvin–Helmholtz systems, cannot be detected as this method is designed to isolate rotationally coherent flow objects only, which it accomplishes well.

The stochastic sensitivity (S^2), a relatively new method, was not expressly designed for detecting Lagrangian coherent structures. That said, the flow attributes observable from this quantity and other sets and characteristics derived from it behave in very similar manner to these Lagrangian coherent structures obtained from other methods. The stochastic sensitivity method is quite flexible in that the S^2 field itself can be used to find flow regions where the uncertainty of a flow trajectory, and therefore the likelihood of more chaotic flow behaviour, is heightened; or S^2 can instead be used to isolate robust sets (containing a wide array of flow objects) within which particle flow is more controlled. The S^2 field is quite similar to the FTLE scalar field for each of our systems, though this is not always guaranteed (Balasuriya 2020*a*), and the robust sets corresponding to small stochastic sensitivity against a lengthscale and velocity uncertainty level of choice can identify a wide variety of rotationally and extensionally coherent flow objects. As a slight disadvantage to this method, the S^2 field exhibits large variations. Not only is it harder to extract coherent structures without scaling the field, but the maximal ridges are more susceptible to velocity uncertainty. This velocity uncertainty does not have a considerable impact on the robust sets, or regions of the domain where stochastic sensitivity is low. This method is also reasonably robust to changes in data resolution, so long as the flow contained within a system is not highly volatile (e.g. in a flow such as the Bickley jet; which is regular across most of the domain except where the eddies meet the zonal jet). However, the computational steps required to produce this quantity are rather extensive and can consume a reasonably large amount of system memory.

The transfer operator and dynamic Laplace operator methods are very similar in the sense that both of these methods seek particle density distributions which remain “robust” (consistent) under a flow. From this, one can obtain a curve or line along which to divide a flow domain into two separate coherent sets based on the overall flow patterns observable within a system. From the results obtained for both of these methods (except in the case of the Unsteady Stuart Vortex system), each of these methods has been able to divide our flow domain into two coherent sets in more or less the same fashion. Systems such as the Bickley Jet or Kelvin–Helmholtz flows manage to divide nicely into two coherent sets, whereas in the case of a system such as the Unsteady Stuart Vortex or the Gulf Stream, these methods divide the flow domain in half in a more or less trivial fashion. There is also the further complication of deciding which singular vector/eigenvector of the transfer operator/dynamic Laplace operator to use in order to find these coherent sets. The second vector is normally the one chosen, however one can also choose a subsequent vector so long as there is a sizeable spectral gap between its respective eigenvalue and the next eigenvalue in the sequence. There are even techniques available in the literature for producing further coherent sets using multiple transfer operator singular vectors or dynamic Laplace operator eigenvectors with tools such as k–means clustering (Froyland & Junge 2018) or the SEBA algorithm (Froyland et al. 2019); however the latter algorithm was primarily used by Froyland et al. (2019) to extract elliptic coherent structures from systems such as the Gulf Stream and numerically simulated turbulence models, which could be achieved in an easier fashion using a method such as the LAVD. Both of these methods are robust to acute velocity uncertainty, though the boundaries of the coherent sets obtained from the transfer operator method become considerably more fuzzy even though the sets retain their shape. We do not observe this with the dynamic Laplace operator method, as these inconsistencies are ironed out by the Laplacian operator matrix used to generate the dynamic Laplace operator. Both of these methods also show resistance to changes in velocity resolution, though as the grid spacing of velocity data becomes unworkably large, the scalar fields for the vectors produced contain a large number of discontinuities and the coherent sets become considerably harder to identify.

The fuzzy c–means clustering method is a sparse LCS detection method used to find k coherent clusters within a flow domain based on the proximity of a collection of flow trajectories within this domain. The method itself is efficient and easy to program as it relies on a pre–prepared clustering algorithm available in MATLAB, and the membership probabilities for the k clusters sought are robust against velocity uncertainty. Of the method's disadvantages, while it is normally robust against changes to the resolution of velocity data, we have seen instances of the clusters formed from this method changing in shape completely when the resolution is lowered, then changing back to their original form following another change in resolution. As mentioned earlier, this can be rectified by sampling more flow trajectories or by changing the sharpness parameter m for this method. The biggest issue with this method, however, is the requirement of the user to pre–determine k, the number of clusters expected to exist within a flow system. For some analytically defined systems, such as the Double Gyre or Bickley jet flows, this can be a relatively easy task. However, in a real world flow such as the Gulf Stream, choosing k is a very difficult task which will require a fair amount of trial and error, and most likely will not isolate key structures such as streams or vortices but will rather divide the flow domain into k portions of more or less equal size that do not speak a great deal about the most influential flow patterns observable within our dynamical system of interest.

The coherent structure colouring (CSC) method can be considered an improvement on the fuzzy c–means clustering method, as the number of clusters does not need to be guessed beforehand, and the method groups together flow trajectories based on how these mix and braid around each other as the flow progresses; rather than simply how close their points remain bundled together over the course of the flow. Out of all of the methods considered in this research, the CSC method has stood out as one of the strongest. It is a relatively easy method to program (especially in comparison to some methods such as the variational LCS method), and in one calculation can give a detailed picture of all of the key coherent structures present within a flow, including streams, jet cores, vortices and even the rough outlines of flow barriers. The method shows considerable resistance against velocity data uncertainty, and works well for some lower levels of resolution, though like with the transfer operator and dynamic Laplace operator methods if the resolution of the data is too small, then discontinuities will form within the CSC coefficient field and coherent structures will be harder to identify. The downsides are the large amount of system memory required if there are many trajectories or time steps and the large computation time if the adjacency matrix entries are calculated iteratively. This generally should not be too much of an issue because as Schlueter-Kuck & Dabiri (2017) have pointed out, a decent enough image of coherent structures within a flow can be produced using a relatively small number of trajectories. Our research has verified this to some extent, as the CSC method has produced viable coherent structures for all of our methods using considerably fewer points than the number of points used for other methods such as the FTLE, LAVD or variational LCS method. More in depth analysis into how many trajectories are required to obtain viable LCS results from each method is recommended, and is suggested as future work.

In light of our findings, we consider the most functional LCS detection methods out of those considered in this study to be the CSC, LAVD and stochastic sensitivity methods. The former two methods rely on relatively simple computation algorithms, produce viable coherent structures amidst the presence of velocity uncertainty and relatively poor (but not unworkably poor) levels of data resolution, and produce two–dimensional Lagrangian coherent structures which give a more complete picture of the most robust flow regions of a dynamical system rather than just detecting the boundaries which separate them. Even though the stochastic sensitivity computational algorithm is complex and shows mild sensitivity to velocity uncertainty and lack of resolution, we also recommend this method as this quantity can be scaled to detect flow barriers and more robust flow regions which depend on computational parameters connected directly to the velocity data of the system of interest. While the FTLE method shows some susceptibility to velocity uncertainty, the algorithms for this method are far easier to use than the variational LCS method and the results produced appear to correlate with similar results obtained from other detection methods as noted earlier. We recommend considering the FTLE method over the variational LCS method for detecting hyperbolic flow barriers, and conducting further investigations into how the FTLE can be used to isolate more robust flow behaviour using minimal values of this quantity. The transfer operator and dynamic Laplace operator methods are robust against velocity data uncertainty and lack of resolution to a reasonable degree. However, these methods mostly do not achieve a great deal in terms of visualising coherent structures, as they seek to obtain robust particle density concentrations used primarily to divide a flow domain into two coherent "halves". Some systems (such as the Bickley Jet) are designed to be divided in this way, however we cannot guarantee that all real world flow domains will divide nicely into two coherent sets. While other techniques such as the SEBA algorithm exist for extracting more detailed coherent flow objects using multiple singular vectors/eigenvectors of these operators, the flow structures detected would look very similar to those obtained from the LAVD and CSC methods, and we therefore recommend the use of these methods instead if this is the desired outcome. We do not recommend the use of the algorithmically complex and susceptible variational LCS method, nor do we recommend the fuzzy c–means clustering method due to the requirement of the number of coherent clusters within a flow to be guessed, and the lack of detail that these clusters can provide especially in a harder to predict real world flow such as the Gulf Stream.

There is much scope for further analysis into the functionality of Lagrangian coherent structure detection methods from the results displayed in this Thesis. Firstly, further resolution analysis can be undertaken to verify the positive relationship between velocity grid mesh size and LCS error obtained in this study; including the derivation of a slope or correlation coefficient relevant to these relationships. Secondly, analysis into the impact of computational parameters on the self–consistency of LCS results can be undertaken. These can be parameters relevant to all methods (e.g. time parameterisation, order of the ODE solving method) or can be method specific (e.g. allowance parameters for variational LCSs, box sizes for the transfer operator matrix, and the sharpness parameter for fuzzy c–means clusters). Thirdly, we can consider the synergy of various existing LCS detection methods to produce more efficient and/or more accurate LCS detection methods. We have attempted this in this research by combining the FTLE with the variational LCS method (see Section 2.2.4). Finally, as the LCS detection methods considered in this study have all been applied to two dimensional flows, these methods can be extended to define and detect LCSs in higher dimensional flows, as has so far been done with the variational LCS (Blazevski & Haller 2014), LAVD (Haller et al. 2016), transfer operator (Froyland et al. 2010, 2012), dynamic Laplace operator (Froyland & Junge 2018) and CSC (Martins et al. 2021) methods extended to three dimensional flows.

Appendix A

Verification of Detection Algorithms

In this Appendix, we detail efforts in validating the MATLAB code produced for each of our Lagrangian coherent structure detection methods on analytic Lagrangian coherent structure results for simplistic flows (where possible) or against results documented within relevant literature. We also validate this code for numerically defined flow systems using simulated velocity representing the simplistic Taylor–Green vortex system.

A.1 Validation of Code - Analytically Defined Flows

Rather than relying on published code for each Lagrangian coherent structure detection method that has been published in the literature, we develop our own MATLAB code for implementation of these methods in order to gain more of an understanding of the fundamentals of each of these methods. To ensure that this code works to the best of our knowledge, we validate this code by testing it either on simplistic flows where Lagrangian coherent structure metrics can be defined analytically or, where this is not possible or not readily straightforward to implement, on resultant coherent structures for toy models published within the literature and accepted to be "true".

We focus on each Lagrangian coherent structure detection method in the order in which they were defined in Chapter 2. We begin with the finite time Lyapunov exponent implemented on the Poiseuille flow system, a simplistic flow consisting of a linear flow channel with consistent linear particle movement to the right–hand direction in x. The velocity for this system takes the form

$$\dot{\mathbf{x}} = \mathbf{u}\left(x, y\right) = \begin{bmatrix} y^2 - 1 \\ 0 \end{bmatrix} \tag{A.1}$$

and will be defined over the domain $\Omega = [0, 10] \times [-1, 1]$. The relative simplicity of this system enables its flow map at any initial point $\mathbf{x}_0 = (x_0, y_0) \in \Omega$ over the flow interval $[t_0, t_0 + T]$ to be computed analytically, and this happens to take the form

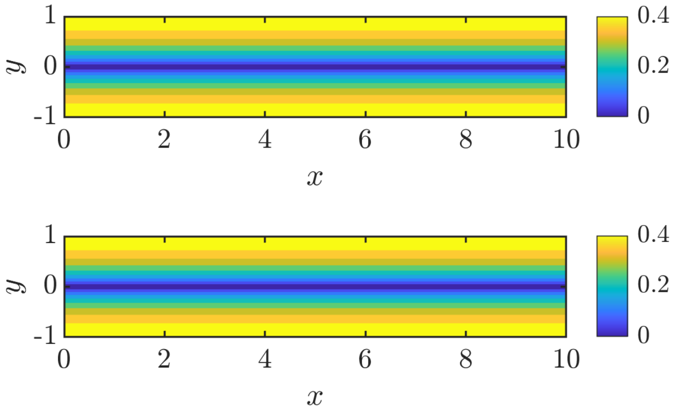

Figure A.1: The FTLE Φ_0^5 for the Poiseuille flow computed using its analytic expression (top) and numerically using our own MATLAB code (bottom). The supremum for the absolute error between these calculations is 5.5095×10^{-14}.

$$\mathbf{F}_{t_0}^{t_0+T}(x_0, y_0) = \begin{bmatrix} x_0 + (y_0^2 - 1)\, T \\ y_0 \end{bmatrix}. \tag{A.2}$$

After some derivation, the finite time Lyapunov exponent for this system

$$\Phi_{t_0}^{t_0+T}(x_0, y_0) = \frac{1}{|T|} \ln\left(\sqrt{(2T^2 y_0^2 + 1) + 2T\,|y_0|\sqrt{T^2 y_0^2 + 1}}\right). \tag{A.3}$$

We test our MATLAB code on a 1000×400 grid of points within Ω defined above, with $t_0 = 0$ and $t_0 + T = 5$. The resultant finite time Lyapunov exponents calculated both from (A.3) analytically and numerically from our code can be seen Figure A.1. The finite time Lyapunov exponent for this system produces a trough along $y = 0$, which represents a coherent flow channel for this system, along with two ridges along $y = -1$ and $y = 1$ which indicate where particle stretching capability is greatest within this system. From the Figure caption the largest recorded value of absolute error between the analytic and numerical quantities for Φ_0^5 is 5.5095×10^{-14}, which is a rather impressively low bound for the error. The likely causes for such a low error include an accurate numerical computation for the flow map using ode45 ((A.2) was not used for this numerical computation), and a refined grid of points helped ensure low errors when computing the gradient of the flow map $\nabla\mathbf{F}_{t_0}^{t_0+T}(\mathbf{x}_0)$.

While we could use the Poiseuille system (A.1) to test the variational method for Lagrangian coherent structure detection, this would involve numerically checking the eigenbasis of the Cauchy–Green strain tensor (2.2), along with analytically checking other conditions necessary for the algorithm including the acceptable set $\mathcal{G}_0$. Instead, we will test this algorithm by attempting to replicate the results for the Double Gyre system

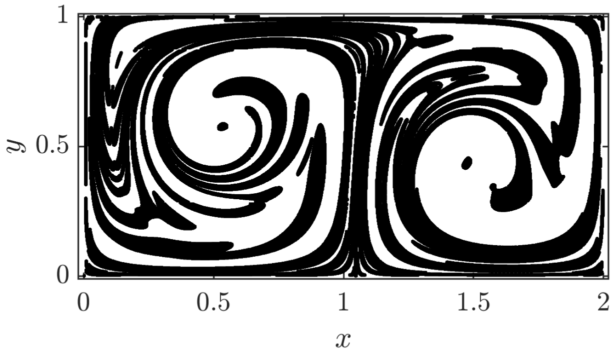

Figure A.2: The field of acceptable points $\mathcal{G}_0$ used to generate hyperbolic variational LCSs using Farazmand and Haller's (2012) algorithm. Compare with Figure 9(a) in Farazmand and Haller (2012).

published in Farazmand & Haller (2012) using a similar set of computational parameters.

The Double Gyre system takes the form

$$\dot{\mathbf{x}} = \mathbf{u}(x, y, t) = \begin{bmatrix} -\pi A \sin(\pi f(x,t)) \cos(\pi y) \\ \pi A \cos(\pi f(x,t)) \sin(\pi y) \frac{\partial f}{\partial x} \end{bmatrix},$$

where

$$f(x,t) = x^2 (\epsilon_* \sin(\omega t)) + x (1 - 2\epsilon_* \sin(\omega t)),$$

$A = 0.1$, $\epsilon_* = 0.1$ and $\omega = \pi/5$, defined over the flow domain $\Omega = [0,2] \times [0,1]$ with the flow period ranging from $t_0 = 0$ to $t_0 + T = 20$. Farazmand & Haller (2012) implement their first variational hyperbolic Lagrangian coherent structure detection algorithm (VLCS Algorithm A) on this system, and publish an acceptable set $\mathcal{G}_0$ and the resultant Lagrangian coherent structures for this system in their paper. Since all of these detection algorithms rely on the same strategy to solve the first Cauchy–Green eigenvector field, with the resultant structures and hyperbolic curves regulated in different ways, all we need to do is implement VLCS Algorithm A on the Double Gyre system (3.1) to ensure that integration of the field $\boldsymbol{\xi}_1$ has been done appropriately.

