REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
12/31/2018	Scientific	01/09/2015 - 09/30/2018

4. TITLE AND SUBTITLE
Coherence, Transport, and Inference in Turbulent Dynamical Systems

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-15-1-2093

5c. PROGRAM ELEMENT NUMBER
1000000479

6. AUTHOR(S)
Erik Matthew Bollt

5d. PROJECT NUMBER
4720000183

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
311

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Office of Naval Research
875 N Randolph St, Suite 1425
Arlington, VA 22203-1995

8. PERFORMING ORGANIZATION REPORT NUMBER
BD025

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
ONR REG Office Boston
495 Summer St, Room 627
Boston, MA 02210-2109

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
N62879

12. DISTRIBUTION/AVAILABILITY STATEMENT
All results are available and publically published.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Anaylzing and understanding features in chaotic and turblent systems is central to developing modern systems that thrive in fluid media. From ships in oceans to airplanes in the air, to chemical mixing and biological processes, both scale of vessels as well as the macro scale of weather and oceanic currents, is it clear that the Naval relevance issues call for ever better technology to analyze the underlying processes. This project concerns analysis of chaotic turbulent systems, from a Langrangian perspective including 1) Transport from a theory of Shape Coherence 2) Inference in Spatiotemporal

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Erik M. Bollt
not classified	not classified	not classified	no limitation	68	19b. TELEPHONE NUMBER *(Include area code)* 315-268-2307

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATE COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33315-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report. e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

Final Report 2018 for

"Coherence, Transport, and Inference in Turbulent Dynamical Systems"

N00014-15-1-2093

Erik Bollt, Clarkson University, PI

Contents

1 Degrees Awarded

- Warren Lord, PhD, Clarkson University, 2014-2018, Inference of networks of causal relationships using Causation Entropy.
- A. AlMomani, MS, Clarkson University 2017, continuing to PhD.
- Kelum Hewa Gajamannage, PhD, Clarkson University, 2011-2016, Manifold Learning and Cooperative Group Behavior.
- Matt Fowler, MS student, Clarkson University, 2013-2015.

Abstract Analyzing and understanding features in chaotic and turbulent systems is central to developing modern systems that can thrive in fluid media. From ships in oceans to airplanes in the air, to chemical mixing and biological processes, and so forth, both the local scale of the vessels as well as the macro scale of weather and oceanic currents, it is clear that the Naval relevance of these issues call for ever better technology to analyze the underlying processes. This project concerns analysis of chaotic and turbulent systems, from a Lagrangian perspective, specifically toward a modern perspective of coherent structures for understanding transport as well as persistence of underlying structures whether they be blooms of plankton in the oceans, or cascade of energy in turbulence. We emphasized two major thrusts in this work, both working toward this major theme of understanding aspects of simplicity embedded in nonlinear systems: 1) Transport from a theory of Shape Coherence. 2) Inference in Spatiotemporal Dynamical Systems: Inverse Problems and System Inference.

2 Goals

2.1 Questions Regarding Coherence and Transport in Turbulent Dynamical Systems

Transport from a theory of Shape Coherence: Coherence has clearly become a central concept of interest in nonautonomous dynamical systems, particularly in the study of turbulent flows, with many recent papers designed toward describing, quantifying and constructing such sets,]. There have been a wide range of notions of coherence, from spectral, to set oriented, and through transfer operators as well as variational principles. A general perspective of set oriented analysis of coherence seems to emphasize a discussion of transport. A number of theories are developed to model and analyze the dynamics in the Lagrangian perspective, such as the geodesic transport barriers [24] and transfer operators method. Whatever the perspectives taken, generally it may be summarized that coherent structures can be taken as a region of simplicity, within the observed time scale and over a stated spatial scale, perhaps embedded within an otherwise possibly turbulent flow. See discussion of curvature in the context of dynamical systems. Here we will review our recent theory of Shape Coherence and describe significant new computational and theoretical directions we plan to develop. These new directions will include relating coherence to a theory of transport for nonautonomous systems, extending to a three-dimensional theory of transport and coherence, deepening connections between geometry, differential geometry, measurable dynamics, and dynamical systems. and then application of these concepts to fluidic and especially oceanographic and atmospheric flow problems and including understanding advection, diffusion-reaction-problems such as plankton bloom growth in the oceans and estuaries.

2.2 Questions Regarding Inference in Spatiotemporal Dynamical Systems

Inverse Problems and System Inference From Remote Sensing: We have been developing convex optimization methods for systems inference, designed for data sets from remote sensing platforms, as summarized in our review article the recent PhD thesis. This includes methods for inferring both advective information (time varying vector fields) and parametric information regarding system identification, especially

for problems from ecological oceanographic systems as information remotely by satellite imagery. This continues to be a promising thread with aspects to develop further that we describe here, to continue to build a practical tool for modeling spatiotemporal systems remotely and therefore to inform the analytic coherence and transport discussion of the first part of this project described above. As we will describe here, we will adapt recent Bayesian learning based data fusion methods to the problem of improving our current estimates to mixed data forms, suitable for improving already developed estimates from remotely observed scalar fields to the possibility of mixed data including from sparsely placed floaters (buoys). Such improvements should significantly help with missing data from a remote sensor based method alone which is subject to occlusion.

Navy relevant technologies as well as Navy relevant environmental operating conditions all depend on excellent mathematical techniques for predicting and analyzing turbulent flows. Therefore, we believe that our program as described here, with some carefully chosen specific problems, will introduce and streamline new methodologies for global analysis of chaotic and complex dynamical systems problems.

3 Accomplished under Goals

3.1 Annotated Bibliography of Journal Publications

- A. AlMomani and E. M. Bollt, "Go With the Flow, on Jupiter and Snow. Coherence From Model-Free Video Data without Trajectories," J Nonlinear Sci (2018). https://doi.org/10.1007/s00332-018-9470-1. We develop remote sensing inference methods for coherent set by an anisotropic directed diffusion operators corresponding to flow on a directed graph, from a directed affinity matrix developed for coherence by spectral graph theory. We analyze weather both on Jupiter and Earth.

- Abd AlRahman R. AlMomani, Jie Sun, and Erik Bollt, "How Entropic Regression Beats the Outliers Problem in Nonlinear System Identification," submitted (2018). We develop a new information theoretic method of system identification that we prove signficiantly out performs standard metric based methods, especially in the environment of a sparse problem but corrupted by significant outlier noise, where metric based methods including compressed sensing give poor results.

- J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M.D. Mahecha, E.H. van Nes, J. Peters, R. Quax, M. Reichstein, M. Scheffer, B. Schlkopf, P. Spirtes, G. Sugihara, J. Sun, K. Zhang, & J. Zscheischler, "Perspective article: Inferring causation from time series with perspectives in Earth system sciences," submitted (2018) In disciplines dealing with complex dynamical systems, such as Earth science, replicated real experiments are rarely feasible. We catalogue methods for increasing amounts of observational and simulated data by data-driven causal inference methods, common in Earth science where causal methods will potentially advance the state-of-the-art.

- Erik Bollt, Qianxiao Li, Felix Dietrich, Ioannis Kevrekidis, "On Matching, and Even Rectifying, Dynamical Systems through Koopman Operator Eigenfunctions," SIAM Journal on Applied Dynamical Systems 17.2 (2018): 1925-1960. Matching dynamical systems, through different forms of conjugacies and equivalences, has long been a fundamental concept in classification of non- linear dynamic behavior. We show a data-driven algorithm developments by Koopman spectral theoretical and computational machinery to develop diffeomorphismic transformations between integrable dynamical systems.

- Erik Bollt, "Open or Closed? Information Flow Decided by Transfer Operators and Forecastability Quality Metric," Chaos: An Interdisciplinary Journal of Nonlin, 28, 075309 (2018). We assert that the concept of information flow, and the related causation inference, are summarized by questions of closure. Two alternative forms of restricted Frobenius-Perron (FP) operators, interpreted as either

closed (determinstic FP) or not closed (the unaccounted outside influence requires the stochastic FP operator).

- Jim Bagrow, Erik Bollt, "An information-theoretic, all-scales approach to comparing networks," submitted, (2018). An important task network science is network comparison, developing a similarity or distance measure between for meaningful comparisons between graphs. We introduce a new measure to compare networks, the Portrait Divergence, that is a Jensen-Shanon divergence based on our graph invariant, the network portrait.

- Donal Harkin, Naratip Santitissadeekorn, Christof Meile, Erik Bollt, and George Waldbusser, "Bayesian analysis for image-based motion estimation with application to environmental flows," submitted, (2018). We develop a Bayesian algorithm for unobserved velocity fields from experimental marine surficial sediment, where flow is caused by pumping activity of burrowing macrofauna. The model reflects mass conservation equations for dissolved tracers, which is simplified in our application to the so-called optical flow equation.

- Joshua Garland, Andrew M. Berdahl, Jie Sun, and Erik Bollt, "The Anatomy of Leadership in Collective Behaviour" Chaos: An Interdisciplinary Journal of Nonlin, 28, 075308 (2018). Mobile animal groups are composed of individuals with different levels and types of influence over group behaviors. We assert that complex interaction rules and dynamics typical of groups imply that leadership is not merely binary classifiers (leader or follower), rather, complex combinations of multiple components.

- Warren M. Lord, Jie Sun, Erik Bollt, "Geometric k-nearest neighbor estimation of entropy and mutual information," Chaos: An Interdisciplinary Journal of Nonlin, 28, 033114 (2018). For efficient statistical estimation of mutual information, we introduce a new class of knn estimators, geometric knn estimators (g-knn), which use more complex local volume elements to better model the local geometry of the probability measures that outperforms the previous standard, Kraskov-St?ogbauer-Grassberger (KSG) estimation.

- K.G.D. Sulalitha Priyankara, Sanjeeva Balasuriya, and Erik Bollt, "Quantifying the Role of Folding in Nonautonomous Flows: the Unsteady Double-Gyre," International Journal of Bifurcation and Chaos, Vol. 27, No. 10 (2017) 1750156. We analyze the stretching + folding = chaos mantra of chaotic dynamics showing curvature can be used to identify fold points in general nonautonomous flows in two dimensions. Our new technique, fold reentrenchment, allows proof of chaos and classifies turbulent behaviors otherwise missed by FTLEs.

- Bruce Alstrom, Pier Marzocca, Erik Bollt, "Nonlinear System Identification of a Rossler System Under Periodic Closed-Loop Control Via Time-Frequency and Bispectral Analysis," Mechanical Systems and Signal Processing (MSSP), 99 (2018) 567585. A fixed gain fixed frequency controller produces quadratic phase-coupling along constant frequency that are perpendicular to diagonals of the bicoherence matrix, demonstrated in a Rossler system, for system identification and control. Under periodic closed-loop control via wavelet bispectral analysis, component mechanisms of synchronization are stabilized.