Figure A.2 shows the acceptable set of initial points $\mathcal{G}_0$ for this system, from which we seed initial integration points for solving the $\boldsymbol{\xi}_1$ field, and within which all strainlines produced must remain. Comparing this qualitatively with Figure 9(a) in Farazmand & Haller (2012), our acceptable set matches nicely with the published Figure as similar shapes and structures are captured within this set. Figure A.3 shows the hyperbolic variational LCS produced for the Double Gyre system with the longest length. We use the same parameters as in Farazmand & Haller (2012), except we extend the fail allowance parameter $\mathcal{L}_f = 1$ instead of 0.2 to allow for extra uncertainty. Our curve matches that featured in Figure 10(b) in Farazmand & Haller (2012), and contains more detail as

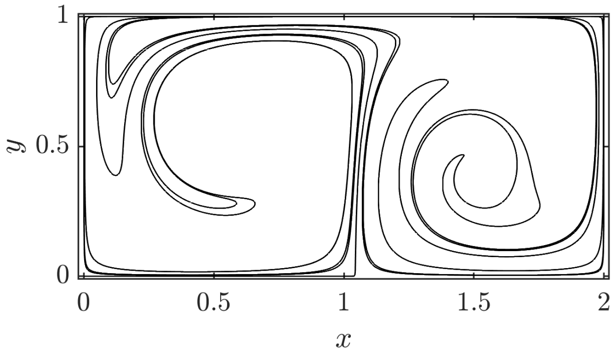

Figure A.3: The hyperbolic variational LCS of longest arc length generated using Farazmand and Haller's (2012) algorithm. Compare with Figure 10(b) in Farazmand and Haller (2012).

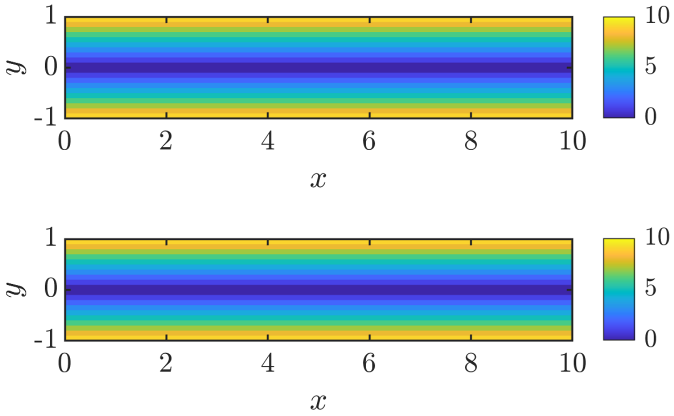

Figure A.4: The LAVD Θ_0^5 for the Poiseuille flow computed using its analytic expression (top) and numerically using our own MATLAB code for solving the integral (2.6) (bottom). The supremum for the absolute error between these calculations is 1.1546×10^{-14}.

we have made $\mathcal{L}_f$ considerably larger than theirs. From these results, we are satisfied that computation of the Cauchy–Green eigenbasis and the solution curve of $\boldsymbol{\xi}_1$ using our MATLAB code is acceptable.

Next, we return to the Poiseuille flow (A.1) to test out our MATLAB algorithms for the Lagrangian averaged vorticity deviation method. While we don't expect any vortical structures within this linear channel flow, this flow can still be used to test out the method as the LAVD quantity can be computed analytically, and takes the form

$$\Theta_{t_0}^{t_0+T}(x_0, y_0) = 2T\,|y_0|\,. \tag{A.4}$$

Figure A.4 shows the Poiseuille flow LAVD Θ_0^5 computed from numerically evaluating the integral (2.6), while Figure A.5 shows the same field computed instead by solving the

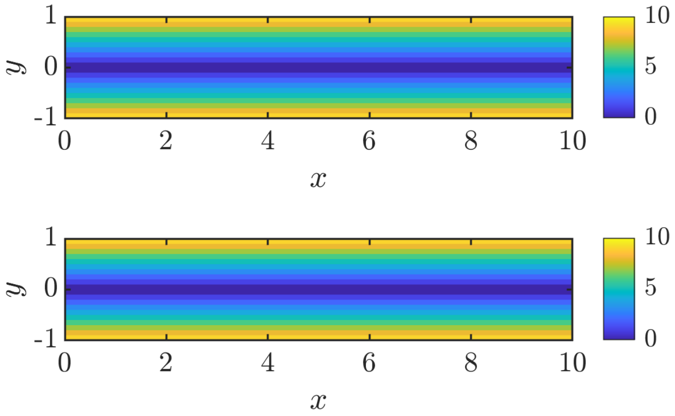

Figure A.5: The LAVD Θ_0^5 for the Poiseuille flow computed using its analytic expression (top) and numerically using our own MATLAB code for solving the extended ODE system (2.7) (bottom). The supremum for the absolute error between these calculations is 5.3291×10^{-15}.

extended ordinary differential equation system (2.7); with both Figures comparing these results to their analytical counterparts. Again, no vortical structures were anticipated to exist within this flow, but the LAVD of this system starts off at zero along the jet positioned at $y = 0$, and gradually gets larger towards the upper and lower edges of the domain Ω. In both cases, the supremum of the error between the analytic and numerical LAVD quantities is extremely low, verifying our MATLAB code for both of these algorithms. We compute the LAVD primarily using the extended ODE system (A.5) due to its relative ease of computation and, as Haller et al. (2016) have claimed, using (A.5) to compute the LAVD is more convenient and less prone to errors than evaluating the integral (2.6) numerically.

Next, we test our algorithm for computing the stochastic sensitivity of a flow system using the Poiseuille flow (A.1) once more. As with the FTLE and LAVD methods, the stochastic sensitivity of the Poiseuille flow can be computed analytically, which after some derivation comes out to be

$$S^2\left(x_0, y_0\right) = \frac{2}{3}y_0^2((t_0+T)^3-t_0^3)+T+\sqrt{\frac{4}{9}y_0^4((t_0+T)^3-t_0^3)^2+y_0^2((t_0+T)^2-t_0^2)^2}. \tag{A.5}$$

This equation looks rather complicated, however it simplifies quite readily as we have $t_0 = 0$, to become

$$S^2\left(x_0, y_0\right) = \frac{2}{3}y_0^2T^3 + T + \sqrt{\frac{4}{9}y_0^4T^6 + y_0^2T^4}. \tag{A.6}$$

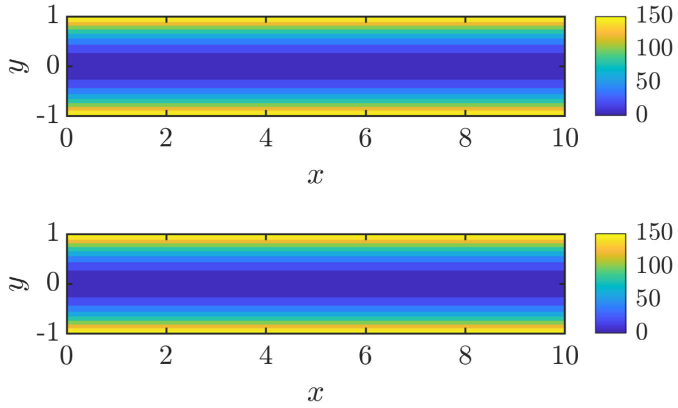

Figure A.6: The Stochastic Sensitivity for the Poiseuille flow over the time interval $[0, 5]$ computed using its analytic expression (top) and numerically using a grid which covers the range of the flow map at time $t_0 + T = 5$ (bottom). The supremum for the absolute error between these calculations is 3.2631×10^{-4}.

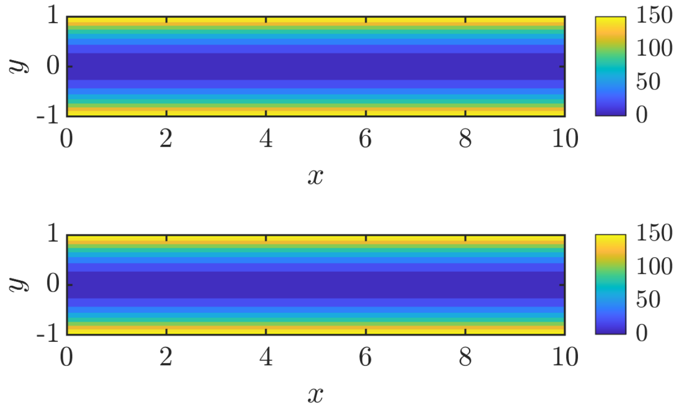

Figure A.7: The Stochastic Sensitivity for the Poiseuille flow over the time interval $[0, 5]$ computed using its analytic expression (top) and numerically by defining a star grid at every flow map point at time $t_0 + T = 5$ (bottom). The supremum for the absolute error between these calculations is 3.5390×10^{-4}.

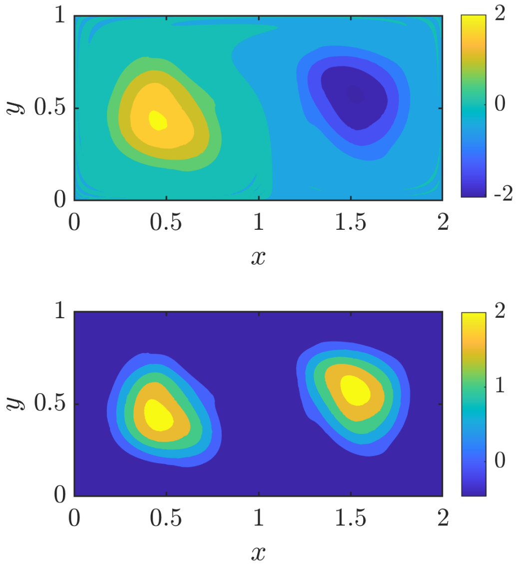

Figure A.8: Scalar fields representing the second (top) and third (bottom) singular vectors of the transfer operator for the second incarnation of the Double Gyre system over the flow interval $[0, 10]$. Compare with Figures 18(a) and 18(b) in Tallapragada and Ross (2013).

Figure A.6 shows the Stochastic Sensitivity computed for the Poiseuille flow by producing a uniform grid covering the range of flow map points at time $t_0 + T = 5$, while Figure A.7 shows the same quantity computed instead by defining a star grid at each flow map point; with each result compared to its analytically defined equivalent in each Figure. The stochastic sensitivity fields do not look too unlike the FTLE scalar field shown in Figure A.1, with lower values present within the linear channel and larger values observable along the upper and lower boundaries of the domain. The MATLAB code for both of our methods appears to be working well once again, though unlike in the case of the FTLE and LAVD methods, the supremum of the error for the stochastic sensitivity is larger, now of order 10^{-4} in both cases. This is likely due to more calculations involved in generating these fields, including several instances of numerical integration and interpolation of relevant quantities, particularly for the grid extension method (our first method); which generates considerably more numerical error than it would for the FTLE or LAVD methods. That said, this level of error is still acceptable enough for verification of the codes for this method. Due to system memory considerations, we primarily use the star grid method for computing the stochastic sensitivity in this research unless otherwise specified.

Next, we turn our attention to the transfer operator method by returning to the above defined Double Gyre system (3.1), with a slight alteration to the parameters and time interval to match those used in Tallapragada & Ross (2013); these being $A = 0.25$, $\epsilon = 0.25$

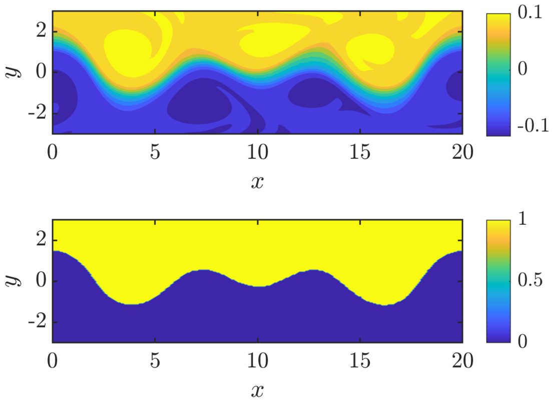

Figure A.9: The scalar field representing the second right singular vector of the transfer operator for the Bickley jet system (top) and two coherent sets relevant to the initial flow time $t_0 = 0$ days produced from this field (bottom). Compare with Figure 2(g) and the first plot in the left column of Figure 6 in Hadjighasem et al. (2017).

and $\omega = 2\pi$, with $t_0 = 0$ and $t_0 + T = 10$. The domain Ω remains the same as used above for the variational LCS method. Figure A.8 shows the scalar fields produced from the second and third right singular values of the transfer operator computed using this version of the Double Gyre flow. The second singular vector field divides the domain Ω at time t_0 approximately in half in producing two coherent sets, while the third singular vector field instead chooses the two gyres as one coherent set with the remainder of Ω forming the other set. The use of singular vectors other than the second when using the transfer operator or dynamic Laplace operator has been investigated in the literature (Froyland & Junge 2018, Froyland et al. 2019) and will be discussed further later on in this Chapter. These scalar fields look comparable with those produced in Tallapragada & Ross (2013), even though the scale of the singular vector values is different and the colours of the third singular vector field have been inverted (indicating an inversion of the sign of the singular vector values in this case). This is acceptable as a singular vector, like an eigenvector, can be multiplied by any scalar without the formation of the vector changing. Tallapragada & Ross (2013) also publish the first six eigenvalues generated for the matrix $\mathbf{P}^\top\mathbf{P}$, which were 1, 0.9997, 0.9995, 0.9987, 0.9981 and 0.9971. Our computations produced the six leading eigenvalues 1, 0.9997, 0.9995, 0.9986, 0.9986 and 0.9985, which differ slightly due to variations in computational techniques and numerical estimation errors, though the differences are not so large that they become a concern.

As an added check for the transfer operator method, we test this on a version of the Bickley jet flow system used in Hadjighasem et al. (2017) as we use the Bickley jet for LCS analysis throughout this Thesis. The Bickley jet velocity system takes the form

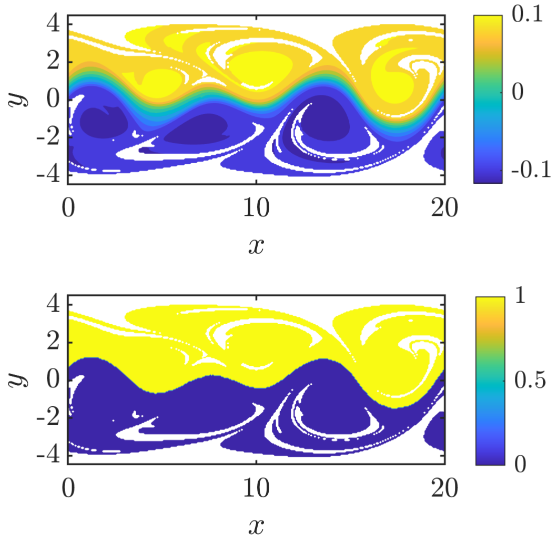

Figure A.10: The scalar field representing the second left singular vector of the transfer operator for the Bickley jet system (top) and two coherent sets relevant to the final flow time $t_0 + T = 11$ days produced from this field (bottom). Compare with Figure 4(a) and the first plot in the right column of Figure 6 in Hadjighasem et al. (2017).

$$\dot{\mathbf{x}} = \mathbf{u}(x, y, t) = \begin{bmatrix} u(x, y, t) \\ v(x, y, t) \end{bmatrix},$$

where

$$u(x, y, t) = U_0 \operatorname{sech}^2(\frac{y}{L_0}) + 2U_0 \operatorname{sech}^2(\frac{y}{L_0}) \tanh(\frac{y}{L_0}) \sum_{i=1}^{3} A_i \cos(k_i(x - c_i t)),$$

$$v(x, y, t) = -U_0 L_0 \operatorname{sech}^2(\frac{y}{L_0}) \sum_{i=1}^{3} A_i \sin(k_i(x - c_i t)),$$

with parameters $U_0 = 62.66\text{ms}^{-1} = 5.4138\text{Mmd}^{-1}$, $L_0 = 1.770\text{Mm}$, $A_1 = 0.075$, $A_2 = 0.4$, $A_3 = 0.3$, $k_i = (2i)/r_e$, $i = 1, 2, 3$, $r_e = 6.371\text{Mm}$, $c_2 = 0.205U_0$, $c_3 = 0.461U_0$ and $c_1 = c_3 + ((\sqrt{5} - 1)/2)(k_2/k_1)(c_2 - c_3) \approx 0.1446U_0$; over the flow domain $\Omega = [0, 20) \times [-3, 3]$, which is defined in units of megametres (Mm) and is periodic in the x-direction; and over the time interval which ranges from $t_0 = 0$ to $t_0 + T = 11$ days.