- Jie Sun, Fernando Quevedo, Erik Bollt, "Statistical Inverse Formulation of Optical Flow with Uncertainty Quantification," Inverse Problems 34 105008 (2018). We develop Bayesian statistical inverse methods for optical flow that unlike classical approaches, includes not only allows for point estimates, but also provides a distribution of solutions for uncertainty quantification, of vector fields, applied to synthetic and real world images sequences, applicable to remote sensing.

- Kelum Gajamannage, Randy Paffenroth, and Erik Bollt, "A Nonlinear Dimensionality Reduction Framework Using Smooth Geodesics," to appear Pattern Recognition, (2018). We develop a nonlinear dimensionality reduction framework for manifold that emphasizes smoothness of geodesics,

for data on manifold that have been noise corrupted. Robustness of this approach for noisy and sparse datasets is demonstrated by the implementation of the method on synthetic and real-world datasets.

- Jie Sun, Fernando J. Quevedo, Erik Bollt, "Data Fusion Reconstruction of Spatially Embedded Complex Networks," submitted (2017). We introduce a kernel Lasso (kLasso) optimization that simultaneously accounts for spatial regularity and network sparsity to reconstruct spatial complex networks from data, that exploits spatial embedding distances to penalize overabundance of spatially long-distance connections. Examples include both synthetic and real-world spatial networks.

- Qianxiao Li, Felix Dietrich, Erik M. Bollt, Ioannis G. Kevrekidis, "Extended dynamic mode decomposition with dictionary learning: a data-driven adaptive spectral decomposition of the Koopman operator," Chaos: An Interdisciplinary Journal of Nonlin (Vol.27, Issue 10), 103111 (2017). Numerical approximation methods for Koopman operators, the dynamic mode decomposition (DMD) and extended-DMD (EDMD), are significantly improved by machine learning methods to construct a trainable dictionary of best observer functions. Using the Duffing oscillator and the Kuramoto Sivashinsky PDE examples, we efficiently study global behaviors.

- Erik Bollt, Jie Sun, Jakob Runge, "Introduction to Focus Issue: Causation inference and information flow in dynamical systems: Theory and applications" Chaos: An Interdisciplinary Journal of Nonlin 28, 075201 (2018). Questions of causation are foundational across science and often relate to problems of control, policy decisions, and forecasts. In nonlinear dynamics and complex systems science, causation inference and information flow are closely related concepts, whereby information or knowledge of states defines complex systems science.

- Felix P. Kemetha, Sindre W. Haugland, Felix Dietrich, Tom Bertaland, Qianxiao Lie, Erik Bollt, Ronen Talmon, Katharina Krischera, and Ioannis G. Kevrekidis, "An Equal Space for Complex Data with Unknown Internal Order: Observability, Gauge Invariance and Manifold Learning," submitted, (2017). Connections between manifold-learning extracts intrinsic coordinates (order) from observations of complex dynamics. Systems modeling considerations allow tuning scales of the data-mining kernels of dynamic models at different levels of coarse-graining. Observability of physical space from temporal data versus spatially resolved lumped representations, yields Equal Space.

- K.G.D. Sulalitha Priyankara, Sanjeeva Balasuriya, and Erik Bollt, "Quantifying the Role of Folding in Nonautonomous Flows: the Unsteady Double-Gyre," submitted (2017). We analyze chaos in the well-known nonautonomous Double-Gyre system. A key focus is on folding, which is possibly the less-studied aspect of the stretching + folding = chaos mantra of chaotic dynamics. Despite the Double-Gyre not having the classical homoclinic structure for the usage of the Smale-Birkhoff theorem to establish chaos, we use the concept of folding to prove the existence of an embedded horseshoe-map. We also show how curvature of manifolds can be used to identify fold points in the Double-Gyre. This method is applicable to general nonautonomous flows in two dimensions, defined for either finite or infinite times. The well studied Double-Gyre system may be considered as a standardized problem to contrast chaos from mixing. It is often invoked as a model for a chaotic system, and used as a testbed in numerical simulations. Strangely, however, there does not appear to be a proof in the literature that the system is actually chaotic. While it is easy to establish that certain stable and unstable manifolds intersect, there are technical impediments in taking the next step to claim that this results in chaos. By using a new technique we call fold reentrenchment, we are able here to show the presence of a Smale horse-shoe embedded in the phase space. In this process, we are focussing on the folding aspect of chaos, in contrast to highly popular methods such as Finite-TimeLyapunov Exponents (FTLEs) which target the quantification of stretching. We

further develop a quantification of folding as a systems propensity to develop curvature, and show how this criterion can be highly informative in analyzing the chaotic nature of general systems.

- Kelum Gajamannage, Erik Bollt, "Detecting Phase Transitions in Collective Motion Using Manifold's Curvature , " Mathematical Biosciences and Engineering (MBE), vol. 14, no. 2, pp. 437-453, (2017) If a given behavior of a multi-agent system restricts the phase variable to an invariant manifold, then we define a phase transition as a change of physical characteristics such as speed, coordination, and structure. We define such a phase transition as splitting an underlying manifold into two sub-manifolds with distinct dimensionalities around the singularity where the phase transition physically exists. Here, we propose a method of detecting phase transitions and splitting the manifold into phase transitions free submanifolds. Therein, we firstly utilize a relationship between curvature and singular value ratio of points sampled in a curve, and then extend the assertion into higher-dimensions using the shape operator. Secondly, we attest that the same phase transition can also be approximated by singular value ratios computed locally over the data in a neighborhood on the manifold. We validate the Phase Transition Detection (PTD) method using one particle simulation and three real world examples.

- Erik Bollt, "Regularized Forecasting of Chaotic Dynamical Systems," Chaos Solitons and Fractals, 94 8-15 (2017). While local models of dynamical systems have been highly successful in terms of using extensive data sets observing even a chaotic dynamical system to produce useful forecasts, there is a typical problem as follows. Specifically, with k-near neighbors, kNN method, local observations occur due to recurrences in a chaotic system, and this allows for local models to be built by regression to low dimensional polynomial approximations of the underlying system estimating a Taylor series. This has been a popular approach, particularly in context of scalar data observations which have been represented by time-delay embedding methods. However such local models can generally allow for spatial discontinuities of forecasts when considered globally, meaning jumps in predictions because the collected near neighbors vary from point to point. The source of these discontinuities is generally that the set of near neighbors varies discontinuously with respect to the position of the sample point, and so therefore does the model built from the near neighbors. It is possible to utilize local information inferred from near neighbors as usual but at the same time to impose a degree of regularity on a global scale. We present here a new global perspective extending the general local modeling concept. In so doing, then we proceed to show how this perspective allows us to impose prior presumed regularity into the model, by involving the Tikhonov regularity theory, since this classic perspective of optimization in ill-posed problems naturally balances fitting an objective with some prior assumed form of the result, such as continuity or derivative regularity for example. This all reduces to matrix manipulations which we demonstrate on a simple data set, with the implication that it may find much broader context.

- Sean Kramer, Erik M. Bollt, "An Observer for an Occluded Reaction-Diffusion System With Spatially Varying Parameters," Chaos: An Interdisciplinary Journal of Nonlinear Science, 27, 035810 (2017) (2017) Spatially dependent parameters of a two-component chaotic reaction-diffusion PDE model describing ocean ecology are observed by sampling a single species. We estimate model parameters and the other species in the system by autosynchronization, where quantities of interest are evolved according to misfit between model and observations, to only partially observed data. Our motivating example comes from oceanic ecology as viewed by remote sensing data, but where noisy occluded data are realized in the form of cloud cover. We demonstrate a method to learn a large-scale coupled synchronizing system that represents spatio-temporal dynamics and apply a network approach to analyze manifold stability. Research in large-scale oceanic phenomena is made possible by remote sensing instruments mounted on ocean-observing satellites. These instruments provide datasets that can be filtered to study sizable ecological events, including harmful algal blooms. The fact that datasets are often patchy when clouds hide regions in the spatial domain is a substantial

difficulty when attempting to parameterize a dynamical system. To attack this problem we extend a recently developed autosynchronization method. Model parameters and states are evolved in a drive-response pattern, on a-priori known model equations, to learn model states and parameters even while data are considerably spatially occluded. It has been shown that, assuming the model structure to be known, a synchronization system can be designed to effectively act as an observer to identify system parameters, even in a large scaled network system. While a discretized PDE can be interpreted as a particular lattice network, the realistic problem of cloud occlusions will cause times where the observer network is essentially disconnected. Our prior work has shown that synchronization can exist even in a large scale network that is not fully connected but rather has a so-called fast blinking structure. The method is analyzed by interpreting the discretized PDE as a large-scale coupled moving neighborhood network.

- Bruce Alstrom, Stephane Moreau, Noureddine Atalla Erik Bollt, Pier Marzocca, "Nonlinear Systems Identification of a NACA 0015 High Lift System Under Dual Location Open Loop Control," 52nd 3AF International Conference on Applied Aerodynamics 27 ? 29 March 2017, Lyon - France, FP58-AERO2017-alstrom (2017) An NACA 0015 airfoil fitted (with flow control capability) with a simple trailing edge flap set at an angle of attack of 16 degrees (deep stall), flap deflection is set at zero degrees at an airspeed of 5 m/s is the subject of this research. The flow conditions analyzed are an unforced stalled condition and an open loop controlled condition with a modulation frequency of 50 Hz. The analysis is performed using classical signal processing tools along with higher order spectral moments. For the un-forced flow condition, the analysis revealed that the autobispectral results are directly related to the RMS pressure distribution especially when there is a vortex shedding frequency present. The analysis also exposed the nonlinear signature of vortex pairing and shredding. For the forced flow condition, it is observed that forcing the flow stabilizes tones that otherwise would migrate downstream of the leading edge. The surface pressure fluctuations are reduced because of convective energy transfers between the forced fundamental of 43 Hz and low frequency components which results in reattachment of the flow.

- Bruce Alstrom, Pier Marzocca, Erik Bollt, "Nonlinear System Identification of a Rossler System Under Periodic Closed-Loop Control Via Time-Frequency and Bispectral Analysis," Mechanical Systems and Signal Processing (MSSP), to appear 2017 This study has two primary objectives; they are to investigate the nonlinear interactions (or quadratic phase-coupling) in a chaotic Rossler system under periodic closed-loop control via wavelet bispectral analysis; and to further identify the component mechanisms of synchronization. It is observed that a fixed gain fixed frequency controller produces quadratic phase-coupling and decoupling along lines of constant frequency and that are perpendicular to the diagonal of the bicoherence matrix. Further, it was also observed that for synchronization to occur, both frequency entrainment and quadratic phase-coupling must be present. It was found that forcing the Rossler system with a constant frequency did not reduce the amplitude of the resulting period-1 orbit at sufficiently high gains. For the controller with a fixed gain and time-varying error signal, it was found that the time varying forcing frequency (adjusted by an extremum seeking feedback loop) linearizes the Rossler system and in doing so, suppresses the phase coherence completely. The time-varying forcing frequency removes the conditions for frequency entrainment by providing broadband attenuation; the result is suppression without synchronization.

- Tian Ma, Nicholas Ouellette, Erik M Bollt, "Stretching and Folding in Finite Time , " Chaos 26, 023112 (2016). Complex flows mix efficiently, and this process can be understood by considering the stretching and folding of material volumes. Although many metrics have been devised to characterize stretching, fewer are able to capture folding in a quantitative way in spatiotemporally variable flows. Here, we extend our previous methods based on the finite-time curving of fluid-element trajectories to nonzero scales and show that this finite-scale finite-time curvature contains information about both stretching and folding. We compare this metric to the more commonly used

finite-time Lyapunov exponent and illustrate our methods using experimental flow-field data from a quasitwo-dimensional laboratory flow. Our new analysis tools add to the growing set of Lagrangian methods for characterizing mixing in complex, aperiodic fluid flows One of the hallmarks of chaotic or turbulent fluid flows is rapid mixing. Vigorously stirred fluids mix so well because they move material quickly and yet are constrained to remain in a finite amount of space. The chaotic mixing process can thus be conceptualized as the result of stretchingthat is, the rapid separation of nearby packets of fluidand foldingthat is, the bending of packets of fluid to maintain the overall volume. Because these two processes are fundamental for mixing, they have been the subject of intense study; however, many more methods have been developed to characterize stretching than folding. Here, we describe a method based on the propensity of the trajectories of fluid elements to curve that allows us to quantify both stretching and folding at the same time. We compare this technique with a more common tool that isolates only stretching and demonstrate our methods on data from an experimental flow.

- Ranil Basnayake, Erik Bollt, Nicholas Tufillaro, Jie Sun, Michelle Gierach "Regularization Destriping of Remote Sensing Imagery with Missing Data Preprocessing, " Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2016-74, 2016 We illustrate the utility of variational destriping for ocean color images from both mulitspectral and hyperspectral sensors. In particular, we examine data from a filter spectrometer, the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi National Polar Partnership (NPP) orbiter, and an airborne grating spectrometer, the Jet Population Laboratorys (JPL) hyperspectral Portable Remote Imaging Spectrometer (PRISM) sensor. We solve the destriping problem using a variational 5 regularization method by giving weights spatially to preserve the other features of the image during the destriping process. The target functional penalizes the neighborhood of stripes (strictly, directionally uniform features) while promoting data fidelity, and the functional is minimized by solving the Euler-Lagrange equations with an explicit finite difference scheme. We show the accuracy of our method from a benchmark data set which represents the Sea Surface Temperature off the Coast of Oregon, USA. Technical details, such as how to impose continuity across data gaps using inpainting, are also described.

4 Technical Report

Here we will now review three of the works cited as part of the annotated bibliography above. These are chosen to highlight because they each highlight details relevant to the major goals.

4.1 Regarding Goal 1: Part 1: Questions Regarding Coherence and Transport in Turbulent Dynamical Systems

Material here associated with PI's publication, A. AlMomani and E. M. Bollt, "Go With the Flow, on Jupiter and Snow. Coherence From Model-Free Video Data without Trajectories," J Nonlinear Sci (2018). https://doi.org/10.1007/s00332-018-9470-1.

4.1.1 Introduction

There has been significant emphasis in recent dynamical systems literature to define, and find "coherent structures," [67, 10, 22, 41, 54, 69, 24, 25, 26]. It could be said that these methods could be divided into those that follow interiors of sets by transfer operators, or those that define a property of boundaries of such sets and follow boundary curves, [48, 49, 37, 38]. Some methods have been developed and put forward without specifically defining the coherency principle that the method is designed to extract. In any case, perhaps most would agree that coherency should be defined in some manner to describe sets (of particles) that "hold together" for some time, or densities of ensembles of particles [43, 72, 56], or measurements thereof [12, 42, 102, 3].

In essentially all of the studies that have appeared in recent literature, no matter what the method, approach or perspective, one starts with a dynamical system. From there follows the quantity to be analyzed. In other words, an underlying flow is assumed in the sense that generally a differential equation is required to proceed, whether explicitly or implicitly through observations of an experiment. For this we will write,

$$\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, t), \tag{4.1}$$

for a vector field, $\mathbf{F} : M \times \mathbb{R} \to M$, (typically $M \subset \mathbb{R}^2$ or perhaps $\mathbb{R}^3$), but this may be developed from a stream function from an underlying partial differential equation for example. In any case, then a flow mapping, $\mathbf{x}(t) = \Phi(\mathbf{x}_0, t_0, t)$ is inferred, even if this means numerical integration of the differential equation. In recent work, aspects of advection and diffusion have been both involved in developing a better understanding of coherence, [24, 25, 17], including for models of stochastic processes. We summarize that universally, previous work either begins with a model of the dynamical system, or at least attempted to empirically develop a model perhaps by optic flow, including our own, [45, 6, 46] or similarly by other means, [36], and recently by Koopman operator methods but requiring a vector field, [19].

In contrast to all the mathematical formalism and machinery behind current studies of coherency, it can be said that people "recognize" coherent sets when they see them; consider that the Great Red Spot of Jupiter is clear to any and all that have seen it, as perhaps the most famous coherent set in the solar system. With this motivation, we will develop here an observer based perspective of coherence. If we do not have a model, as the dynamical system is known only by remote sensing observations, then in practice the flow mapping, $\Phi(\mathbf{x}_0, t_0, t)$ is at best inferred, but generally not available, and often likewise nonlinear systems require numerical integration to infer the flow at sampled points. Here we will approach questions of coherence in the setting that we have only remote observations, but no model. Developing a model of the flow either directly, $\Phi(\mathbf{x}_0, t_0, t)$, or as a model of the vector field (say by optic flow), may not always be practical or the best way to proceed.

Figure 1: Jupiter Portrait as viewed from the spaceship Cassini. "This true color mosaic of Jupiter was constructed from images taken by the narrow angle camera onboard NASA's Cassini spacecraft on December 29, 2000, during its closest approach to the giant planet at a distance of approximately 10 million kilometers (6.2 million miles)." [59].

Take as case in point that the Great Red Spot (GRS) was observed and identified as persistent over many years without ever needing to develop a great deal of the formalism associated with our modern descriptions and algorithms of coherent sets. No transfer operators [43, 56], no Koopman operators [12, 102], and no vector field were required [46]. See Fig. 18, as seen in the year 2000 from the Cassini space probe, a joint NASA, European Space Agency (ESA), and Italian space agency Agenzia Spaziale Italiana (ASI) mission [60]. The solar system's largest and most persistent planetary hurricane storm, the vortex structure called the GRS is clearly visible in this image. There are also belts and zones as persistent latitudinal structures, as well as many other smaller storms, (but still massive by Earth standards). There are also other embedded objects, that are clearly present and notable by the naked eye, without ever needing a digital computational engine to identify. It is as clear today to the casual observer of these modern images, as it was to the Renaissance era astronomer Giovanni Domenico Cassini himself that there are coherent structures on Jupiter, [20]. See Fig. 19 where Cassini's sketches show some of the same structures as viewed across several years from 1665-1677, were clear enough that he was able to see them despite what were low quality telescopes by any modern standard, and many of these structures persist today, hundreds of years later. It is important to distinguish between the concept of a feature that we may notice in a single image as compared to a feature that persists over several successive images, over time. Persistence over time is more akin to what is meant by coherency; we will contrast image segmentation concepts versus motion segmentation concepts. In Sec. 4.1.6, this contrast leads us to a directed affinity.

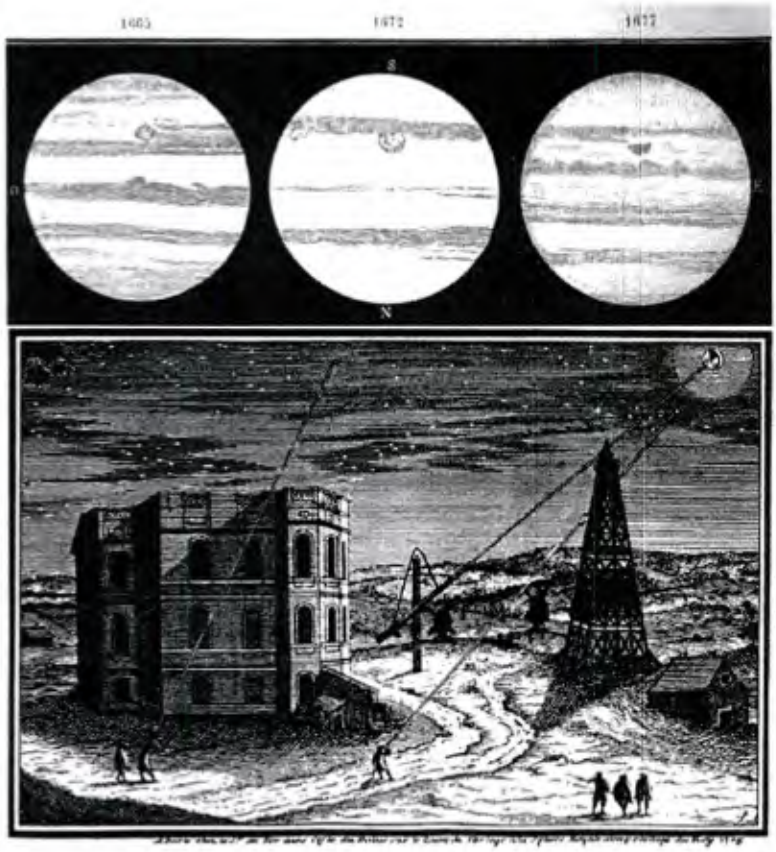


Figure 2: Jupiter as sketched by Giovanni Domenico Cassini (Top) in his own hand from 1665-1677, from the Memoires de l'Académie Royale des Sciences de Paris, [20]. Note that North is drawn, and so labelled, on the bottom. We see that Cassini was seeing and sketching similar scenes over the several years, including apparently the large storm. (Bottom) A sketch of the observatory in Paris.

It is our goal here to define the manner in which we *know* a coherent structure when we see it, and to develop appropriate mathematical formalism behind this sense. Furthermore, and also central to this work, since an observer can gain impression of persistence of certain structures, then our perspective should be developed to be directly comparable to imagery, without needing to go through steps of modeling the imagery by computed vector fields, then integrating the vector fields numerically to develop a flow map, before only then involving the methods of geometric dynamical systems; note that we admit this is counter even to our own previous efforts [45, 46] which have followed this exact prescription for data driven remote sensing starting with specializing optic flow methods and comparably even for the study of the atmospheres of Jupiter by [35]. The principle we choose in this current work is that we could start and end with the images themselves, as representing pointwise measurements in time.