Figure A.9 shows the scalar field pertaining to the second right singular vector of the transfer operator, and two coherent sets produced for this system by dividing along the zero contour of the singular vector field. In line with the results shown in Tallapragada & Ross (2013), the transfer operator method divides the flow domain by the coherent

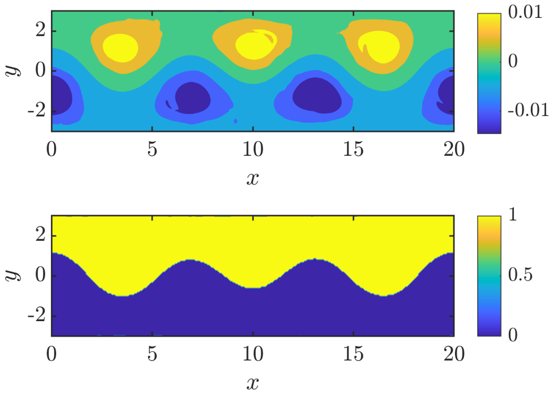

Figure A.11: The scalar field representing the second eigenvector of the dynamic Laplace operator for the Bickley jet system (top) and two coherent sets relevant to the initial flow time $t_0 = 0$ days produced from this field (bottom). Compare with the first plot in the right column of Figure 15 in Froyland and Junge (2018) and the second plot in the left column of Figure 3 in Froyland et al. (2019).

jet stream centred along $y = 0$, with the two vortex layers within the system identified as these sets. The singular vector field shows some more detail of the coherent vortices which exist within each vortex layer. Figure A.10 shows the scalar field and coherent sets obtained from the left singular vector of the transfer operator matrix, which show the evolution of the coherent sets computed for the initial time t_0 following advection through the Bickley jet system over T units of time. The coherent sets have considerably deformed in shape, but as Figure A.10 attests, practically no mixing of particles between these coherent sets is observed, producing two almost invariant coherent flow objects and fulfilling the objective of the transfer operator method. We have included these time advected coherent sets here for the sake of observation and testing the method, however throughout this research only the coherent sets obtained for the initial time t_0 from the right singular vector of the transfer operator $\mathbf{P}$ will be considered. Tallapragada & Ross (2013) do not detail the eigenvalues obtained for $\mathbf{P}^\top\mathbf{P}$ in this case, however in comparing the coherent sets visually, we were able to obtain similar results from similar computational parameters and techniques. The values of the singular vector entries do again look different, however this is once again acceptable as a singular vector or an eigenvector of a matrix can be multiplied by a scalar without affecting the vector.

We test the dynamic Laplace operator defined in matrix form using the Bickley jet flow once again to replicate an experiment considered in Froyland & Junge (2018) and Froyland et al. (2019). We use the same parameters for the flow, except this time we make $A_1 = 0.0075$ and $A_2 = 0.15$; and focus on the time interval which ranges from $t_0 = 0$ to $t_0 + T = 40$ days. Figure A.11 shows the scalar field pertaining to the second eigenvector of the dynamic Laplace operator and two coherent sets produced by partitioning along the

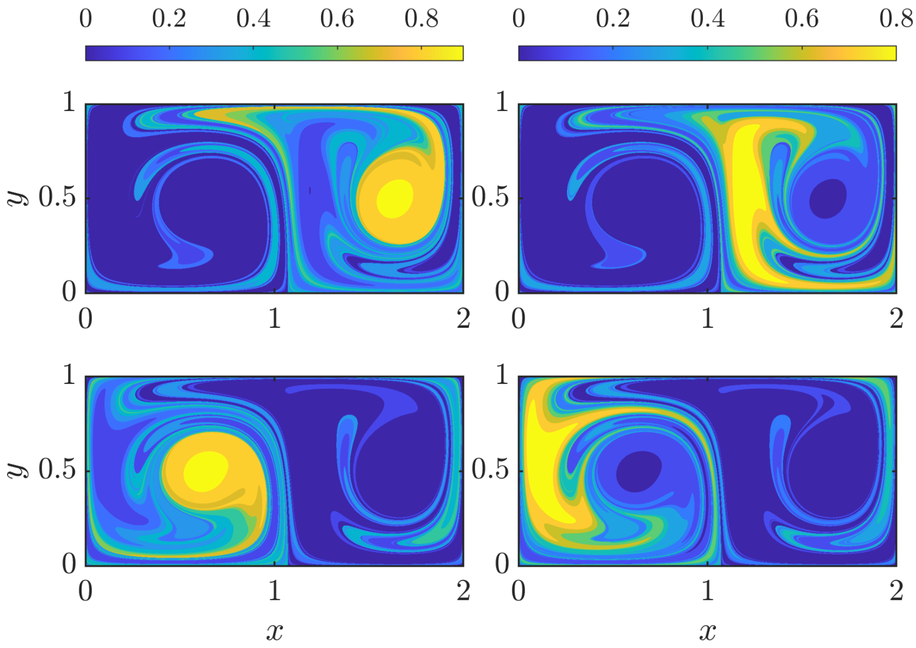

Figure A.12: Membership probability scalar fields for four FCM clusters within a third incarnation of the Double Gyre system over the flow interval $[2.5, 42.5]$ computed over a uniform grid of initial points. Compare with Figure 5 in Allshouse and Peacock (2015).

zero contour level of the eigenvector scalar field once more. Our results considerably match the published Figures bar a few slight differences, with similar coherent sets produced from the transfer operator method which differ slightly due to a change in the flow interval and some of the Bickley jet flow parameters. Again, the eigenvector field itself displays more detail regarding the coherent objects within this flow particularly by enabling visualisation of the coherent vortices within each of the vortex layers. The first six eigenvalues produced for the dynamic Laplace operator here were 0 (to within 10^{-15}), -3.9665, -6.4184, -6.4748, -6.5077 and -6.5522. While these differ slightly from the spectrum of eigenvalues provided in Froyland & Junge (2018) and Froyland et al. (2019), this is again likely due to estimation errors and differences in computational techniques. The results aren't wildly different, the first eigenvalue is close to machine zero and the spectral gap between the second and third eigenvalues has been maintained. We obtain similar eigenvalues and the same visual results displayed in Figure A.11 using the sparse definition of the dynamic Laplace operator and similar computational parameters to those detailed in Froyland & Junge (2018).

We next test the fuzzy c–means clustering method, by returning to the Double Gyre system (3.1) and our original parameter choices $A = 0.1$, $\epsilon = 0.1$ and $\omega = \pi/5$. We attempt to replicate the results published in Allshouse & Peacock (2015) for the flow interval defined from $t_0 = 2.5$ to $t_0 + T = 42.5$, using initial points defined in a uniform

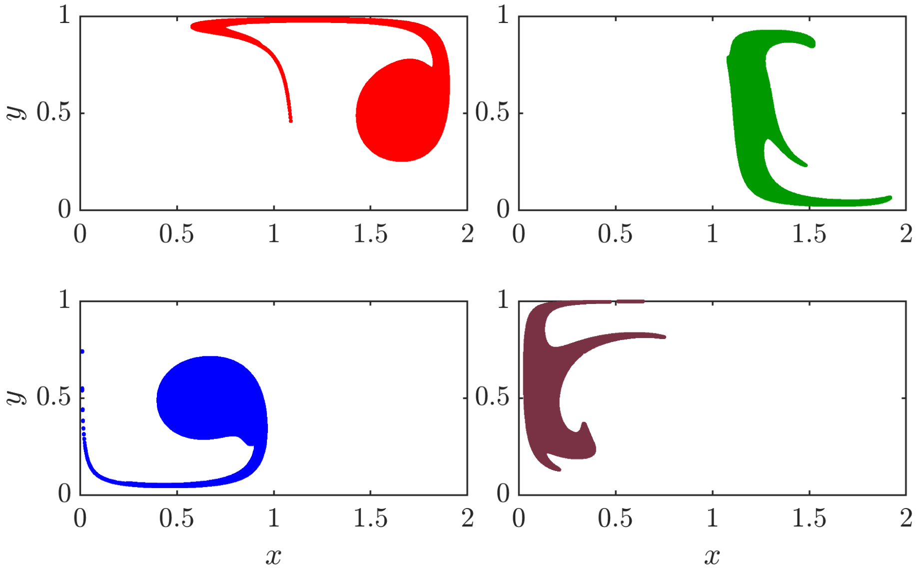

Figure A.13: Four fuzzy c–means clusters relevant to a third incarnation of the Double Gyre system over the flow interval [2.5, 42.5] computed over a uniform grid of initial points. Compare with Figure 5 in Allshouse and Peacock (2015).

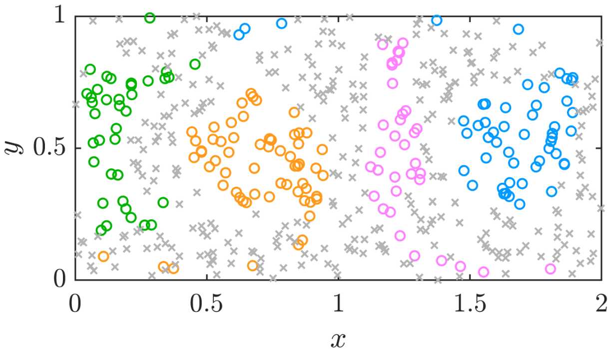

Figure A.14: Four fuzzy c–means clusters relevant to a third incarnation of the Double Gyre system over the flow interval [2.5, 42.5] computed from 500 uniformly distributed initial points. Compare with Figure 6 in Allshouse and Peacock (2015).

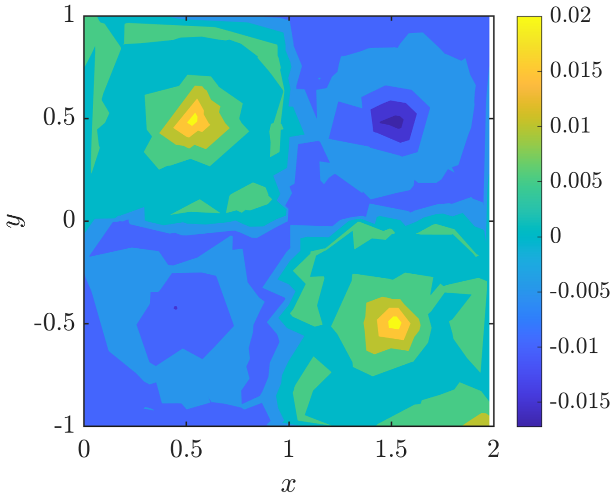

Figure A.15: The coherent structure colouring scalar field for the steady Quadruple Gyre system over the flow interval $[0, 40]$ computed from 300 uniformly distributed initial points. Compare with Figure 1(c) in Schlueter–Kuck and Dabiri (2017).

grid and a set of sparse initial points defined under a uniform distribution. In our first example, we compute four fuzzy c–means clusters for this system using a uniform grid of initial points. Figure A.12 shows the membership probability fields for four FCM clusters of this system, while in Figure A.13 we explicitly isolate these four clusters using a membership probability threshold of 70%, as was the case in Allshouse & Peacock (2015). The clusters we obtain look very similar to those obtained in Allshouse & Peacock (2015), with both gyres and their "tails" isolated along with deformed oblong type structures which sit to the left of each gyre. The clusters are presented in a different order to that presented in Allshouse & Peacock (2015), which was done to highlight the fact that the FCM Clustering algorithm was found to produce coherent clusters in an arbitrary order after each iteration. In Figure A.14, we test this algorithm on 500 (rather than 300 as in Allshouse & Peacock (2015)) sparse initial points chosen randomly under a uniform distribution. The plot indicates the initial locations of each trajectory at time $t_0 = 2.5$, with circular plot points of the same colour belonging to one of four FCM clusters at the 70% membership probability threshold once again. Points indicated by grey "x" marks do not belong to any of these clusters at this probability threshold. We were once again able to produce results similar to those published in Allshouse & Peacock (2015), however, unless otherwise specified, we apply the FCM clustering algorithm to trajectories computed from initial points defined over a uniform grid.

Finally, we test the coherent structure colouring method by replicating some experiments performed on analytically defined flow models in Schlueter-Kuck & Dabiri (2017). Specifically, we consider the Quadruple Gyre system, which has the same velocity as the Double Gyre system (3.1), except we extend the domain Ω in y in the fashion

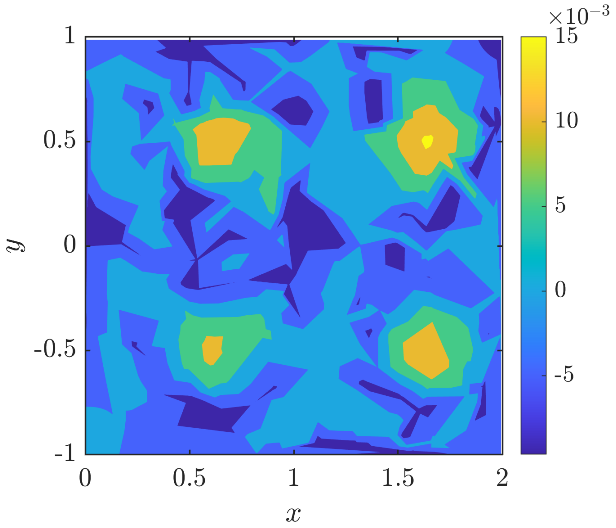

Figure A.16: The coherent structure colouring scalar field for the unsteady Quadruple Gyre system over the flow interval $[2.5, 42.5]$ computed from 300 uniformly distributed initial points. Compare with Figure 2(a) in Schlueter–Kuck and Dabiri (2017).

$\Omega = [0, 2] \times [-1, 1]$ to produce four rotating gyres instead of two with diagonally adjacent gyres rotating in the same direction. In this system, $A = 0.1$ again and we consider a steady (autonomous) version of this system with $\epsilon = \omega = 0$, and an unsteady (non–autonomous) version where we once again let $\epsilon = 0.1$ and $\omega = \pi/5$. We test the CSC method for both of these systems using a uniformly distributed set of 300 sparse initial points as was done in Schlueter-Kuck & Dabiri (2017). Figure A.15 shows the coherent structure colouring scalar field for the steady Quadruple Gyre system over the time interval $[0, 40]$, which identifies the top–left and bottom–right cells as coherent sets corresponding to similar positive values of the CSC coefficient, while the top–right and bottom–left cells have been identified with negative CSC values. These similarities are most likely due to the direction of gyre rotation being the same within cells of similar CSC value as suggested in Schlueter-Kuck & Dabiri (2017). Figure A.16 shows the CSC scalar field for the unsteady version of this system over the time interval $[2.5, 42.5]$, which in this case identifies the four gyres by similar positive CSC coefficient values and the remainder of Ω by negative CSC values, as in this case the gyres retain their shapes as they rotate, but fluid moves between the four cells as ϵ no longer has a value of 0. The scalar fields produced along with the CSC coefficient values recorded are strongly similar to those published in Schlueter-Kuck & Dabiri (2017), bar some minor differences. Like with the FCM Clustering method, unless otherwise specified we apply the CSC method to trajectory data defined from initial points within a uniform grid as the relevant velocity data will be available for all data sets considered in this study.