In this paper we structure the presentation as follows. In Sec. 4.1.2 we relate to the concept that observations in a color image are spatial measurements, which when evolved in time, relate to measurements along orbits, noting an observer-centric perspective. In Sec. 4.1.6 we formalize the idea of motion segmentation as a partition in space, along time, and we relate this to coherency. We contrast ideas from the image processing community about image segmentation which is inherently a symmetric concept in almost any algorithmic approach one might take, as compared to motion segmentation that leads necessarily to a not symmetric description due to the arrow of time inherent in the concept. We review in Sec 4.1.4-4.1.5, the especially popular methods of k-means clustering and also spectral clustering, respectively, as related to image segmentation. Note that since they are inherently symmetric concepts how they should not be directly applied to motion segmentation or coherency questions. So in Sec. 4.1.6

we develop a directed affinity which naturally incorporates the asymmetry of directed time in a manner that describes coherency as a spatial and time oriented version of "particles hold together", with details relating to the graph Laplacian of a directed graph underlying the directed graph version of spectral graph theory. In Sec. 4.2.2-4.2.4 we include background material regarding the nCut problem, for clustering, and leading to spectral clustering for the directed graph problem with the corresponding special case of a weighted directed graph Laplacian, as used in the directed spectral segmentation. In Sec. 4.1.7, we address experimental and numerical results by showing motion segmentation by directed spectral segmentation methods naturally finds convincing coherent sets, in data from Jupiter, from a lake effect snow effect data set from a local storm near our own university, and from the highly popular double-gyre system often used for benchmarking coherent set analysis.

4.1.2 Measure Along Orbits

Suppose the dynamical system, Eq. 4.1, over the phase space M may not be directly known to us, and we have a measurable function,

$$\mathbf{c} : M \to \mathbb{R}^d, \tag{4.2}$$

where d is the number of scalar measurements made. In the case of the image shown in Fig. 18, $d = 3$, and $\mathbf{c}$ samples three color intensities from a standard color scale (such as RGB as shown) at each point $\mathbf{z}$ in the field of view; that is at a given time, $c_j(\mathbf{z}) : M \to \mathbb{R}$, $j = 1, 2, 3$, measures any one of the color plane values at $\mathbf{z}$. For a hyperspectral imager, generally $d > 3$. For intrinsic and not directly measured quantities, such as concentrations of a particular chemical gas, say call that at a given time, $h_i : M \to \mathbb{R}$, describes pressures, temperatures, or gas concentrations, for example, [23]. In the case of Jupiter, for example, the colors and intensities at each point represent chemical concentrations of various chemicals in the clouds, densities, depth of the cloud layers, and other properties as inferred by reflectance [23]. Then measured quantities are collection of functions $\{h_i\}_{i=1}^K$, hidden to us but combined into the function, $\mathbf{c}(\mathbf{z}) = \mathbf{q}(h_1(\mathbf{z}), .., h_K(\mathbf{z}))$, by some unknown to us function $\mathbf{q}$ related to the underlying physics.

What allows us to describe patterns in images as coherent, is that they persist in some form across many frames of the "movie", meaning as the system is observed through multiple times, and gradually evolving, what is seen is close enough to the original that we recognize it; in [49, 48] we suggested the concept of shape coherence as sets that almost maintain shape over time.

In [28, 26] a concept of coherent pairs was developed that roughly states that a coherent pair of sets A and B should be such that $\Phi(A) \approx B$, but also $\Phi^{-1}(B) \approx A$ with some notion of diffusion or randomness to reward those set pairs when the boundary does not grow too large. Here we have simplified the notation of the flow from Eq. 4.1, as $z = \Phi(z_0) = \Phi(\mathbf{z}_0, t_0, t)$ suppressed noting t_0, t taken to be fixed for sake of discussion. This was formalized with understanding a diffusive operator associated with numerical estimation of the transfer operators, in [26] and then a complete theory of the boundary sets, isoperimetry ("bottleneckyness" or Cheeger constants) was developed for this situation, [28]. The idea of studying the boundary of sets then also relates to the concepts of recent formulations of geodesic Lagrangian Coherence Structures (LCS) and transport barriers in terms of studying strain and also length minimizing curves, [37, 38]. In some sense both aspect of stretching and folding associated with curvature may have a role, [50].

It is interesting to relate observing as described here to the notion of measuring along orbits, related to the Koopman operator, [12]. Considering $z = \Phi(\mathbf{z}_0)$ as the "down stream" image of an initial condition $\mathbf{z}_0$, then to measure (the colors) down stream from $\mathbf{z}_0$ is a concept defined by the Koopman operator formalism, [12, 42] which we recall[12, 8], $\mathcal{K} : \mathcal{F} \to \mathcal{F}, \mathcal{K}[h](\mathbf{z}) = h \circ \Phi(\mathbf{z})$. where $\mathcal{F}$ is taken to be a space of measurable functions over M, and $h \in \mathcal{F}$. Recall that, "...the Koopman operator maps functions of state space to functions of state space and not states to states" [102]. Several measurable functions such as $\mathcal{G} = (g_1, .., g_K)$, Eq. 4.2, has been called, a "vector valued observable," [102]. A Koopman operator applied to each is inherited by the vectorized version of the Koopman operator, $\mathcal{K}_t[\mathbf{c}] = (\mathbf{K}_t[c_1], .., \mathbf{K}_t[c_k])$. Then measuring color at A may be written as $\mathbf{c}(A)$ as a remote measurement related to gases, pressures

and so forth in the scene. Measurement downstream would be, $K[\mathbf{c}](A)$ as the push forward of A. Since the adjoint $\mathcal{K}^*$ has properties of a pull back operator, (and associate with the Frobenius-Perron transfer operator, [8]), then to have measurements on the push forward match those on the pull back, approximately, is to demand, $\mathcal{K}^*\mathcal{K}[\mathbf{c}](A) \approx A$, but likewise, the stated in reverse there should be an approximate matching for a coherent pairing, $\mathcal{K}\mathcal{K}^*[\mathbf{c}](B) \approx B$. Note then that, $\mathcal{K}^*$ may be defined in terms of the pull back $\mathcal{K}^*[\rho](\mathbf{z}) = \rho \circ \Phi^{-1}(\mathbf{z})$, when Φ^{-1} exists as it will if it is a flow, but alternatively, the Frobenius-Perron operator is, $\mathcal{K}^*[\rho](\mathbf{z}) = \int_M \delta(\mathbf{z} - \Phi(\mathbf{y}))\rho(\mathbf{y}d\mathbf{y}$, and the Koopman operator can be written, $\mathcal{K}[h](\mathbf{z}) = \int_M \delta(\mathbf{y} - \Phi(\mathbf{z}))h(\mathbf{z}d\mathbf{y}$, where for sake of brevity, we are suppressing statement of the space of functions where this is appropriate, and the corresponding discussion of bilinear relating the operator and its adjoint, [12, 8]. In [28] a spectral method was developed associate with these eigenfunction type statements for the operator, $\mathcal{K}^*\mathcal{K}$, and this description is expanded upon in [4]. An average of both forward and backward time coherent pairings was offering in [29], including a statement that these concepts are associated with keeping small boundaries. These have proven to be a very effective and powerful approach, however, they require Lagrangian trajectory data. Even recent clustering methods such as the k-means approach in [30], or the spectral approach in [34] require Lagrangian trajectory data (stated roughly as measurements following along with orbits). There has been related work in spatiotemporal feature extraction and forecasting from the Koopman perspective, [31, 32], but also adaptations of the Koopman operator for image texture analysis[70] and also for video segmentation, [1]. It contrast, it can be said that remotely sensed "movie" data is inherently Eulerian (stated roughly as measurements associated with fixed positions in space).

With this background, we now proceed to contrast image segmentation methods toward developing a spectral motion segmentation method. Notice that when only movie data is available, then we specifically lack the Lagrangian trajectory data to explicitly carry forward any of the several operator methods or boundary methods or LCS methods from the literature. So in this case, we proceed to build a proxy operator, that rewards concepts of like measurements, and close distance, and in many ways this proxy operator serves the role of a transfer operator estimator, perhaps likely a Bayesian estimator, which we plan to pursue as a question in future work. Only the DMD method (Dynamic Mode Decomposition) [12], [42]-[66] can also directly handle movie data, but is also somewhat different in approach to how the operator is estimated by a least squares approach. As for now, notice that stated as an anzatz, we are emphasizing continuity in space and continuity in time measurements of the underlying but unknown flow.

4.1.3 Image Segmentation and Symmetric Affinity, versus Motion Segmentation and Not Symmetric Affinity

"Following along measured observations," for clustering is not necessarily the same as following along orbits for coherence, but they are easily confused, even if these ideas sometimes may coincide. And rightly, they both might be called coherence depending on the setting. We have quoted the phrase "following along measured observations," because this is roughly describing a cluster of like measurements that tracks in time, but may not specifically be exactly attached to the underlying advective flow. Specifically in the lake effect snow Example 4.1.9, we illustrate how an advective-reactive-diffusive process may give a very different notion of coherency than simple an advective or advective-diffusive process.

In the image processing community, the general problem of clustering "like objects" between frames of a movie is called *motion segmentation*, also known as *image motion segmentation*, and in turn the outcome of these have been used in the image processing community to infer *motion tracking*. Motion tracking of objects, or tokens, is not necessarily the same problem as the inference of the underlying flow (*tracking [13] would traditionally be applied in image processing of say a movie of people moving to reveal the underlying motions of individuals, or groups of individuals, as a "token".*) which is easily seen when considering the weather event illustrated in Fig. 20. We will cast this work as one of motion tracking, and then contrast to the Lagrangian coherent structures problem, [62]. We argue that only the former is

truly accessible by remote sensing.

First we review the static time problem of *image segmentation,* generally as clustering problems, and in the language of our image data from remote sensing. Consider clustering within a single scene, meaning a single frame of a movie. Suppose a grid of positions where color (or some other collection of pointwise measured quantities) is sampled, at each of $\{\mathbf{z}_i\}_{i=1}^M$ for M (usually uniformly spaced grid of) pixels over $\{\mathbf{z}_i\}_{i=1}^M \subset \mathcal{M} \subset \mathbb{R}^2$. So $\mathcal{M}$ is the framed image. At each of these, observe $\mathbf{c}(\mathbf{z}) : \mathcal{M} \to \mathbb{R}^d$, (generally say $d = 3$ colors at each position) to form an observation matrix,

$$X_{i,r} = \mathbf{c}(\mathbf{z}_i). \tag{4.3}$$

Since $\mathbf{c}$ is a vector valued measured observation with d observations (colors), then X is $M \times d$. For many frames sampled across time, we will write, $X_{i,r}[k]$,where $1 \leq r \leq d$ "colors," and for each time t_k, $0 \leq k \leq N-1$.

Our goal is to partition the space based on a notion of coherency, across time and space. By spatial partition the space of sampled data, we mean, given data $\{\mathbf{z}_i\}_{i=1}^M$, there is an assignment into labels, $\mathcal{S} = \{S_l\}_{l=1}^k$ that serves as a function from the pixel positions to (colored) labels. How this assignment should be done appropriately is a matter we now discuss and we describe how it should relate to how the measured ⌋ values vary across time. Usually a clustering is useful if it associates "like" (in some suitable sense) c-measured values of the data. Perhaps the two most commonly useful image segmentation methods are called k-means[40] and spectral segmentation[61] respectively.

Image segmentation may be formulated as a spectral graph partitioning problem, [61], which we review in Sec. 4.1.5. However, these methods need a major adjustment when applied to image sequences (movies) for motion segmentation, despite that traditionally they have been applied to movies with some degree of success [68]. The key difference is what underlies a notion of coherent observations, remembering that the arrow of time has directionality. We require *affinity matrices that are not symmetric,* and when considered as graphs, they are *directed graphs.* Therefore much of the theoretical underpinning of the standard spectral partitioning needs some adjustment, since it relies on symmetric matrices and undirected graphs. We will need a graph Laplacian for weighted directed graphs.

4.1.4 On k-means Image Segmentation by Color Alone

A simple and common form of clustering that one might choose would be a k-means clustering of an image scene [40] based just on the pointwise measurements alone (say colors for example) as a solution to the partitioning problem, to find a partition $\mathcal{S}$ such that

$$\mathcal{S} = \underset{S}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{X_{j,:} \in S_i} \|X_{j,:} - \mu_i\|, \tag{4.4}$$

where $\|\cdot\|$ is the Euclidean norm of the color values, and μ_i are means in each color channel. We see the k-means method is a solution of a partitioning problem. An image such as that of the colors of Jupiter is shown in Fig. 4 (b and c) for an example of a static time segmentation of a Jupiter image with $d = 3$ colors $\mathbf{c}(\mathbf{z})$, measured pointwise where $\{\mathbf{z}_i\}_{i=1}^N$ are the pixel positions on the image. The k-means problem solution has a direct method of updating the cost function Sec. 4.4 as membership of indexed values in each partition element is adjusted, thus shifting the group mean, until optimality is found sufficiently, as reviewed in many standard texts [18, 40, 2].

Beyond the k-means clustering concepts there are spectral clustering methods, and in fact even the k-means method has a spectral formulation,[18]. We will see that spectral methods seem to perform well from a clustering perspective alone, but also spectral descriptions of clustering will allow us to interpret our notions of motion segmentation more naturally in terms of coherence as analogies to spectral decompositions of transfer operators lead to dynamical systems concepts of coherence as already emphasized in the literature, [46, 24, 62]. So we proceed to recall the spectral concepts of clustering for image segmentation.

4.1.5 On Spectral Segmentation by Color Alone

There have been several complementary views of clustering by spectral methods, by graph cuts [68, 39], as random walkers, [52], and comparably as a diffusion process as described by diffusion map [15] and comparably as an eigensystem. Many of these come back to some version of a max-flow min-cut algorithm that we will review in Sec. 4.2.2, [63], and in turn as related to the conductance also called Cheeger-constant as a measure of "bottleneckyness" of the underlying graph. In this section, we review the computations for the simpler case of weighted *undirected* graphs, appropriate for image segmentation, but in the subsequent section we will relate our motion segmentation problem to the graph problem of weighted *directed* graphs to account for the directed aspect of the arrow of time.

Proceeding computationally, image segmentation may be formulated as a graph partitioning problem, and as such, doing so with color alone means formulating the data set; assign data set [61],

$$X = [X_{1,:}^T | X_{2,:}^T | ... | X_{M,:}^T], \tag{4.5}$$

So, for color alone, X is $d \times N$. Columns of X are the color channels at each pixel position $\mathbf{z}_i$, and we write $X_i = X_{i,:}^T$. If distance is based on color alone, and so as in Sec. 4.4, we write a pairwise distance function. Let

$$D_{i,j} = \|X_i - X_j\| = \sqrt{\sum_{l=1}^{d} (X_{i,k} - X_{j,k})^2}, \tag{4.6}$$

describe a matrix of distance function values across the sample of points, for distance function, $d(\mathbf{z}_i, \mathbf{z}_j)$, and $d : M \times M \to \mathbb{R}^+$. Next as done in many general spectral clustering methods, [61, 15] and as specialized to image segmentation [68, 61], a pairwise symmetric affinity matrix may be defined,

$$W_{i.j} = e^{-D_{i,j}^2/2\sigma^2}. \tag{4.7}$$

The value of $\sigma > 0$ may be chosen as a resolution parameter, as it emphasizes how to emphasize distance between measurements. It is convenient to emphasize the "practical" sparsity, by reassigning $W_{i,j} = 0$ if $W_{i,j} < \epsilon$ for a small threshold, $\epsilon > 0$. This can be interpreted as generating a weighted graph, $G = (V, E)$, where vertices $V = \{1, 2, ..., pq\}$ have edges between them whenever $W_{i,j} > 0$ and with weights accordingly.

A degree matrix, corresponding to the weighted symmetric directed graph is,

$$\mathcal{D}(i,i) = \sum_j W_{i,j}, \ \mathcal{D}_{i,j} = 0, i \neq j, \tag{4.8}$$

Shi and Malik [68] realized noted that the max-cut is equivalent to,

$$\min_x ncut(x) = \min_y \frac{y^T(\mathcal{D} - W)y}{y^T D y}, \tag{4.9}$$

as can be proved through the Courant-Fischer theorem, [8], and [68] for the image segmentation setting. This then brings us to the generalized eigenvalue eigenvector problem,

$$(\mathcal{D} - W)y = \lambda D y \tag{4.10}$$

where the second smallest eigenvalue and corresponding eigenvector solve the optimization problem. This could be written in terms of a symmetric normalized graph Laplacian, L, by noting that Eq. 4.10 transforms into,

$$\mathcal{D}^{-1/2}(\mathcal{D} - W)\mathcal{D}^{-1/2}x = \lambda x, \tag{4.11}$$

or,

$$Lx = \lambda x, \tag{4.12}$$

if,

$$L = \mathcal{D}^{-1/2}(\mathcal{D} - W)\mathcal{D}^{-1/2}, \tag{4.13}$$

by substitution,

$$y = \mathcal{D}^{-1/2}x. \tag{4.14}$$

The affinity matrix eigenvalue problem has an interpretation as a stochastic matrix eigenvalue problem, by [52, 51, 52],

$$P = \mathcal{D}^{-1}W. \tag{4.15}$$

Meila and Shi [52] noted that the affinity matrix W relates to random walks through a graph according to this stochastic matrix P, and this relates closely to a diffusive process underlying the diffusion map method, [15, 58]. This random walker interpretation connection between eigenvalues of P and W is reviewed further in Sec. 4.2.3

Now the smallest eigenvalue of Eq. 4.10 corresponds to the greedy partition (one element of the A-B partition is empty) so the second smallest eigenvalue corresponds to the Cheeger-balanced partition, the best bi-partition. Then one could proceed by recursively bi-partitioning [47]. We follow the concept of [61] which is to choose the k smallest eigenvalues *after the zero eigenvalue* and corresponding eigenvectors and then to cluster these by use of k-means from there. This is what we see in Fig. 4(b).

4.1.6 Motion Segmentation, and Directed Affinity, Following Along Measured Observations

Now we develop a directed affinity matrix $\mathcal{W}$ (note the change of font to distinguish from the symmetric counterparts W in Eq. 4.7). Replace X in Eq. 4.5 with,

$$X(t) = [X_{1,:}(t)^T|X_{2,:}(t)^T|...|X_{M,:}(t)^T], \tag{4.16}$$

where $X_{i,:}(t)$ denotes the column vector of d colors at $\mathbf{z}_i$, pixel location i, at time t in the movie sequence. Generally the colors at pixel i will be changing over time. Then let,

$$\begin{aligned} D_1(i,j,a,\tau) &= \sum_{l=0}^{\tau-1} \|X_i(t+la) - X_j(t+(l+1)a)\| \\ &= \sum_{l=0}^{\tau-1} \sqrt{\sum_{l=1}^{d} (X_{i,k}(t+(l)a) - X_{j,k}(t+(l+1)a))^2}. \end{aligned} \tag{4.17}$$

This compares the scene at pixel position i, through τ-time instances, $l = 0, a, 2a, ..., (\tau-1)a$, to the scene at pixel j through τ-time instances one step in the future, $l = 1, a, 2a, ..., \tau a$. Note that the norm, the inner sum, is the same as the color measuring norm in Eq. 4.6.

Now we measure the spatial distance between the pixels, as they appear naturally in the scenes represented by the figures. Let,

$$D_2(i,j)^2 = \|\mathbf{z}_i - \mathbf{z}_j\|^2 = (z_{1,i} - z_{1,j})^2 + (z_{2,i} - z_{2,j})^2, \tag{4.18}$$

where $\mathbf{z}_i = (z_{1,i}, z_{2,i})$ denotes the spatial coordinates of pixel number i. This is the standard spatial Euclidean norm.

Adding these two norms defines a spatial and time delayed color distance function,

$$D(i,j,a,\tau)^2 = D_1(i,j,a,\tau)^2 + \alpha D_2(i,j)^2. \tag{4.19}$$

Note that $D_1(i,j,0,0)$ is identical to the distance function in Eq. 4.6 used for image segmentation, but including $a > 0$, $\tau > 0$ allows us to consider motion segmentation and so coherency. Finally an affinity matrix follows,

$$\mathcal{W}_{i,j} = e^{-D(i,k,a,\tau)^2/2\sigma^2}. \tag{4.20}$$

Notice we have suppressed including all the parameters in writing $\mathcal{W}_{i,j}$, and that besides time parameters a and τ that serve as sampling and history parameters, together the parameters α and σ serve to balance spatial scale and resolution of color histories, and comparable to the role of σ in Eq. 4.7.

Contrasting W in Eq. 4.7 to $\mathcal{W}$ in Eq. 4.20 we see we the difference of symmetric versus generally asymmetric matrices reflecting the arrow of time. Such a difference is fundamental and naturally must be included in any concept of coherence. Clustering in this setting then reflects the concept of coherence, as a scene that retains its "appearance", but for now we continue with the idea that maintaining appearance is a sensible idea of coherence.

We proceed to cluster the system summarized by affinity $\mathcal{W}$ by interpreting the problem as random walks through the weighted *directed* graph, $G = (V, E)$ generated by $\mathcal{W}$ as a weighted adjaceny matrix. Stated equivalently, this is like a directed diffusion problem. See Sec. 4.2.3 for the comparable discussion in the symmetric case. So let,

$$\mathcal{P} = \mathcal{D}^{-1}\mathcal{W}, \tag{4.21}$$

where analogously to the symmetric case, $\mathcal{D}(i,i) = \sum_j \mathcal{W}_{i,j}$, $\mathcal{D}_{i,j} = 0, i \neq j$. So $\mathcal{P}$ is a row stochastic matrix representing probabilities of a Markov chain through the directed graph G, where,

$$\mathcal{P}_{i,j} = p(j(t+a)|i(t)), \tag{4.22}$$

and with this in mind, there is an interpretation of this directed graph partition by spectral methods as a naive-Bayes image classifier, by an unknown transfer operator, and we plan to pursue this perspective in the future; a similar observation that the symmetric diffusion map method relates to a Bayesian update has been made in [71].