A.2 Taylor–Green Vortex (Validation of Code for Numerically Defined Velocity)

From the results detailed in the previous Section, we have developed a confidence that the MATLAB code and algorithms written for each Lagrangian coherent structure detection method work in accurately detecting these structures for a wide variety of flows. However, we have only tested this code on analytically defined toy models thus far, and since this research also considers the analysis of Lagrangian coherent structures within systems defined by Eulerian velocity data over a finite amount of time steps, we must test our code on a numerically simulated velocity data set whose analytically defined velocity is easy to compute. Therefore in this Section we apply the same MATLAB code to a relatively simple flow system defined using both analytically defined velocity equations and Eulerian velocity data. The system we consider is the steady–state (autonomous) Taylor–Green vortex system, which takes the analytic form

$$\dot{\mathbf{x}} = \mathbf{u}\left(x, y\right) = \begin{bmatrix} \cos\left(x\right)\sin\left(y\right) \\ -\sin\left(x\right)\cos\left(y\right) \end{bmatrix}. \tag{A.7}$$

We seek out Lagrangian coherent structures relevant to this system over the time interval ranging from $t_0 = 0$ to $t_0 + T = 5$ using all of the Lagrangian coherent structure detection methods considered for this study. We apply each method to the version of this system with analytically defined velocity, and using Eulerian velocity data defined over a finite number of time steps. We obtain this data through simulation of velocity adherent to the Navier–Stokes system of equations with (A.7) chosen as our initial condition. We simulate this data over a 2π–periodic domain $\Omega = [0, 2\pi) \times [0, 2\pi)$ using a uniform initial grid of 256×256 points, solve the Navier–Stokes equations using a third–order Runge–Kutta integration scheme in MATLAB with an integration time step of $\Delta t = 0.01$, and save velocity data after 1 unit of time for a total flow period of 10 units of time. By conducting this experiment, not only do we validate our LCS detection code for each method, but we validate the code used to simulate velocity data under the Navier–Stokes equations which we will use to simulate other turbulent flow systems for Lagrangian coherent structure analysis throughout this Thesis.

We begin with the FTLE for the steady Taylor–Green flow, computed from both analytic and numerically defined velocity and displayed in Figure A.17. From both of these Figures, the FTLE is low within each Taylor–Green vortex cell with the boundaries of these cells identifiable by the maximal ridges of the field. These maximal ridges do not touch or intersect, due to lower levels of the FTLE on either side of the central points of each maximal ridge, which happen to be stagnation points for the Taylor–Green vortex field. Both of these FTLE fields produce similar looking scalar fields with consistent values of the quantity, and the maximal absolute error between these quantities being of order 10^{-5} thereby providing us with a confidence that our FTLE algorithm works for

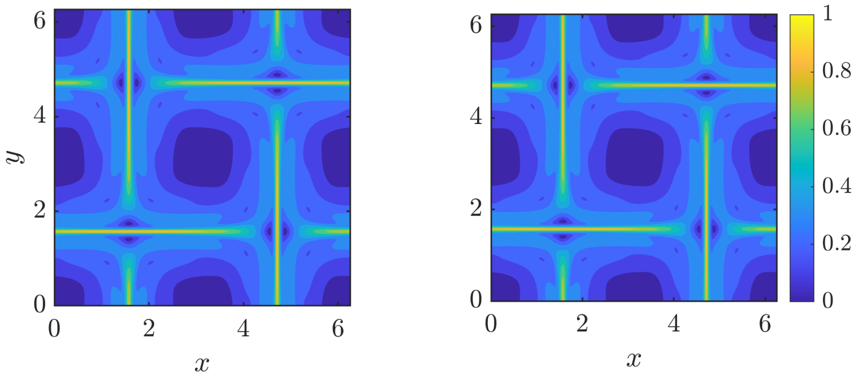

Figure A.17: The FTLE Φ_0^5 for the Taylor–Green vortex system computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between these quantities is 1.2307×10^{-5}.

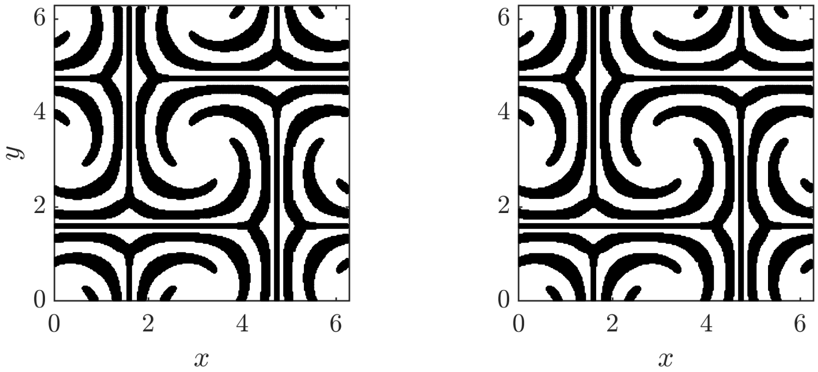

Figure A.18: The acceptable set of initial points $\mathcal{G}_0$ used to generate hyperbolic variational LCSs for the Taylor–Green vortex system computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right).

numerical data just as well as it does for analytically defined velocity.

We next turn our attention to the hyperbolic variational LCS algorithm, which we validate once again just by testing VLCS Algorithm A. As was the case in the previous Section, we are simply focused on ensuring that integration of the first eigenvector field $\boldsymbol{\xi}_1$ is executed appropriately as all of the VLCS detection algorithms rely on this. Figure A.18 shows the acceptable set $\mathcal{G}_0$ computed for VLCS Algorithm A, which from observation appears to be more or less identical when computed using analytical or numerical Taylor–Green velocity. The hyperbolic VLCSs for this system are shown in Figure A.19, computed from 500 randomly sampled points within $\mathcal{G}_0$ and given the allowance parameter $\mathcal{L}_f = 1$, and appear to not only coincide with the maximal ridges of the FTLE field, but are also comprised of curves which curl or spiral towards the centre of each vortex cell. The

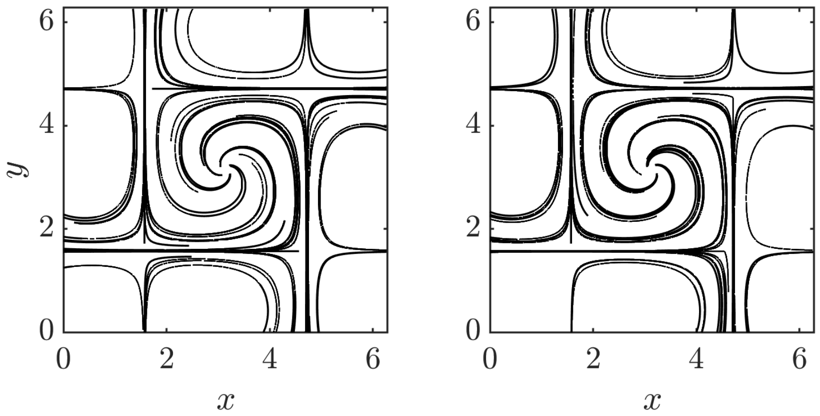

Figure A.19: Hyperbolic variational Lagrangian coherent structures for the Taylor–Green vortex system computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right).

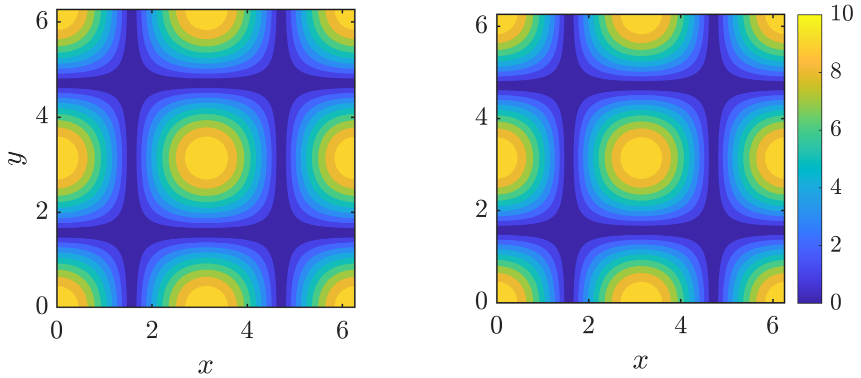

Figure A.20: The LAVD Θ_0^5 for the Taylor–Green vortex system computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between these quantities is 4.0×10^{-3}.

structures produced look more or less similar between the two plots, except in the plot produced from numerical velocity data some of the hyperbolic LCS curves which run along boundaries of the vortex cells bend to the side as they approach a corner of the cell rather than ending just before they reach these corners.

Next we consider the LAVD, with the scalar fields produced from this method shown in Figure A.20. This method has been able to identify each vortex cell much more clearly, with the rotational coherence of flow trajectories strong close to the centres of these cells and weak along the boundaries between these cells. Aesthetically, the LAVD computed from analytic velocity and numerically defined velocity is strongly similar although the largest maximal error recorded between the quantities is of order 10^{-3}. This error is

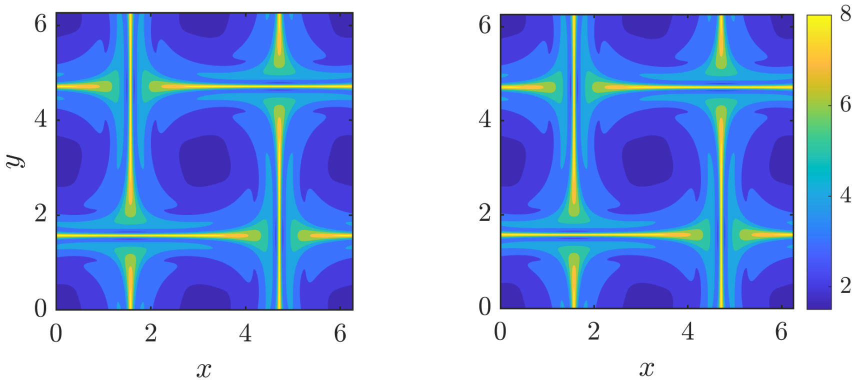

Figure A.21: The natural logarithm of the stochastic sensitivity for the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between the raw S^2 quantities is 2.8×10^{-1}.

likely owed to the fact that the vorticity of flow trajectories has to be interpolated along with the relevant velocity vectors to compute the LAVD, and is nonetheless not a greatly concerning value for the error.

We next examine the stochastic sensitivity of the Taylor–Green vortex flow from the results shown in Figure A.21. We plot the natural logarithm of the raw S^2 quantity in this Figure to enhance the visibility of coherent structures within this field. From both scalar fields, the stochastic sensitivity field looks similar to the FTLE fields seen in Figure A.17, with low stochastic sensitivity and more predictable flow observable within the centres of each vortex cell and maximal ridges of this field forming boundaries between these cells. However, unlike in the FTLE field where the maximal ridges approach the stagnation points of the field but never touch them, the maximal ridges in the $\ln\left(S^2\right)$ field "fan open" as they approach a stagnation point. There are also no observable "blobs" of low stochastic sensitivity measure on either side of the maximal points of this quantity like in the FTLE field, rather very thin ridges of low S^2 value surround each maximal ridge. The scalar fields produced look similar and correspond to similar S^2 values, however the largest absolute error recorded for the quantity is of order 10^{-1}. This is an expected observation due to the large gradients observable within the S^2 field, particularly where this quantity is extremal.

The optimal vector produced from the second right singular vector of the transfer operator also looks similar when computed from analytical or numerical Taylor–Green velocity as attested by the results shown in Figure A.22. To compute the transfer operator for this system, we divide Ω into 150×150 rectangular boxes of equal size and sample 1000 initial flow points within each box. From these results, the transfer operator method produces one coherent set from the central vortex cell and the other coherent set from the vortex cells on each corner of the domain. Each coherent set contains portions of the

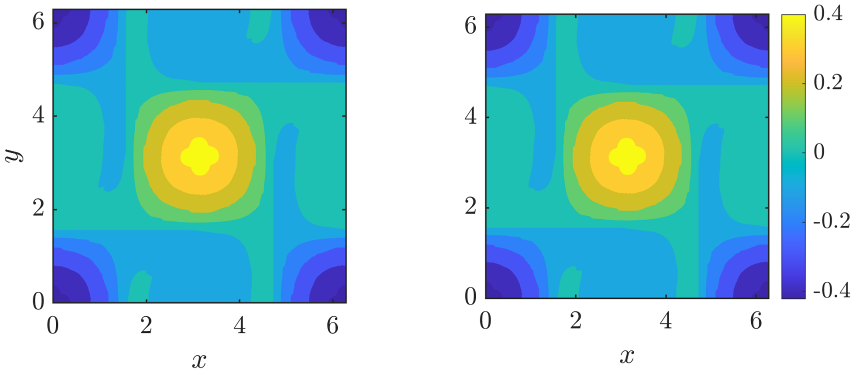

Figure A.22: Scalar fields for the second right singular vector of the transfer operator for the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between the singular vector entries is 4.5616×10^{-4}.

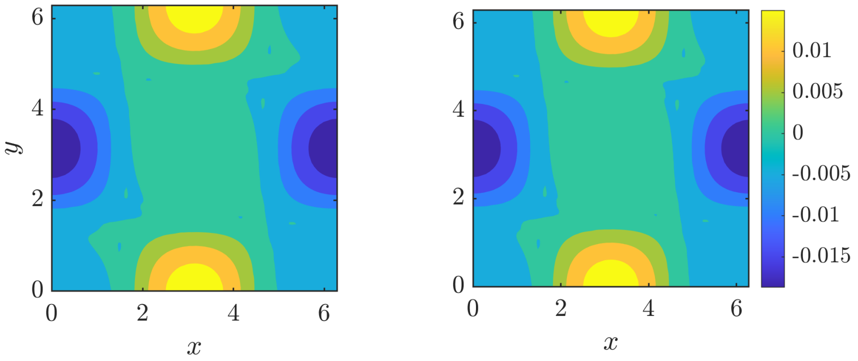

Figure A.23: Scalar fields for the second eigenvector of the dynamic Laplace operator for the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between the singular vector entries is 1.2423×10^{-4}.

other vortex cells where flow trajectories spill or leak between cells after beginning flow on the boundaries between vortex cells, for instance. Once again, the two fields are visually similar and the maximal error between coefficients of the optimal vectors computed is of order 10^{-4}, which is appropriate enough.

The coherent sets produced from the second eigenvector of the dynamic Laplace operator also look similar between analytic and numerically defined versions of the Taylor–Green vortex system, as observed from the results shown in Figure A.23. We use the same box configuration for Ω and the same number of points sampled within each box defined above for the transfer operator method. The two coherent sets obtained from this method differ considerably from those generated from the transfer operator method, with

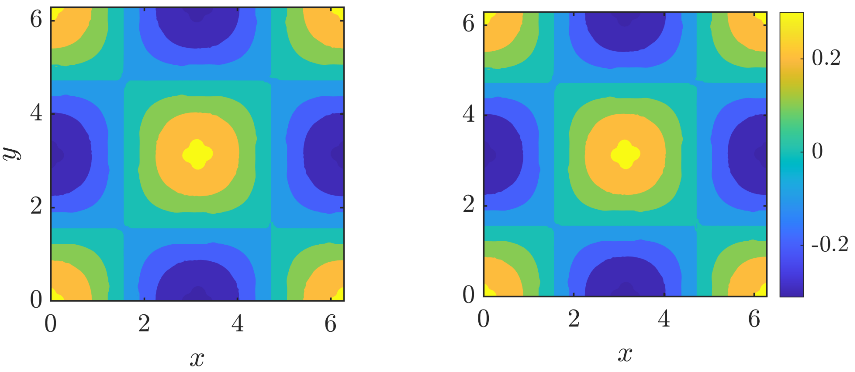

Figure A.24: Scalar fields for the fourth right singular vector of the transfer operator for the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between the singular vector entries is 5.4791×10^{-4}.

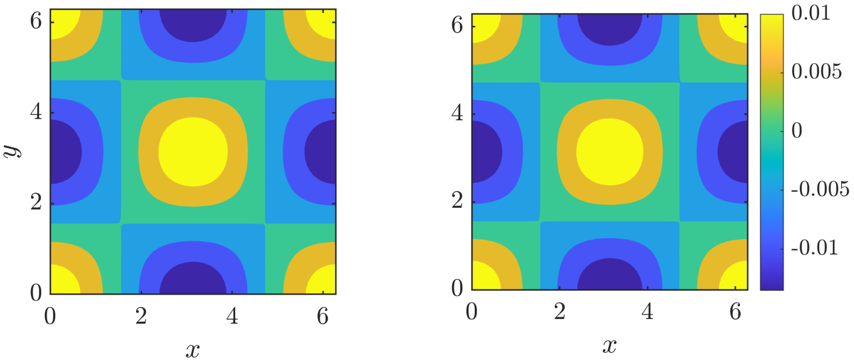

Figure A.25: Scalar fields for the fourth eigenvector of the dynamic Laplace operator for the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between the singular vector entries is 7.3726×10^{-6}.

the middle column of vortex cells identified as one coherent set, and the two remaining columns identified as the other coherent set. The two fields once again are similar aesthetically, while the largest absolute error recorded between the eigenvector entries is of order 10^{-4}, which is acceptable enough.

It is worth noting that the coherent sets produced from the fields shown in Figures A.22 and A.23 do not give the most comprehensive picture of the Taylor–Green vortex flow, though this could have something to do with the lack of a spectral gap, or large jump, between the second and third eigenvalues/singular values recorded for these operators (Froyland et al. 2010, Froyland & Junge 2018). The first six singular values recorded for the transfer operator $\mathbf{P}$ were 1, 0.9998, 0.9998, 0.9998, 0.9995 and 0.9995;

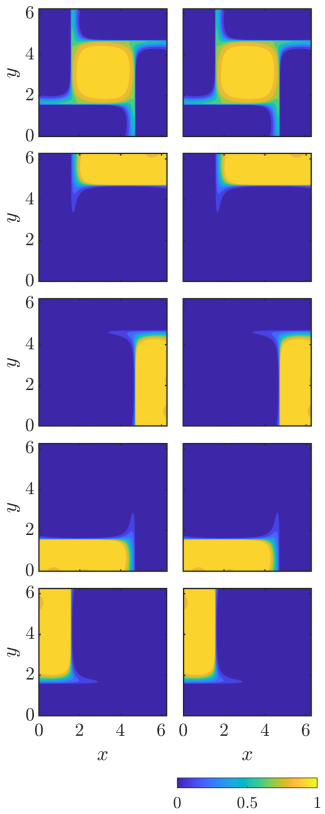

Figure A.26: Membership probability scalar fields for five FCM clusters within the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The suprema for the absolute errors between the membership probabilities of each cluster are (respectively): 8.7869×10^{-6}, 9.5237×10^{-6}, 9.5593×10^{-6}, 9.5743×10^{-6} and 9.5387×10^{-6}.

and the first six eigenvalues recorded for the dynamic Laplace operator were 0, -3.9443, -3.9476, -4.0712, -9.0016 and -9.6093. In both of these sets, a spectral gap is observable by the difference between the fourth and fifth value recorded (the jump is much larger for the dynamic Laplace operator eigenvalues), indicating that the fourth singular vector/eigenvector should be used to obtain coherent sets using this method. When we use the fourth singular vector of the transfer operator or the fourth eigenvector of the dynamic Laplace operator, we can clearly identify the coherent vortex cells for this system, as demonstrated by the results shown in Figures A.24 and A.25. The maximum absolute error for the singular vector entries for the transfer operator method computed between analytic and numerical velocity data is still of order 10^{-4}, while in the case of the fourth dynamic Laplace operator eigenvector the maximal error is now of order 10^{-6}, a considerable improvement on the second eigenvector.

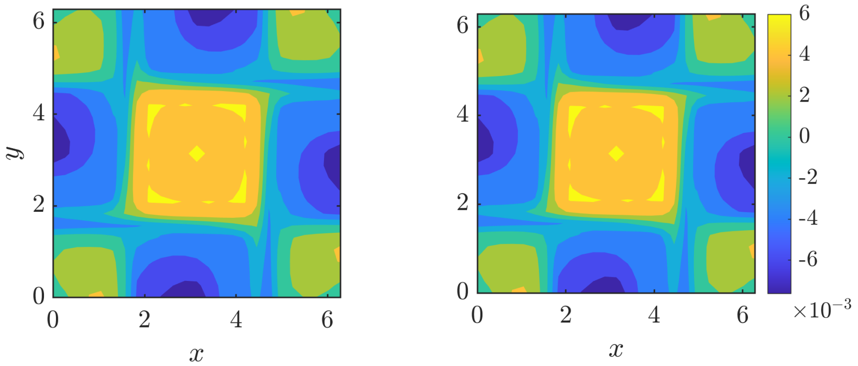

Figure A.27: Coherent structure colouring scalar fields for the Taylor–Green vortex system over the flow interval $[0, 5]$ computed using its analytic velocity expression (left) and simulated velocity data which represents the system (right). The supremum for the absolute error between the CSC coefficients is 1.53×10^{-2}.

We next turn our attention to the FCM clustering method, with the membership probability fields for five Taylor–Green vortex clusters detailed in Figure A.26. We choose to extract five clusters from 256×256 trajectories whose initial points are organised in a uniform grid, and choose the sharpness parameter m to have a value of 1.5. The algorithm identifies the central vortex cell as one coherent cluster, and four oblong shaped clusters which each comprise two of the vortex cells which surround the central cell. Once again, the membership probabilities produced from analytic and numerical velocity data appear to be similar, both visually and seeing as the maximal absolute error recorded between the analytic and numerical membership probabilities are of order 10^{-6} for each cluster.

Finally, we focus on the coherent structure colouring method, with the resultant scalar fields for this system featured in Figure A.27. We compute the CSC scalar field for the Taylor–Green vortex system using $25 \times 25 = 625$ flow trajectories, whose initial points are arranged in a uniform grid. From the scalar fields produced, the CSC coefficients with positive sign allow us to visualise the central vortex cell and all diagonally adjacent cells, while the negative CSC coefficients allow us to visualise the remaining cells which are horizontally or vertically adjacent to the central cell. The shapes and structures identified within each field look the same from the results obtained from analytic and numerical Taylor–Green velocity, although the largest absolute error recorded between the quantities is of order 10^{-2}. This is likely due to differences between the flow maps computed from analytic and numerical velocity at different time steps between t_0 and $t_0 + T$, which impact the elements of the CSC adjacency matrix $\mathbf{A}$ and all subsequent calculations. Nonetheless, we have achieved acceptable Lagrangian coherent structure results for the numerically defined Taylor–Green velocity flow from each method in comparison to the results produced from the analytically defined velocity, which not only verifies our Navier–Stokes solver for velocity simulation, but also assures that our LCS algorithms work just as well for numerically defined velocity systems as they do for analytic toy models.

Appendix B

Convergence of Noise Realisations

In this Appendix, we briefly provide evidence justifying our decision that a sample size of 100 was sufficient for producing means and standard deviations of stochastic realisations of various LCS quantities, as discussed in Chapter 4. We also explain why we have selected the particular histogram bin configurations for the VLCS histograms outlined in this Chapter.

B.1 Bickley Jet

Figure B.1 shows the scalar fields for the mean and double the standard deviation of the Bickley Jet FTLE, computed from 25, 100 and 200 stochastic realisations using the same parameters detailed in Chapter 4. As we can see from this Figure, convergence of the stochastic realisations of this quantity to a suitable mean has occurred relatively quickly, as there appears to be little (if any) observable difference between the fields computed for each sample size. We observe this for all of the other LCS detection methods as well, as we see from Figures B.3–B.12. In Figure B.2, we show different bin configurations for the VLCS likelihood histograms. If we choose the 90×30 case, we can see a jagged shape of the zonal jet, however we lose a considerable amount of detail into the structures close to and within the coherent Bickley jet vortices. In the 300×100 case, the resolution of the Figure becomes too refined and we are left with a large number of dark blue coloured bins representing areas of little or zero probability of being traversed by a hyperbolic LCS. If the separation radius ρ used in the VLCS–B algorithm were to be lowered, this would reduce the number of these boxes but it would also significantly increase the computational run time necessary to complete each stochastic realisation of this algorithm. We therefore settle for the sensible centre between these two cases, where we define 180×60 histogram bins of equal size.

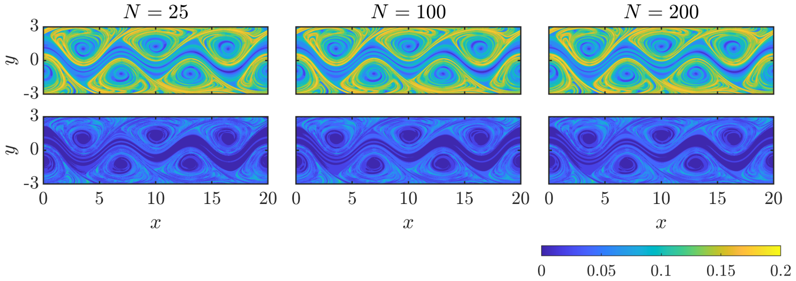

Figure B.1: The mean (top row) and twice the standard deviation (bottom row) of $N = 25$ (left column), 100 (centre column) and 200 (right column) stochastic realisations of the FTLE for the Bickley Jet flow from $t_0 = 0$ days to $t_0 + T = 40$ days.

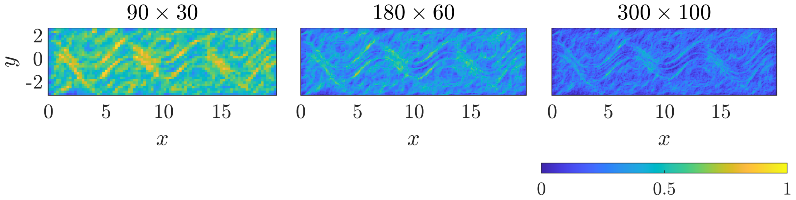

Figure B.2: The likelihood of a hyp–VLCS passing through one of $90 \times 30 = 2700$ (left), $180 \times 60 = 10800$ (centre) and $300 \times 100 = 30000$ (right) bins generated from 100 stochastic realisations of the variational LCS method for the Bickley Jet flow from $t_0 = 0$ days to $t_0 + T = 40$ days.

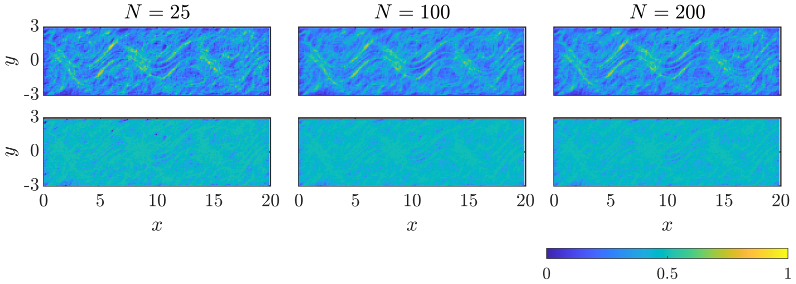

Figure B.3: The likelihood (top row) and twice the variation in this likelihood (bottom row) of a hyp–VLCS passing through one of $180 \times 60 = 10800$ bins generated from $N = 25$ (left column), 100 (centre column) and 200 (right column) stochastic realisations of the variational LCS method for the Bickley Jet flow from $t_0 = 0$ days to $t_0 + T = 40$ days.

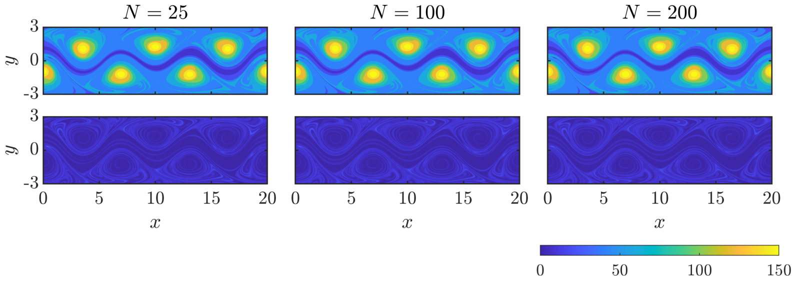

Figure B.4: Fields of the mean and twice the standard deviation of the Bickley Jet LAVD arranged in the same fashion as seen in Figure B.1.

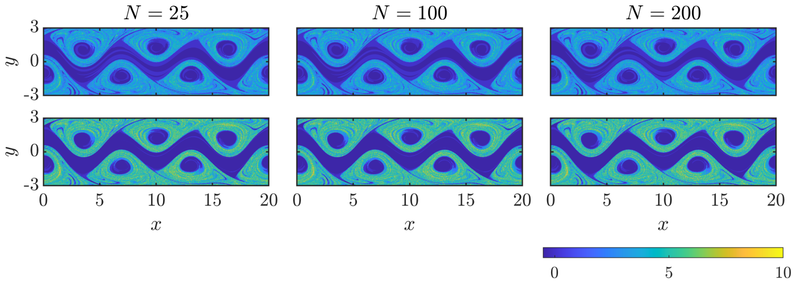

Figure B.5: Fields of the mean and twice the standard deviation of the quantity $\ln\left(\sqrt{hv_r S^2}\right)$ for the Bickley Jet flow arranged in the same fashion as seen in Figure B.1.

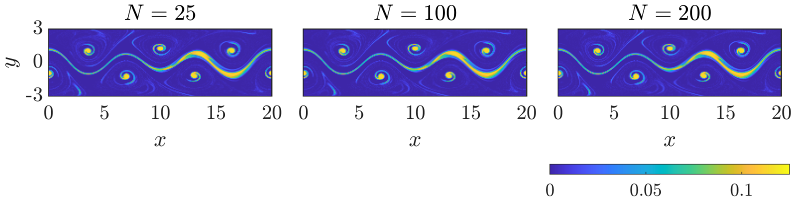

Figure B.6: Sample probability density function fields for the robust set $R(0.3, 5.6 \times 10^{-3}, 2.22 \times 10^{-2})$ of the Bickley Jet flow generated from 25 (left), 100 (centre) and 200 (right) realisations of the stochastic sensitivity method.

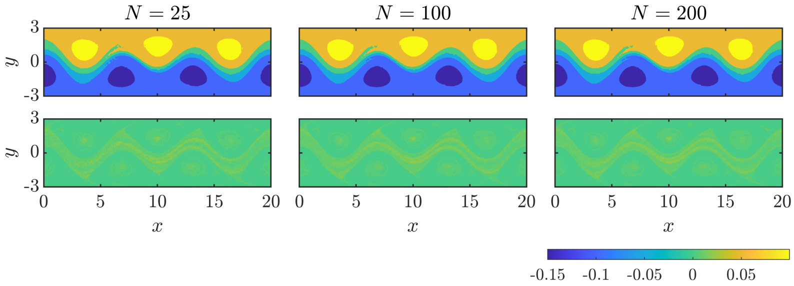

Figure B.7: Fields of the mean and twice the standard deviation of the transfer operator optimal vector for the Bickley Jet flow arranged in the same fashion as seen in Figure B.1.

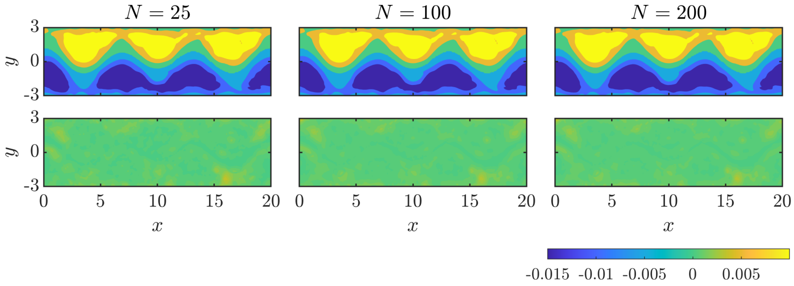

Figure B.8: Fields of the mean and twice the standard deviation of the dynamic Laplace operator second eigenvector for the Bickley Jet flow arranged in the same fashion as seen in Figure B.1.

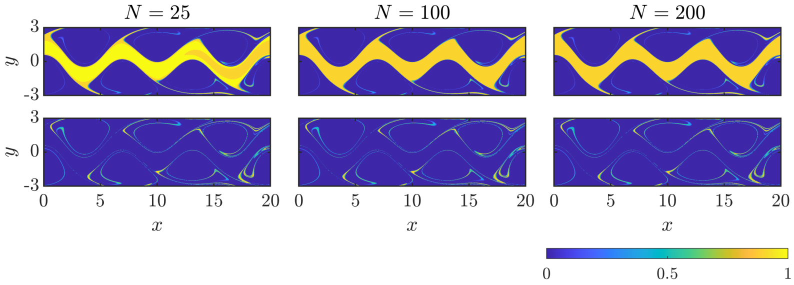

Figure B.9: Fields of the mean and twice the standard deviation of the membership probability for the first FCM cluster of the Bickley Jet flow arranged in the same fashion as seen in Figure B.1.