We may cluster the directed graph by concepts of spectral graph theory for directed graphs, following the **weighted directed graph Laplacian** described by Fan Chung, [14], and a similar computation has been used for transfer operators in [25, 33] and as reviewed, [8]. The Laplacian of the directed graph G is defined, [14],

$$\mathcal{L} = I - \frac{\Pi^{1/2}\mathcal{P}\Pi^{-1/2} + \Pi^{-1/2}\mathcal{P}^T\Pi^{1/2}}{2}. \tag{4.23}$$

See discussion of the symmetric spectral graph theory in, Sec. 4.2.2-4.2.3, and the ncut problem solution standard description by Courant-Fischer theory, and how that adapts to this weighted directed graph Laplacian case, in Sec. 4.2.4. Note that $\mathcal{P}$ is row stochastic implies that it row sums to one, or stated as the right eigenvector is the ones vector, $\mathcal{P}\mathbf{1} = \mathbf{1}$, but the left eigenvector corresponding to left eigenvalue 1 represents the steady state row vector of the long term distribution,

$$u = u\mathcal{P}, \tag{4.24}$$

which for example if $\mathcal{P}$ is irreducible, then $u = (u_1, u_2, ..., u_{pq})$ has all positive entries, $u_j > 0$ for all j, or say for simplicity $u > 0$. Let Π be the corresponding diagonal matrix,

$$\Pi = diag(u), \tag{4.25}$$

and likewise,

$$\Pi^{\pm 1/2} = diag(u^{\pm 1/2}) = diag((u_1^{\pm 1/2}, u_2^{\pm 1/2}, ..., u_{pq}^{\pm 1/2})), \tag{4.26}$$

which is well defined for either $\pm$ sign branch when $u > 0$. The the first smallest eigenvalue larger than zero, $\lambda_2 > 0$ such that,

$$\mathcal{L}v_2 = \lambda_2 v_2, \tag{4.27}$$

allows a bi-partition, by,

$$y = \Pi^{-1/2}v_2, \tag{4.28}$$

by sign structure. As before, analogously to the Ng-Jordan-Weiss symmetric spectral image partition method, [61], the first k eigenvalues larger than zero, and their eigenvectors, can used to associate a multi-part partition, by assistance of k-means clustering these eigenvectors.

4.1.7 Numerical Results of Motion Segmentation by Time Directed Affinity and Spectral Partition

Now we show three example problems indicating the efficacy of the directed spectral partition method, from a directed affinity from Eq. 4.19-Eq. 4.20. These will be, the Cassini remotely observed movie of Jupiter, a local lake effect snow event, and a synthetic data set from the double-gyre, in that order.

4.1.8 Directed Spectral Partition: Jupiter

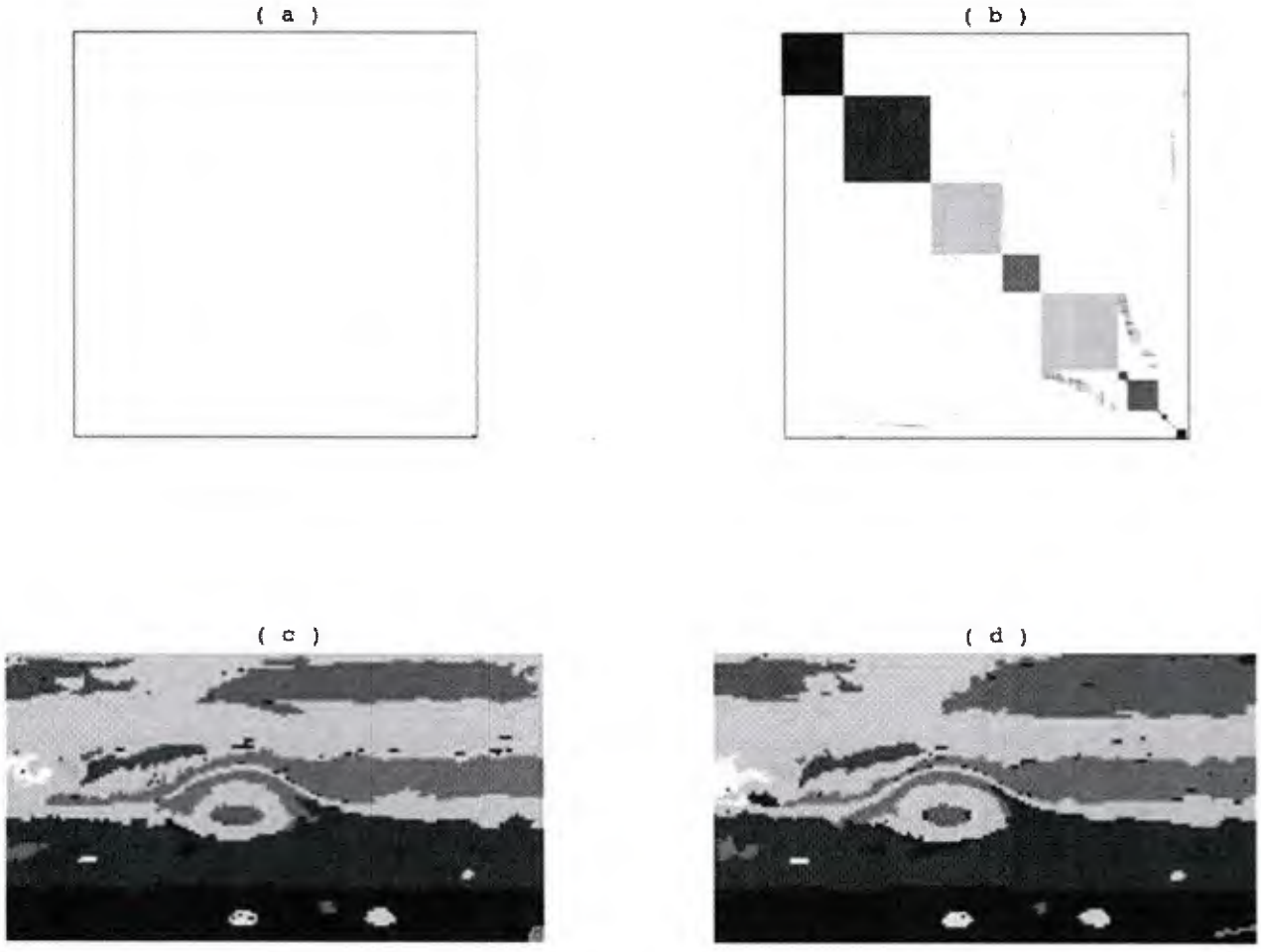


Figure 3: Given a small scene surrounding the Great Red Spot, and course grained (for ease of computation and clarity of presentation in this figure), (a) The affinity matrix, $\mathcal{W}$, Eq. 4.20, (b) Affinity matrix sorted according to spectral partition by methods of Sec. 4.1.6, Eq. 4.20-Eq. 4.28. (c) Coloring by each block of the sorted affinity matrix, partitions the scene according to regions that are found in multiple frames. (d) The partioned scene after $t = T$ time.

The results of partitioning using the directed affinity matrix $\mathcal{W}$ is shown in Fig. 3 from a scene of the GRS, and including the affinity matrix and a permutation that brings it to block structure as indicated by colors matching the colored scene. Fig. 4 again shows a scene of the GRS of Jupiter and its segmentation according to comparing the different methods of k-means to a single scene, a spectral method from a single scene, and finally our directed spectral method. We see that our method (d), the regions found by the directed method are most coherent in the sense of showing across time what is clearly visible in a movie, and perhaps difficult to fully appreciate in a static figure here.

The entire Cassini Jupiter data set is shown by directed spectral partition for coherence as shown in Fig. 5. Most notable are the banded longitudinal structures, the many circular vortex storms, and the largest being the GRS. Also of interest here, see Fig. 6 that shows a directed spectral partition of the entire data set from Jupiter, as shown from the northern pole. The longitudinal cloud structure can be seen in this global projection to rotate in a manner that reminds us of a twist map[53]. Note the GRS is seen in the 7 o'clock position o this figure.

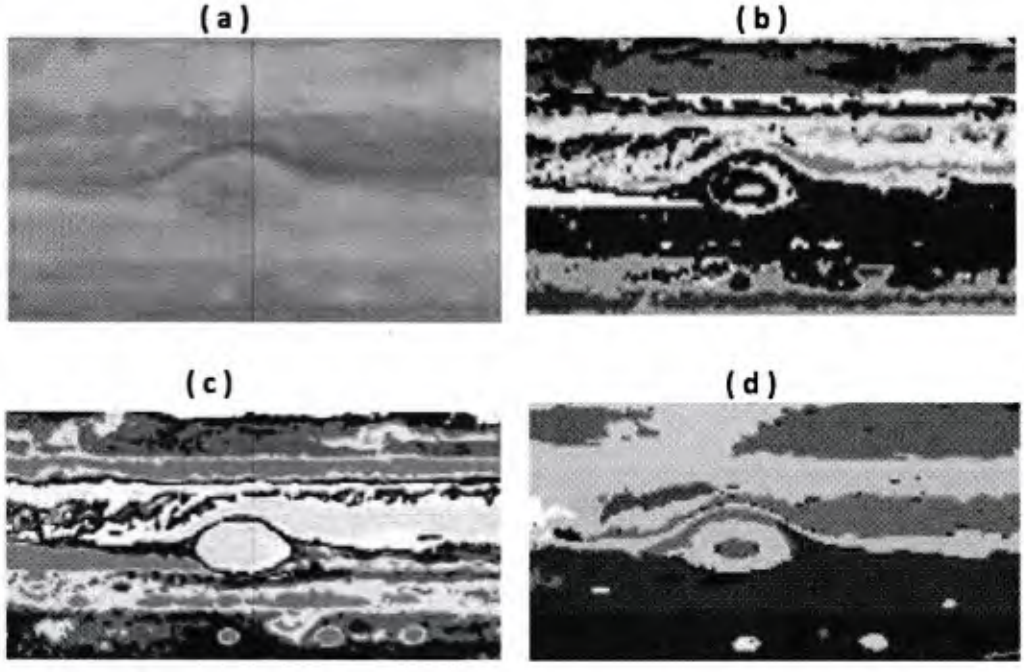


Figure 4: (a) A small scene surrounding the Great Red Spot, (b)A k-means clustering based on color only by affinity matrix. (c) Based on spectral partitioning with color alone affinity (d) Based on directed affinity matrix, as in Fig. 3.