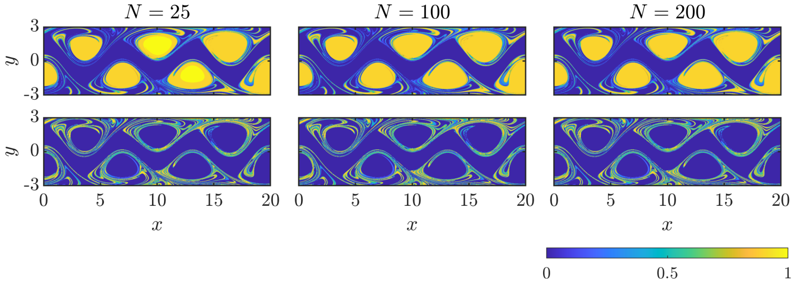

Figure B.10: Fields of the mean and twice the standard deviation of the membership probability for the second FCM cluster of the Bickley Jet flow arranged in the same fashion as seen in Figure B.1.

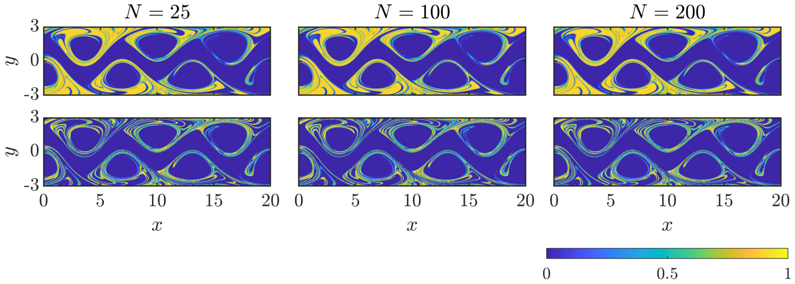

Figure B.11: Fields of the mean and twice the standard deviation of the membership probability for the third FCM cluster of the Bickley Jet flow arranged in the same fashion as seen in Figure B.1.

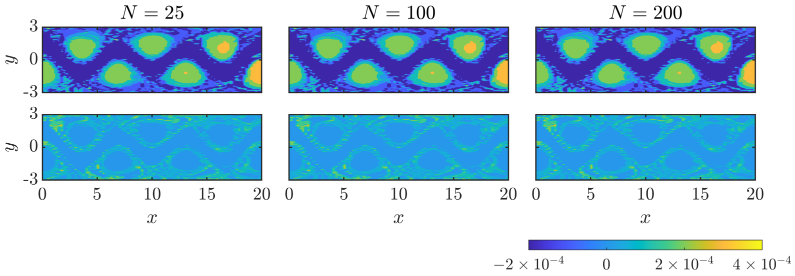

Figure B.12: Fields of the mean and twice the standard deviation of the Bickley Jet CSC arranged in the same fashion as seen in Figure B.1.

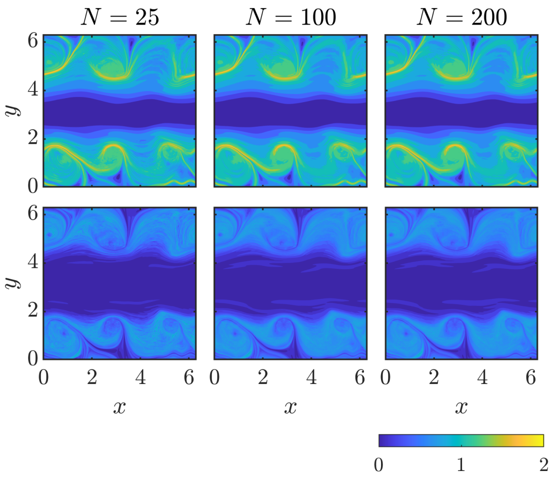

Figure B.13: The mean (top row) and twice the standard deviation (bottom row) of $N = 25$ (left column), 100 (centre column) and 200 (right column) stochastic realisations of the FTLE for the KH2 flow from $t_0 = 12$ seconds to $t_0 + T = 15$ seconds.

B.2 Kelvin–Helmholtz (Version 2)

We next consider the KH2 flow system, with Figure B.13 showing the scalar fields for the mean and double the standard deviation of the FTLE for this flow computed from 25, 100 and 200 stochastic realisations of this quantity using the same parameters detailed in Chapter 4. Like with the Bickley Jet flow, convergence of these realisations happens quickly with not too many discernible differences observable between the resultant fields, aside from some very minor differences observable in both the mean and standard deviation fields between the $N = 25$ and $N = 100$ cases of these fields. This is a common observation across all our LCS detection methods, as the results shown in Figures B.15–B.25 attest. In Figure B.14, we show how the VLCS likelihood field changes in form as the histogram bin configuration changes. Using similar reasoning as that applied to the Bickley Jet flow, we choose the 50 × 50 bins case as it gives sufficient detail into the formation of the likelihood field without leaving a large number of bins corresponding to zero or low probability of being traversed through by a hyperbolic VLCS.

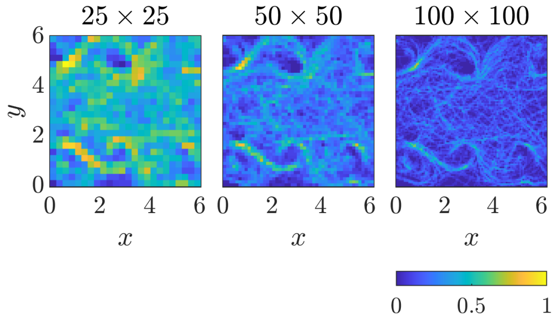

Figure B.14: The likelihood of a hyp–VLCS passing through one of $25 \times 25 = 625$ (left), $50 \times 50 = 2500$ (centre) and $100 \times 100 = 10000$ (right) bins generated from 100 stochastic realisations of the variational LCS method for the KH2 flow from $t_0 = 12$ seconds to $t_0 + T = 15$ seconds.

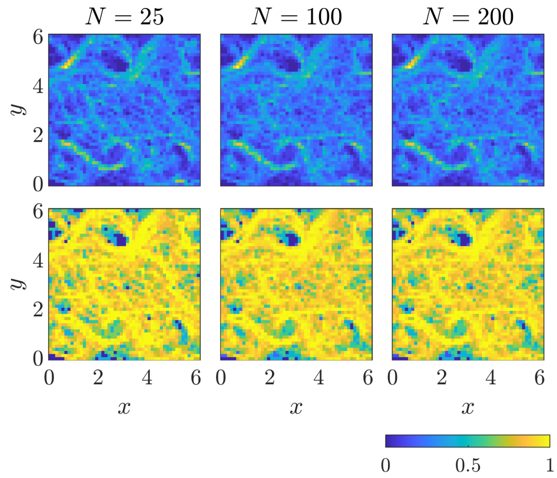

Figure B.15: The likelihood (top row) and twice the variation in this likelihood (bottom row) of a hyp–VLCS passing through one of $50 \times 50 = 2500$ bins generated from $N = 25$ (left column), 100 (centre column) and 200 (right column) stochastic realisations of the variational LCS method for the KH2 flow from $t_0 = 12$ seconds to $t_0 + T = 15$ seconds.

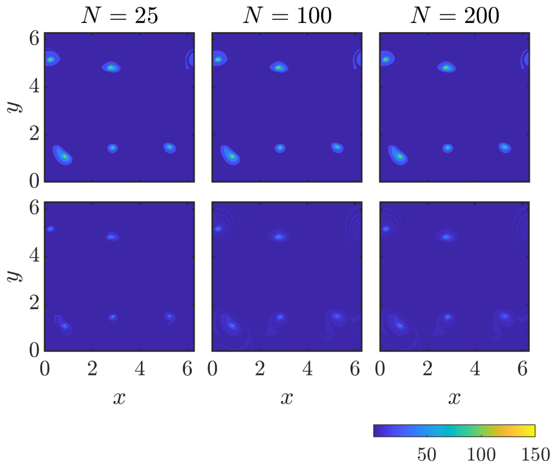

Figure B.16: Fields of the mean and twice the standard deviation of the KH2 LAVD arranged in the same fashion as seen in Figure B.13.

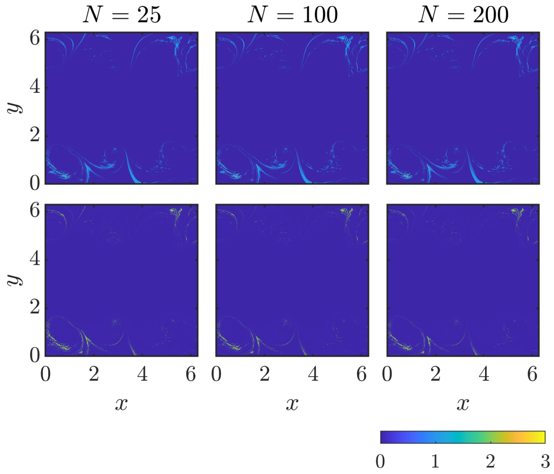

Figure B.17: Fields of the mean and twice the standard deviation of the quantity $\ln\left(\sqrt{hv_r S^2}\right)$ for the KH2 flow arranged in the same fashion as seen in Figure B.13.

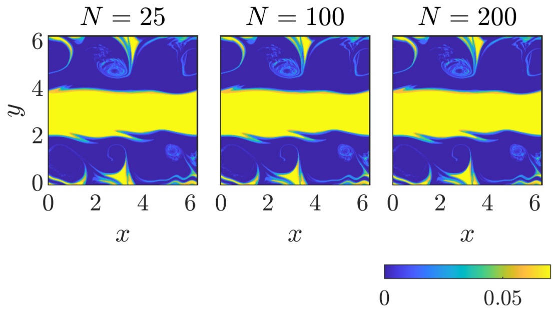

Figure B.18: Sample probability density function fields for the robust set $R(0.25, 1.63 \times 10^{-2}, 6.1 \times 10^{-3})$ of the KH2 flow generated from 25 (left), 100 (centre) and 200 (right) realisations of the stochastic sensitivity method.

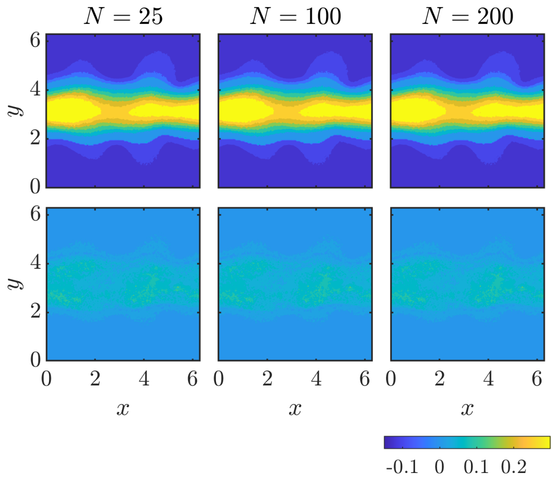

Figure B.19: Fields of the mean and twice the standard deviation of the transfer operator optimal vector for the KH2 flow arranged in the same fashion as seen in Figure B.13.

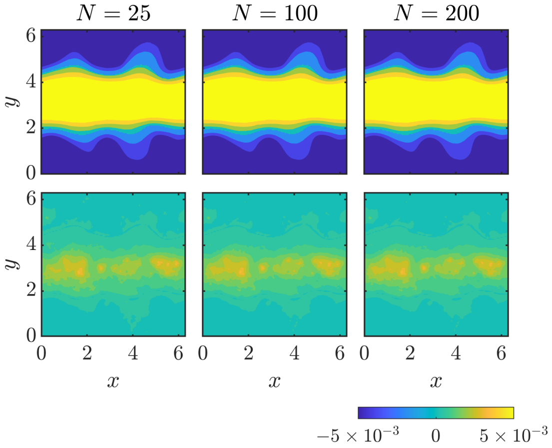

Figure B.20: Fields of the mean and twice the standard deviation of the dynamic Laplace operator second eigenvector for the KH2 flow arranged in the same fashion as seen in Figure B.13.

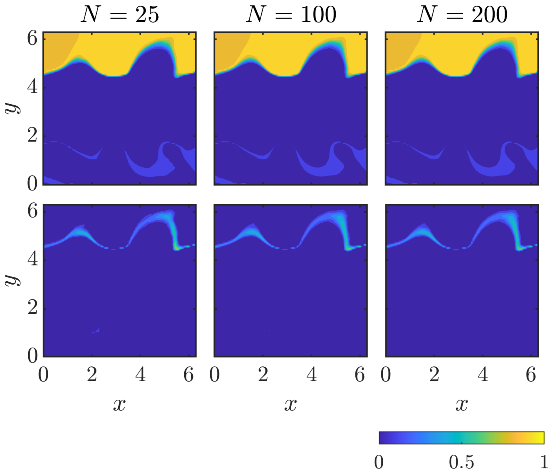

Figure B.21: Fields of the mean and twice the standard deviation of the membership probability for the first FCM cluster of the KH2 flow arranged in the same fashion as seen in Figure B.13.

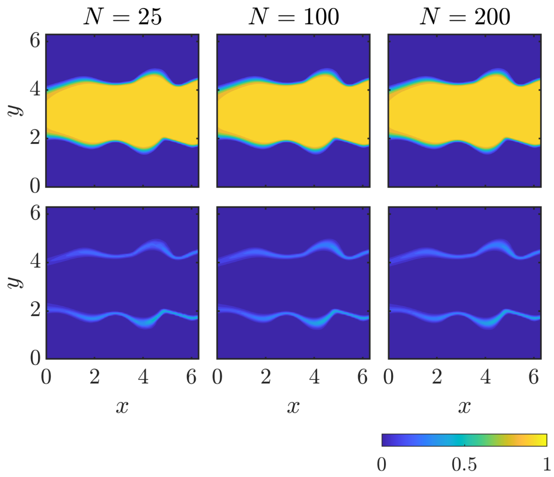

Figure B.22: Fields of the mean and twice the standard deviation of the membership probability for the second FCM cluster of the KH2 flow arranged in the same fashion as seen in Figure B.13.

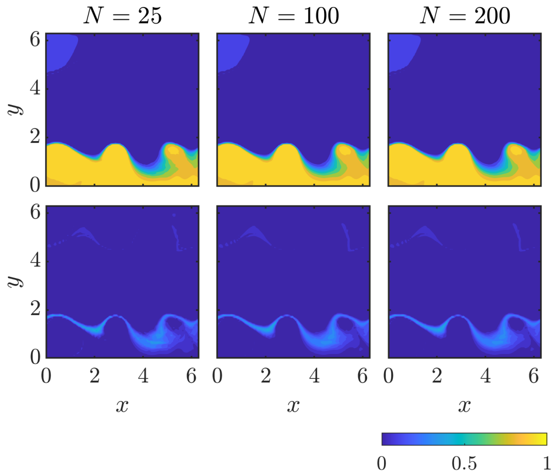

Figure B.23: Fields of the mean and twice the standard deviation of the membership probability for the third FCM cluster of the KH2 flow arranged in the same fashion as seen in Figure B.13.

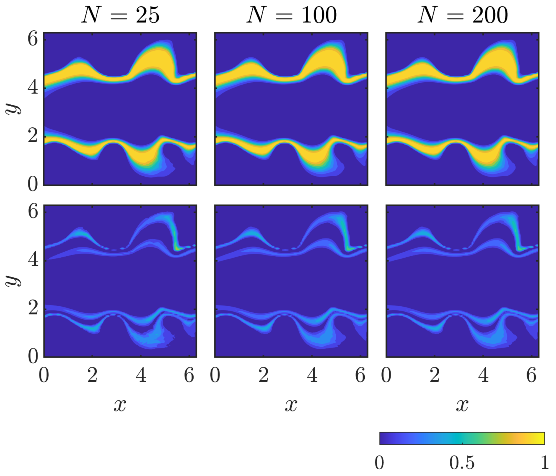

Figure B.24: Fields of the mean and twice the standard deviation of the membership probability for the fourth FCM cluster of the KH2 flow arranged in the same fashion as seen in Figure B.13.

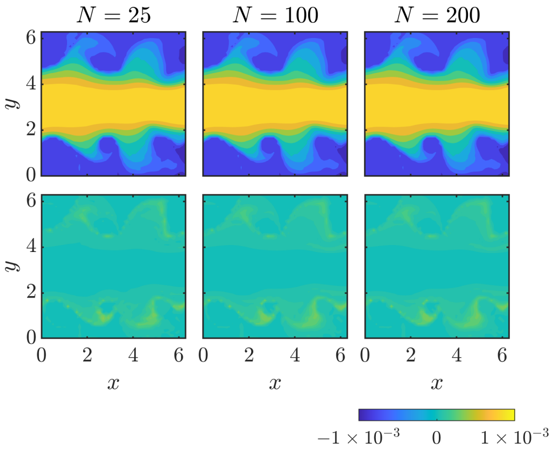

Figure B.25: Fields of the mean and twice the standard deviation of the KH2 CSC arranged in the same fashion as seen in Figure B.13.

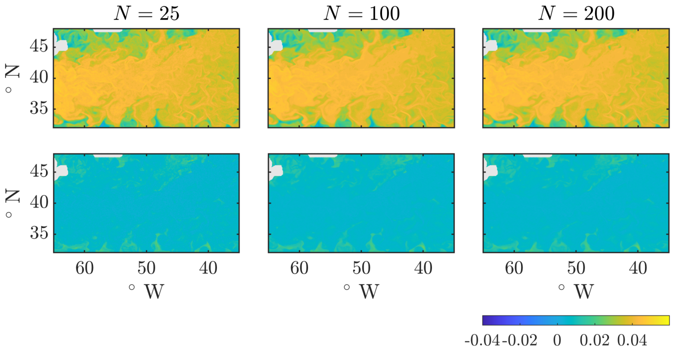

Figure B.26: The mean (top row) and the standard deviation (bottom row) of $N = 25$ (left column), 100 (centre column) and 200 (right column) stochastic realisations of the FTLE for the Gulf Stream flow from t_0 = midnight, January 15th 2015 to $t_0 + T$ = midnight, April 15th 2015 (UTC).

B.3 Gulf Stream

Finally we consider the Gulf Stream flow, starting with the mean and standard deviation fields of the FTLE for this system generated from sample sizes of 25, 100 and 200 stochastic realisations of this quantity on display in Figure B.26. Once again, this verifies that 100 is a sufficient number of stochastic realisations for this method as there are little observable differences within the scalar fields for each of these sample sizes, with some minor differences observable between the $N = 25$ and $N = 100$ cases. This fast convergence of the stochastic realisations is again observable across all of our detection methods, as we see from the results shown in Figures B.28–B.39. In Figure B.27, we display different likelihood fields for the VLCS method corresponding to different histogram bin configurations for the domain Ω. We again desire a bin configuration that provides greater clarity of the shapes observable within the likelihood field, but do not want to reduce the size of the bins to such a low point that we are left with a large number of bins left uncharted by the VLCSs, so we settle for the 240 × 160 bin configuration.

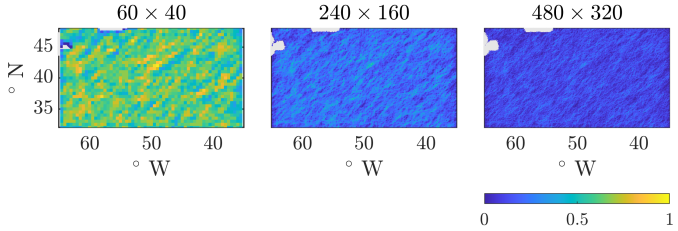

Figure B.27: The likelihood of a hyp–VLCS passing through one of $60 \times 40 = 2400$ (left), $240 \times 160 = 38400$ (centre) and $480 \times 320 = 153600$ (right) bins generated from 100 stochastic realisations of the variational LCS method for the Gulf Stream flow from $t_0 =$ midnight, January 15th 2015 to $t_0 + T =$ midnight, April 15th 2015 (UTC).

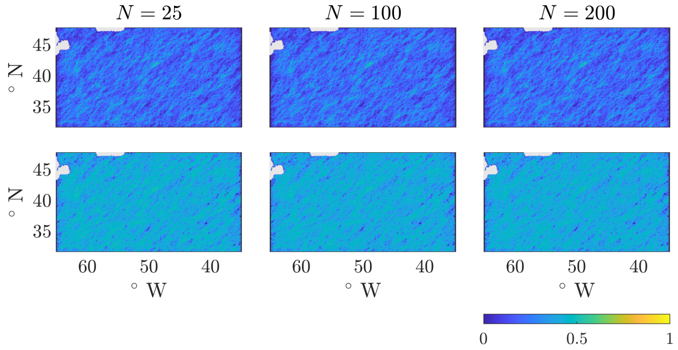

Figure B.28: The likelihood (top row) and the variation in this likelihood (bottom row) of a hyp–VLCS passing through one of $240 \times 160 = 38400$ bins generated from $N = 25$ (left column), 100 (centre column) and 200 (right column) stochastic realisations of the variational LCS method for the Gulf Stream flow from $t_0 =$ midnight, January 15th 2015 to $t_0 + T =$ midnight, April 15th 2015 (UTC).

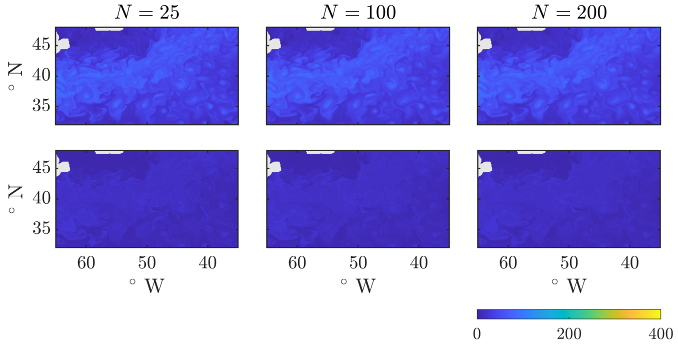

Figure B.29: Fields of the mean and the standard deviation of the Gulf Stream LAVD arranged in the same fashion as seen in Figure B.26.

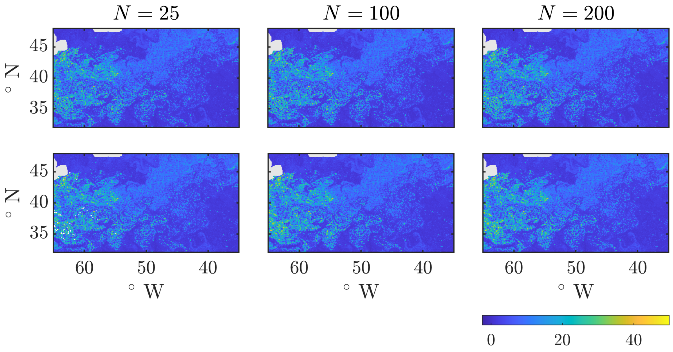

Figure B.30: Fields of the mean and the standard deviation of the quantity $\ln\left(\sqrt{hv_rS^2}\right)$ for the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

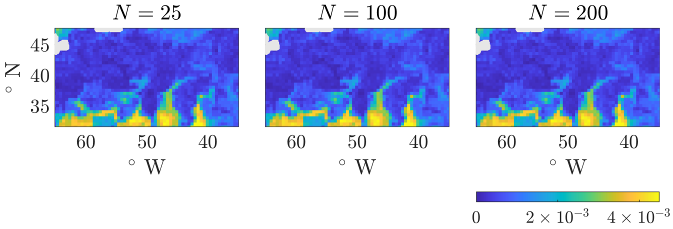

Figure B.31: Sample probability density function fields for the robust set $R(6, 1.27 \times 10^{-2}, 0.125)$ of the Gulf Stream flow generated from 25 (left), 100 (centre) and 200 (right) realisations of the stochastic sensitivity method.

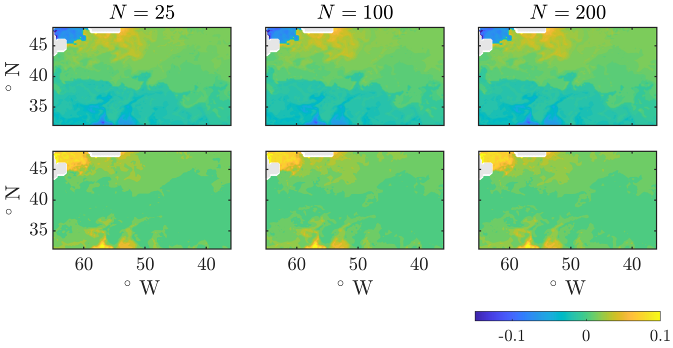

Figure B.32: Fields of the mean and the standard deviation of the transfer operator optimal vector for the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

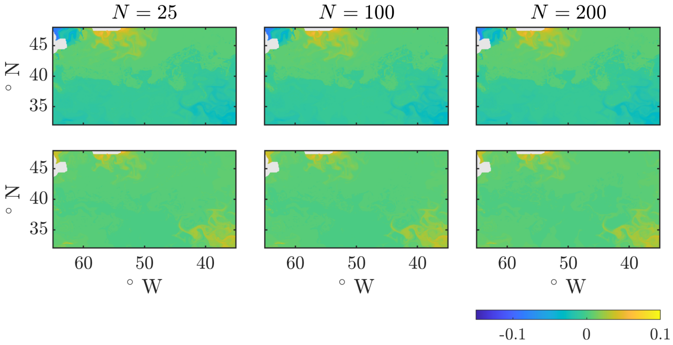

Figure B.33: Fields of the mean and the standard deviation of the dynamic Laplace operator third eigenvector for the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

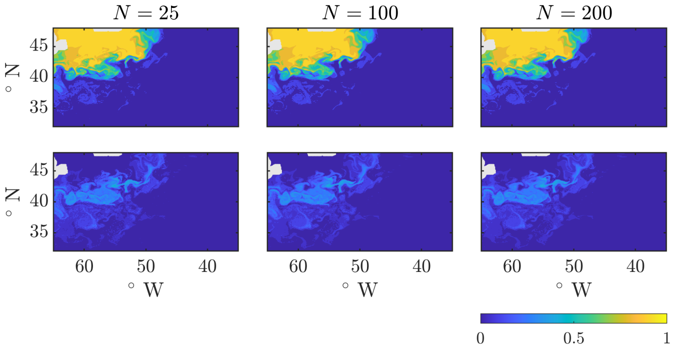

Figure B.34: Fields of the mean and the standard deviation of the membership probability for the first FCM cluster of the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

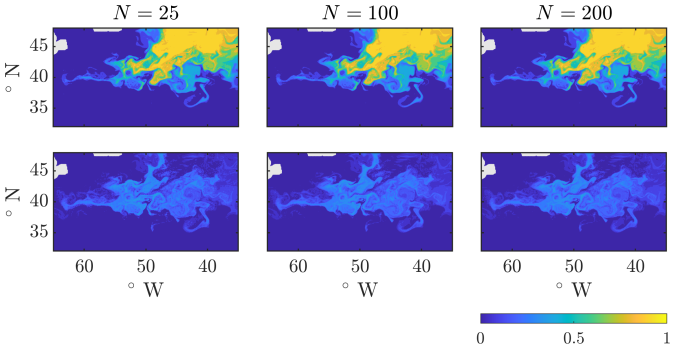

Figure B.35: Fields of the mean and the standard deviation of the membership probability for the second FCM cluster of the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

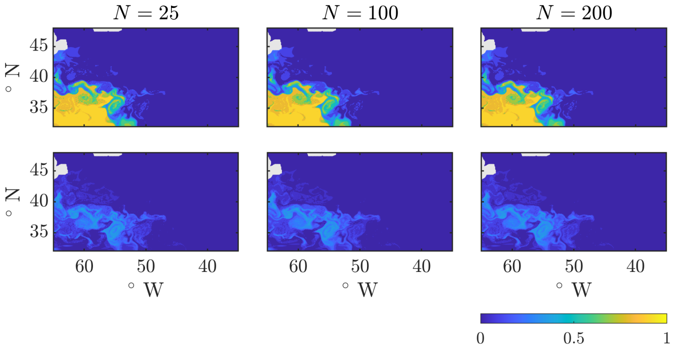

Figure B.36: Fields of the mean and the standard deviation of the membership probability for the third FCM cluster of the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

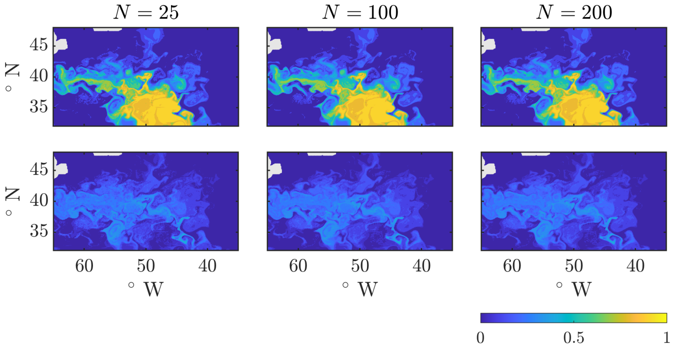

Figure B.37: Fields of the mean and the standard deviation of the membership probability for the fourth FCM cluster of the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

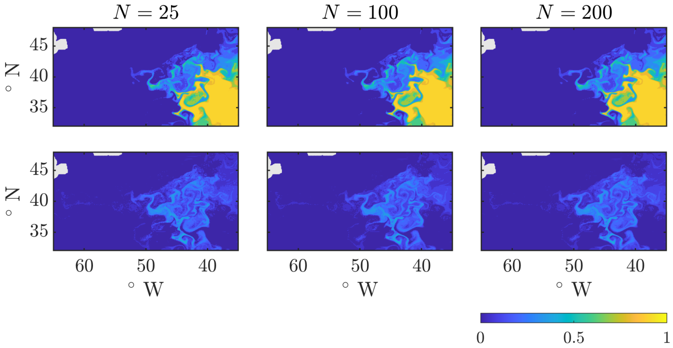

Figure B.38: Fields of the mean and the standard deviation of the membership probability for the fifth FCM cluster of the Gulf Stream flow arranged in the same fashion as seen in Figure B.26.

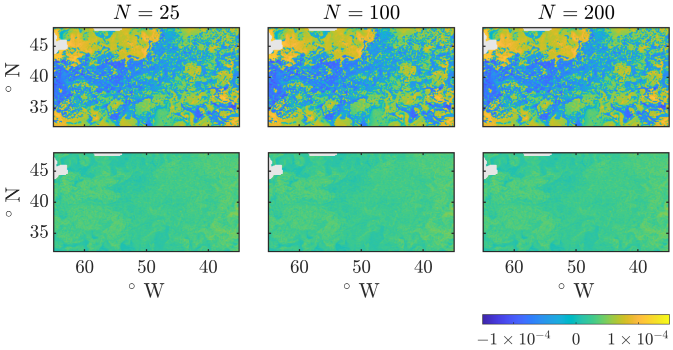

Figure B.39: Fields of the mean and the standard deviation of the Gulf Stream CSC arranged in the same fashion as seen in Figure B.26.

Bibliography

Allshouse, M. R. & Peacock, T. (2015), 'Lagrangian based methods for coherent structure detection', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**(9), 097617.

Allshouse, M. R. & Thiffeault, J.-L. (2012), 'Detecting coherent structures using braids', *Physica D: Nonlinear Phenomena* **241**(2), 95–105.

Balasuriya, S. (2020*a*), 'Stochastic Sensitivity: A Computable Lagrangian Uncertainty Measure for Unsteady Flows', *SIAM Review* **62**(4), 781–816.

Balasuriya, S. (2020*b*), 'Uncertainty in finite–time Lyapunov exponent computations', *Journal of Computational Dynamics* **7**(2), 313–337.

Balasuriya, S. & Bollt, E. M. (2021), 'Globally optimal stretching foliations of dynamical systems reveal the organizing skeleton of intensive instabilities', **Submitted**.