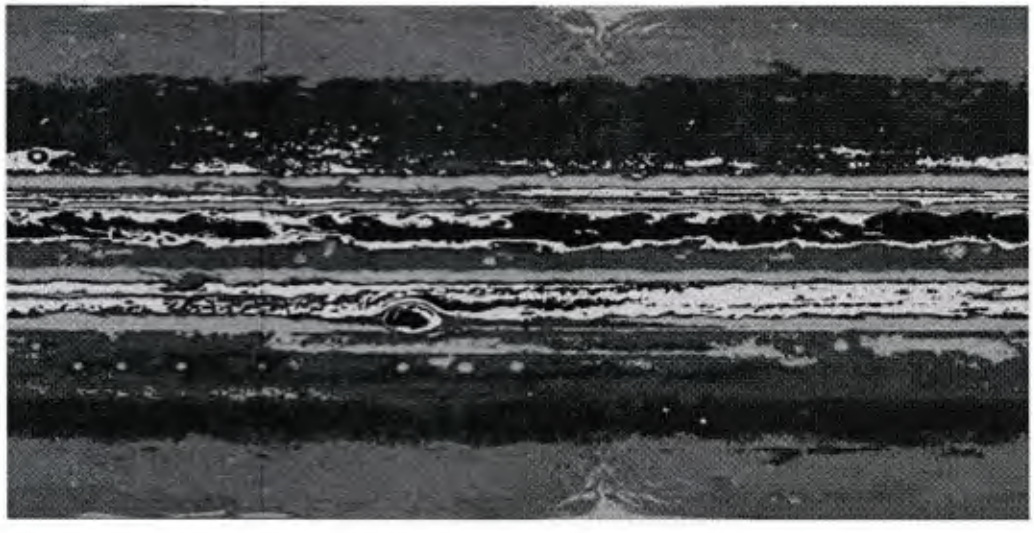


Figure 5: Directed spectral partition of Jupiter of the entire Cassini data set. Compare to Fig. 6.

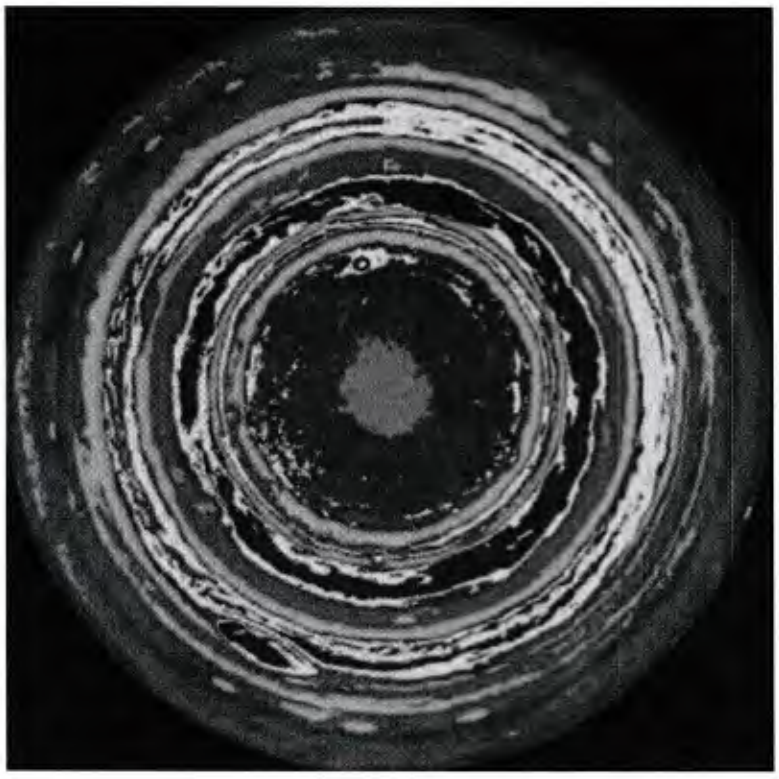


Figure 6: Directed spectral partition of Jupiter as shown on a projection as seen from the northern pole. Compare to Fig. 5.

4.1.9 Directed Spectral Partition: Lake Effect Snow

A lake effect snow is a common scenario during winter months nearby Clarkson University. It comes from energetic but cold air flowing across a relatively warmer expanses of water, in this case the Great Lake, Lake Ontario. These local storms are caused by the moist warm air rising into the cold air, that falls as snow nearby as the prevailing airmass sweeps over the colder downwind land. The hallmark of such an event is a storm that seems to be "parked", sometimes dropping snow for days in one locale, where even perhaps 50 miles away towns may enjoy sunshine. Generally, in such regions prone to the events, they happen many times each winter. They can be persistent and seemingly stationary, lasting for days and dropping massive amounts of snow in a highly localized event, such as for example 8 feet of snow dropped in 10 days over nearby Oswego, NY, in 2007, [21]. We analyzed one such nearby event for which we had convenient data from 2014. See Fig. 20-Fig. 8.

The methods herein successfully identify the lake-effect snow storm as a visually apparent coherent structure. In some ways, storms generally can be described as coherent structures, expressing energy. In this case of a lake effect snow, a particular interpretation is particularly interesting. This is a coherent set that is stationary, even though the underlying flow is advecting, strongly, from west to east. So clearly the coherent structure here is not the outcome of a purely advective process, or even an advective-diffusive process, as assumed in the formulation of most other studies of coherent structures. This one is more akin to the full system which is like an advective-reactive-diffusive process. The reactive part is due to the heat bath (literally) associated with the warm lake reacting with the cold advecting air and then later with the even colder land mass downstream. So what we see, and experience, is a derivative of all three aspects of the process. If we are stationary, say in a village in a region receiving 10 ft of snow in a week, then the coherent structure is a big deal, and very hard to miss, but not understood at all in terms of advection alone as normally described in this literature of coherent sets.

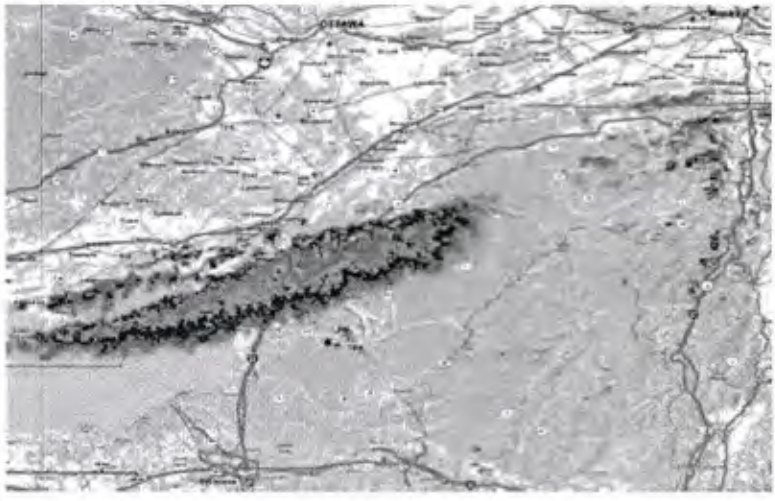


Figure 7: Lake effect snow on NEXRAD level III National Reflectivity Mosaic and Data radar imagery, near Watertown, NY indicating precipitation, of snow, during a period of intense lake effect snow during a 4-day period from Nov 18-21, 2014, shown here on 2014/11/18 at 12EST. The northern side of this localized storm skirts near Clarkson University, Potsdam, NY location of the writing of this article, and Syracuse, NY, Ottawa, ON and Montreal, ON are shown for map perspective. The striking feature of a lake effect snow event from the view of NEXRAD is the energetic snow that seems to stream off the lake but does not move downstream with other tracers and evidence that moves more so with the underlying advection of particles in the flow.

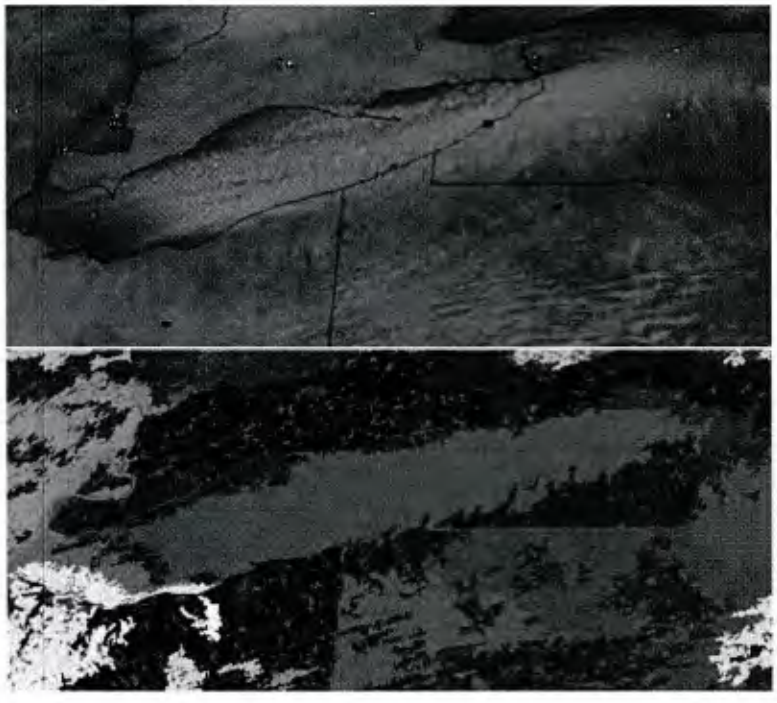


Figure 8: (Top) Lake effect Satellite Image from SSEC, For the same lake effect snow event shown in Fig. 20. (Bottom) Coherence based on a Directed Affinity Matrix.

4.1.10 Directed Spectral Partition: Double Gyre

The double gyre system as introduced by Shadden *et. al.* [67] has become ubiquitous [8, 62] as a benchmark for testing methods for finding or defining coherent structures. We take the standard version, as a nonautonomous Hamiltonian system,

$$\begin{aligned} \dot{x} &= -\pi A \sin(\pi f(x,t)) \cos(\pi y) \\ \dot{y} &= \pi A \cos(\pi f(x,t)) \sin(\pi y) \frac{df}{dx} \end{aligned} \tag{4.29}$$

with standard parameters, where $f(x,t) = \epsilon \sin(\omega t)x^2 + (1 - 2\epsilon \sin(\omega t))x$, $\epsilon = 0.1$, $\omega = 2\pi/10$ and $A = 0.1$. which has become a benchmark problem.

Taking a synthetic data set as a movie of evolving density, we deduce by methods herein clearly similar to many other studies of coherence in the double gyre,[26, 27, 38, 8, 48] by other transfer operator or geometric methods.