Balasuriya, S., Kalampattel, R. & Ouellette, N. T. (2016), 'Hyperbolic neighbourhoods as organizers of finite–time exponential stretching', *Journal of Fluid Mechanics* **807**(1), 509–545.

Balasuriya, S., Ouellette, N. T. & Rypina, I. I. (2018), 'Generalized Lagrangian coherent structures', *Physica D: Nonlinear Phenomena* **372**(1), 31–51.

Balibrea-Iniesta, F., Lopesino, C., Wiggins, S. & Mancho, A. M. (2016), 'Lagrangian Descriptors for Stochastic Differential Equations: A Tool for Revealing the Phase Portrait of Stochastic Dynamical Systems', *International Journal of Bifurcation and Chaos* **26**(13), 1630036.

Benaichouche, A. N., Oulhadj, H. & Siarry, P. (2013), 'Improved spatial fuzzy c–means clustering for image segmentation using PSO initialization, Mahalanobis distance and post–segmentation correction', *Digital Signal Processing* **23**(1), 1390–1400.

Bezdek, J. C., Ehrlich, R. & Full, W. (1984), 'FCM: The fuzzy c–means clustering algorithm', *Computers and Geosciences* **10**(2–3), 191–203.

Bezdek, J. C., Hathaway, R. J., Sabin, M. J. & Tucker, W. T. (1987), 'Convergence theory for fuzzy c–means: Counterexamples and repairs', *IEEE Transactions on Systems, Man, and Cybernetics* **17**(5), 873–877.

Blazevski, D. & Haller, G. (2014), 'Hyperbolic and elliptic transport barriers in three–dimensional unsteady flows', *Physica D: Nonlinear Phenomena* **273–274**(1), 46–62.

Bollt, E. M., Luttman, A., Kramer, S. & Basnayake, R. (2012), 'Hyperbolic and elliptic transport barriers in three–dimensional unsteady flows', *International Journal of Bifurcation and Chaos* **22**(3), 1230012.

Byrne, D. A., Gordon, A. L. & Haxby, W. F. (1995), 'Agulhas Eddies: A Synoptic View Using Geosat ERM Data', *Journal of Physical Oceanography* **25**(5), 902–917.

Crowdy, D. G. (2004), 'Stuart vortices on a sphere', *Journal of Fluid Mechanics* **498**(1), 381–402.

del Castillo-Negrete, D. & Morrison, P. J. (1993), 'Chaotic transport by Rossby waves in shear flow', *Physics of Fluids A: Fluid Dynamics* **5**(1), 948–965.

d'Ovidio, F., Fernández, V., Hernández-García, E. & López, C. (2004), 'Mixing structures in the Mediterranean Sea from finite–size Lyapunov exponents', *Geophysical Research Letters* **31**(17), L17203.

Farazmand, M., Blazevski, D. & Haller, G. (2014), 'Shearless transport barriers in unsteady two–dimensional flows and maps', *Physica D: Nonlinear Phenomena* **278–279**(1), 44–57.

Farazmand, M. & Haller, G. (2012), 'Computing Lagrangian coherent structures from their variational theory', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22**(1), 013128.

Froyland, G. (2005), 'Statistically optimal almost–invariant sets', *Physica D: Nonlinear Phenomena* **200**(1), 205–219.

Froyland, G. (2013), 'An analytic framework for identifying finite–time coherent sets in time–dependent dynamical systems', *Physica D: Nonlinear Phenomena* **250**(1), 1–19.

Froyland, G. (2015), 'Dynamic isoperimetry and the geometry of Lagrangian coherent structures', *Nonlinearity* **28**(1), 3587–3622.

Froyland, G., Horenkamp, C., Rossi, V., Santitissadeekorn, N. & Sen Gupta, A. (2012), 'Three–dimensional characterization and tracking of an Agulhas Ring', *Ocean Modelling* **52–53**(1), 69–75.

Froyland, G., Horenkamp, C., Rossi, V. & van Sebille, E. (2015), 'Studying an Agulhas ring's long–term pathway and decay with finite–time coherent sets', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**(8), 083119.

Froyland, G. & Junge, O. (2018), 'Robust FEM–Based Extraction of Finite-Time Coherent Sets Using Scattered, Sparse, and Incomplete Trajectories', *SIAM Journal of Applied Dynamical Systems* **17**(2), 1891–1924.

Froyland, G. & Kwok, E. (2017), 'A dynamic Laplacian for identifying Lagrangian coherent structures on weighted Riemannian manifolds', *Journal of Nonlinear Science* **30**(1), 1889–1971.

Froyland, G. & Padberg-Gehle, K. (2012), 'Finite–time entropy: A probabilistic approach for measuring nonlinear stretching', *Physica D: Nonlinear Phenomena* **241**(19), 1612–1628.

Froyland, G. & Padberg-Gehle, K. (2015), 'A rough–and–ready cluster–based approach for extracting finite–time coherent sets from sparse and incomplete trajectory data', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**(1), 087406.

Froyland, G., Rock, C. P. & Sakellariou, K. (2019), 'Sparse eigenbasis approximation: Multiple feature extraction across spatiotemporal scales with application to coherent set identification', *Communications in Nonlinear Science and Numerical Simulation* **77**(1), 81–107.

Froyland, G., Santitissadeekorn, N. & Monahan, A. (2010), 'Transport in time–dependent dynamical systems: Finite–time coherent sets', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **20**(1), 043116.

Froyland, G., Stuart, R. M. & van Sebille, E. (2014), 'How well–connected is the surface of the global ocean?', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **24**(1), 033126.

Garaboa-Paz, D., Eiras-Barca, J. & Pérez–Muñuzuri, V. (2017), 'Climatology of Lyapunov exponents: the link between atmospheric rivers and large–scale mixing variability', *Earth System Dynamics* **8**(1), 865–873.

Garcia Llamas, C., Spille, C., Kastens, S., Garaboa-Paz, D., Schlüter, M. & von Kameke, A. (2020), 'Potential of Lagrangian Analysis Methods in the Study of Chemical Reactors', *Chemie–Ingenieur–Technik* **92**(5), 540–553.

Guo, H., He, W., Peterka, T., Shen, H.-W., Collis, S. & Helmus, J. (2016), 'Finite–Time Lyapunov Exponents and Lagrangian Coherent Structures in Uncertain Unsteady Flows', *IEEE Transactions on Visualization and Computer Graphics* **22**(6), 1672–1682.

Hadjighasem, A., Farazmand, M., Blazevski, D., Froyland, G. & Haller, G. (2017), 'A critical comparison of Lagrangian methods for coherent structure detection', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **27**(1), 053104.

Hadjighasem, A. & Haller, G. (2016), 'Geodesic Transport Barriers in Jupiter's Atmosphere: A Video–Based Analysis', *SIAM Review* **58**(1), 69–89.

Haller, G. (2001), 'Distinguished material surfaces and coherent structures in three–dimensional fluid flows', *Physica D: Nonlinear Phenomena* **149**(4), 248–277.

Haller, G. (2002), 'Lagrangian coherent structures from approximate velocity data', *Physics of Fluids* **14**(6), 1851–1861.

Haller, G. (2011), 'A variational theory of hyperbolic Lagrangian coherent structures', *Physica D: Nonlinear Phenomena* **240**(1), 574–598.

Haller, G. (2016), 'Dynamic rotation and stretch tensors from a dynamic polar decomposition', *Journal of the Mechanics and Physics of Solids* **86**(1), 70–93.

Haller, G. & Beron-Vera, F. J. (2012), 'Geodesic theory of transport barriers in two–dimensional flows', *Physica D: Nonlinear Phenomena* **241**(1), 1680–1702.

Haller, G. & Beron-Vera, F. J. (2013), 'Coherent Lagrangian vortices: the black holes of turbulence', *Journal of Fluid Mechanics* **731**(1), R4.

Haller, G., Hadjighasem, A., Farazmand, M. & Huhn, F. (2016), 'Defining coherent vortices objectively from the vorticity', *Journal of Fluid Mechanics* **795**(1), 136–173.

Haller, G. & Yuan, G. (2000), 'Lagrangian coherent structures and mixing in two–dimensional turbulence', *Physica D: Nonlinear Phenomena* **147**(3–4), 352–370.

Haza, A. C., Ozgokmen, T. M., Griffa, A., Garraffo, Z. D. & Piterbarg, L. (2012), 'Parameterization of particle transport at submesoscales in the Gulf Stream region using Lagrangian subgridscale models', *Ocean Modelling* **42**(1), 31–49.

Joseph, B. & Legras, B. (2002), 'Relation between Kinematic Boundaries, Stirring and Barriers for the Antarctic Polar Vortex', *Journal of the Atmospheric Sciences* **59**(7), 1198–1212.

Kang, D. & Curchitser, E. N. (2013), 'Gulf Stream eddy characteristics in a high–resolution ocean model', *Journal of Geophysical Research: Oceans* **118**(1), 4474–4487.

Karrasch, D., Huhn, F. & Haller, G. (2014), 'Automated detection of coherent Lagrangian vortices in two–dimensional unsteady flows', *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **471**(2173), 1–17.

Kent, S. (2008), 'Lagrangian Coherent Structures: Generalizing Stable and Unstable Manifolds to Non-Autonomous Dynamical Systems', *University of Arizona Program in Applied Mathematics* **NA**(NA), 1–15.

Leclair, M., Lowe, R., Zhang, Z., Ivey, G. & Peacock, T. (2020), 'Uncovering Fine–Scale Wave–Driven Transport Features in a Fringing Coral Reef System via Lagrangian Coherent Structures', *Fluids* **5**(1), 190.

Lee, H. G. & Kim, J. (2015), 'Two–dimensional Kelvin–Helmholtz instabilities of multi–component fluids', *European Journal of Mechanics B/Fluids* **49**(1), 77–88.

Lekien, F. & Ross, S. D. (2010), 'The computation of finite–time Lyapunov exponents on unstructured meshes and for non–Euclidean manifolds', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **20**(1), 017505.

Lesieur, M., Staquet, C., Le Roy, P. & Comte, P. (1988), 'The mixing layer and its coherence examined from the point of view of two–dimensional turbulence', *Journal of Fluid Mechanics* **192**(1), 511–534.

Lin, H., Xiang, Y., Qin, S., Xu, H. & Liu, H. (2020), 'Lagrangian analysis of the fluid transport induced by the interaction of two co–axial co–rotating vortex rings', *Journal of Hydrodynamics* **32**(6), 1080–1090.

Liu, Y., Wilson, C., Green, M. A. & Hughes, C. W. (2018), 'Gulf stream transport and mixing processes via coherent structure dynamics', *Journal of Geophysical Research: Oceans* **123**(1), 3014–3037.

Ma, T. & Bollt, E. M. (2013), 'Relatively Coherent Sets As a Hierarchical Partition Method', *International Journal of Bifurcation and Chaos* **23**(7), 1330026.

Mancho, A. M., Wiggins, S., Curbelo, J. & Mendoza, C. (2013), 'Lagrangian descriptors: A method for revealing phase space structures of general time dependent dynamical systems', *Communications in Nonlinear Science and Numerical Simulation* **18**(12), 3530–3557.

Martins, F. A. C., Sciacchitano, A. & Rival, D. E. (2021), 'Detection of vortical structures in sparse Lagrangian data using coherent–structure colouring', *Experiments in Fluids* **62**(4), 69.

Mattner, T. W. (2011), 'Large–eddy simulations of turbulent mixing layers using the stretched-vortex model', *Journal of Fluid Mechanics* **671**(1), 507–534.

Metcalfe, R. W., Orszag, S. A., Brachet, M. E., Menon, S. & Riley, J. J. (1987), 'Secondary instability of a temporally growing mixing layer', *Journal of Fluid Mechanics* **184**(1), 207–243.

Mezić, I., Loire, S., Fonoberov, V. A. & Hogan, P. (2010), 'A New Mixing Diagnostic and Gulf Oil Spill Movement', *Science* **330**(6003), 486–489.

Nolan, P. J., Foroutan, H. & Ross, S. D. (2020), 'Pollution Transport Patterns Obtained Through Generalized Lagrangian Coherent Structures', *Multidisciplinary Digital Publishing Institute: Atmosphere* **11**(2), 168.

Olascoaga, M. J. & Haller, G. (2012), 'Forecasting sudden changes in environmental pollution patterns', *Proceedings of the National Academy of Sciences of the United States of America* **109**(13), 4738–4743.

Onu, K., Huhn, F. & Haller, G. (2015), 'LCS Tool: A computational platform for Lagrangian coherent structures', *Journal of Computational Science* **7**(1), 26–36.

Orselli, I. B. M., Goyet, C., Kerr, R., de Azevedo, J. L. L., Araujo, M., Galdino, F., Touratier, F. & Garcia, C. A. E. (2019), 'The Effect of Agulhas Eddies on Absorption and Transport of Anthropogenic Carbon in the South Atlantic Ocean', *Multidisciplinary Digital Publishing Institute: Climate* **7**(6), 84.

Rockwood, M., Huang, Y. & Green, M. (2018), 'Tracking coherent structures in massively–separated and turbulent flows', *Physical Review Fluids* **3**(1), 014702.

Rypina, I. I., Brown, M. G., Beron-Vera, F. J., Koçak, H., Olascoaga, M. J. & Udovydchenkov, I. A. (2007), 'On the Lagrangian Dynamics of Atmospheric Zonal Jets and the Permeability of the Stratospheric Polar Vortex', *Journal of the Atmospheric Sciences* **64**(1), 3595—-3610.

Schlueter-Kuck, K. L. & Dabiri, J. O. (2017), 'Coherent structure colouring: identification of coherent structures from sparse data using graph theory', *Journal of Fluid Mechanics* **811**(1), 468–486.

Schroeder, P. W., John, V., Lederer, P. L., Lehrenfeld, C., Lube, G. & Schöberl, J. (2019), 'On reference solutions and the sensitivity of the 2D Kelvin–Helmholtz instability problem', *Computers and Mathematics with Applications* **77**(1), 1010–1028.

Senatore, C. & Ross, S. D. (2011), 'Detection and characterization of transport barriers in complex flows via ridge extraction of the finite time Lyapunov exponent field', *International Journal for Numerical Methods in Engineering* **86**(1), 1163–1174.

Shadden, S. C., Dabiri, J. O. & Marsden, J. E. (2006), 'Lagrangian analysis of fluid transport in empirical vortex ring flows', *Physics of Fluids* **18**(4), 047105.

Shadden, S. C., Lekien, F. & Marsden, J. E. (2005), 'Definition and properties of Lagrangian coherent structures from finite–time Lyapunov exponents in two–dimensional aperiodic flows', *Physica D: Nonlinear Phenomena* **212**(1), 271–304.

Stuart, J. T. (1967), 'On finite amplitude oscillations in laminar mixing layers', *Journal of Fluid Mechanics* **29**(3), 417–440.

Suara, K., Khanarmuei, M., Ghosh, A., Yu, Y., Zhang, H., Soomere, T. & Brown, R. J. (2020), 'Material and debris transport patterns in Moreton Bay, Australia: The influence of Lagrangian coherent structures', *Science of the Total Environment* **721**(1), 137715.

Tallapragada, P. & Ross, S. D. (2013), 'A set oriented definition of finite–time Lyapunov exponents and coherent sets', *Communications in Nonlinear Science and Numerical Simulation* **18**(1), 1106–1126.

Teeraratkul, C., Irwin, Z., Shadden, S. C. & Mukherjee, D. (2021), 'Computational investigation of blood flow and flow–mediated transport in arterial thrombus neighborhood', *Biomechanics and Modeling in Mechanobiology* **1**(1), 1–15.

Verma, H., Agrawal, R. K. & Sharan, A. (2016), 'An improved intuitionistic fuzzy c–means clustering algorithm incorporating local information for brain image segmentation', *Applied Soft Computing* **46**(1), 543–557.

Wei, X., Zhan, H., Cai, S., Zhan, W. & Ni, P. (2018), 'Detecting the transport barriers in the Pearl River estuary, Southern China with the aid of Lagrangian coherent structures', *Estuarine, Coastal and Shelf Science* **205**(1), 10–20.

Welford, B. P. (1962), 'Note on a Method for Calculating Corrected Sums of Squares and Products', *Technometrics* **4**(3), 419–420.