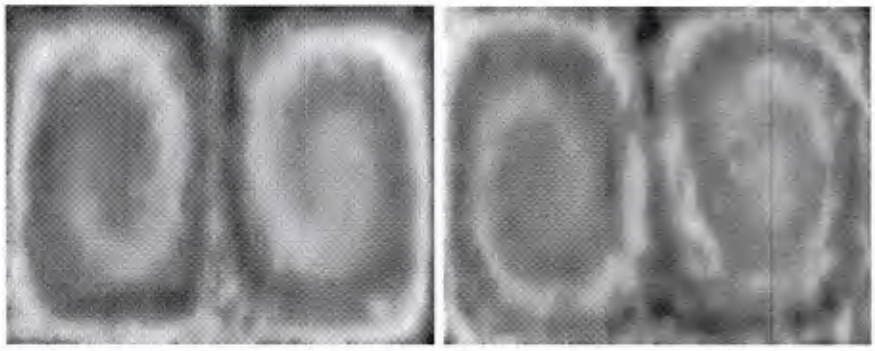


Figure 9: Double Gyre snapshots of evolving density profiles as a "movie" used to find coherent sets shown in Fig. 10.

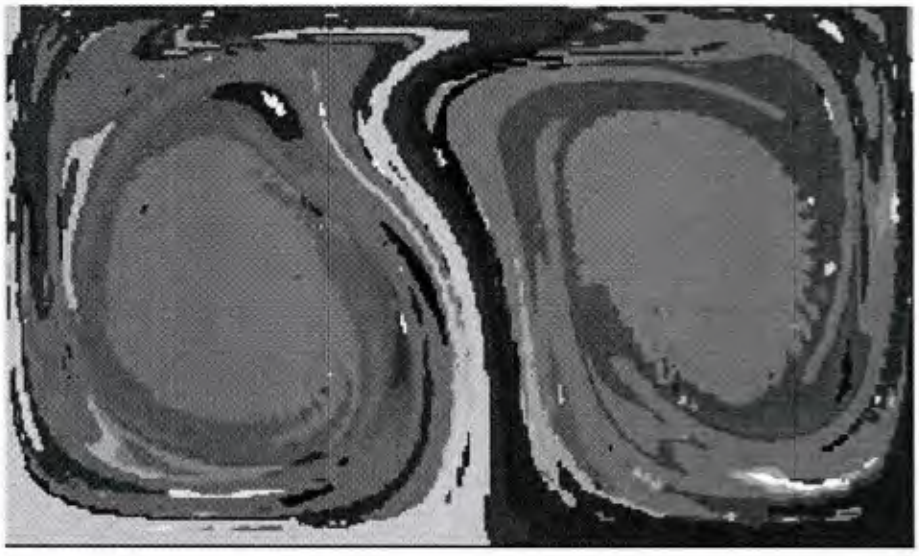


Figure 10: Double gyre Eq. 4.29, coherence based on directed spectral method shows left and right coherence structures comparable to other methods. Raw data is shown in Fig. 9.

4.2 Concluding remarks for Goal 1, Part 1

We have presented a perspective to infer coherence from remotely sensed "movie" data. This is inherently an Eulerian form of a data set since measurements (color) is always associated to a specific location, rather than Lagrangian measurements along orbits. However, most coherence discussions in the literature are formulated in terms of Lagrangian formulations. Our methods are inherently spectral in nature, and our details have aspects closely associated with the Meila-Shi spectral image partitioning, but these are also notably similar to the diffusion map methods. Since our problem has a definitive arrow of time as expected for the time varying aspect of a movie, then the standard symmetric requirement for a spectral method break down. Fortunately the directed graph version of spectral graph theory allows us to handle the weighted directed graphs that we deduce. We remark that the affinity matrix used in this discussion, $\mathcal{W}$ from Eq. 4.20, has an interpretation as a Naive Bayes observation of an unknown transfer operator, and a comparable interpretation of the symmetric version of such an exponential kernel as it appears in the diffusion map literature relates to a Bayesian method has been made and shown to yield useful data specification results, [71], and we plan to pursue this interesting feature in the future. We have remarked that there are aspects of this question that may associate with a Koopman mode interpretation, but our methods do not resemble the DMD modes analysis since thiose operators are deduced by a least squares optimization, whereas we hope in the future to interpret our operators as Bayes estimates that emphasize continuity in space and continuity in time.

Our examples have included especially demonstration that the storms and banded structures of Jupiter that are readily apparent by casual inspection. This simply reflects that our concept of coherence here is more so motion tracking, or motion segmentation than the coherence in literature. By this, we mean that coherency has mostly been associated with advective, or advective diffusive processes. However, as we illustrate with the lake effect snow data set, many processes also include a reactive aspect. As such, this particular data set demonstrates that the measured quantities, such as cloud cover, storm activity, and the like, can remain stationary even while the underlying advection part of the process may be strongly blowing past the process focus. Our methods should be contrasted to standard coherency since by motion tracking we are focusing on measured quantities of interest. We do not take this to be either a strength or a weakness of the standard methods our our own, but rather we just bring it forward as a point of interest. Perhaps there is a connection to the concept of Burning-Invariant-Manifolds (BIM) for reaction diffusion processes, [11, 55, 5] It is our hope that this work will serve as a useful direction to bring spectral methods and clustering methods from data-analytics to dynamical systems concepts of coherency as inferred from real data sets.

4.2.1 Appendix for Goal 1 Part 1

4.2.2 On nCut, The Symmetric Case

Given a graph $G = (E, V)$ generated under the assumption of a $n \times n$ symmetric weight matrix W, then a bi-segmentation of the n vertices of G is a bi-partition v and the compliment $v^c = V \backslash v$. Then the standard definition graph theoretic definition of an *ncut* is in terms of volumes of the weighted sets. Let,

$$vol(A) = \sum_{i \in A} D(i,i), \tag{4.30}$$

the total weighted degrees from the degree matrix Eq. (4.8). This has also been called, $assoc(A, V)$, [68]. The normalized cut of the partition, labelled $nCut(v)$ of the graph by $v \subset V$ is defined,

$$nCut(v) = (\frac{1}{vol(v)} + \frac{1}{vol(v^c)}) \sum_{i \in v, j \in v^c} W_{i,j}. \tag{4.31}$$

where,

$$cut(v) = \sum_{i \in v, j \in v^c} W_{i,j}, \tag{4.32}$$

denotes the total strength of edges between v and v^c. A "good" minimal ncut has relatively small weight between the two subsets but strong internal connections.

It can be shown [68] that if $x = \{-1, 1\}^n$ is an characteristic function for $v \cup v^c$ then the strong problem $\min_x ncut(x)$ has a relaxation, allowing $y_i \mathbb{R}$ rather than $x_i \in \{-1, 1\}$,

$$\min_v ncut(v) = \min_{y, y^T D1=0} \frac{y^T(\mathcal{D} - W)y}{y^T D y}, \tag{4.33}$$

Proof. The second part of the equality is a Raleigh quotient that is solved by the eigensystem,

$$(D - W)y = \lambda D y, \tag{4.34}$$

as related to,

$$Ly = \lambda y, \tag{4.35}$$

by Eq. 4.10-Eq. 4.14. Recall that from (3.10) we have:

$$D^{-1/2}(D - W)D^{-1/2}x = \lambda x \tag{4.36}$$

We see that the Laplacian matrix $L = D - W$ is symmetric positive semidefinite, which gives that $D^{-1/2}(D - W)D^{-1/2}$ is also symmetric positive semidefinite and its eigenvectors are pairwise orthogonal, and we can see that $x_0 = D^{1/2}1$ is eigenvector of Eq. 4.36 with $\lambda_0 = 0$ eigenvalue. Then, all other eigenvectors are orthogonal to x_0. Then:

$$x_1^T x_0 = y_1^T D1 = 0 \tag{4.37}$$

where x_1 is the second smallest eigenvector of Eq. 4.36 and y_1 is the second smallest eigenvector of Eq. 4.34. From the Courant-Fischer theorem [65] we have,

$$\lambda_1 = \min_{x \neq 0, x \perp x_0} \frac{x^T A x}{x^T x} \tag{4.38}$$

with $A = D^{-1/2}(D - W)D^{-1/2}$ from Eq. 4.36, then we have:

$$\lambda_1 = \min_{x \neq 0, x \perp x_0} \frac{x^T D^{-1/2}(D - W)D^{-1/2}x}{x^T x} \tag{4.39}$$

recall that $x = D^{1/2}y$, so we have:

$$\lambda_1 = \min_{y^T D1=0} \frac{y^T(D - W)y}{y^T D y}$$

□

The relationship of this problem to a random walk is discussed further in Eq. 4.2.3.

4.2.3 On Random Walks and Affinity

It has been shown [52, 51] that partitioning the graph $G = (E, V)$ generated in the case of a symmetric affinity matrix W has a random walk interpretation by developing the symmetric stochastic matrix, $P = D^{-1}W$. This relationship could be interpreted as a major idea behind the diffusion map method [15, 58].

The undirected graph corresponding to the symmetric W of Eq. 4.7 can be interpreted in a diffusion sense by $P_{i,j} = p(j|i)$ as the probability of a random walker moving to j from i. For pixels $(x_{k_1}, x_{k_2}, ..., x_{k_r})$ to be grouped as visited by random walkers in that order, according to W of Eq. 4.7, by P as a Markov

chain, we are asking what is $p(x_{k_2}, ..., x_{k_r}|x_{k_1})$ which equals $\Pi_{i=2}^{r} p(x_{k_{i-1}}, x_{k_i})$ by independence of jumps in a Markov chain. In a Markov chain with stochastic matrix P, then the eigenvalue problem $Py = \lambda y$ will have a largest eigenvalue $\lambda = 1$, and corresponding eigenvector $y = \mathbf{1}$ but the second eigenvector describes strongly connected sets [52] and Meila and Shi showed that minimizing the probability of diffusing between two sets equivalently to the ncut problem Eq. 4.33 of the W matrix, which is useful for connecting concepts of random walks and the spectral graph theory derivative of the graph Laplacian. The following theorem is supporting evidence relating the two problems.

Proposition 4.1 *If λ and y are eigenvalue-vector solutions of $Py = \lambda y$ then, $(1 - \lambda)$ is an eigenvalue of $(D - W)y = \lambda Dy$ and y is an eigenvector of $(D - W)y = \lambda Dy$.*

Proof. Since D is invertible, the proof follows immediately from,

$$\begin{aligned} (D - W)y &= \lambda Dy \\ D^{-1}(D - W)y &= D^{-1}\lambda Dy \\ (I - P)y &= \lambda y \\ Py &= (1 - \lambda)y. \end{aligned} \tag{4.40}$$

□

4.2.4 On Cuts and Directed Spectral Graph Theory

The problem with using the standard spectral graph theory for partitioning problems, to our scenario of motion tracking, and finding coherent sets, is that our affinity matrices yield not symmetric matrices. So the discussion in the previous two appendices is not directly applicable. Fortunately, there is a generalization that can handle our needs for a not symmetric cut problem, stated in Sec. 4.1.6. The Laplacian matrix of the directed graph from Fan Chung,[14], in Eq. 4.23 we repeat,

$$\mathcal{L} = I - \frac{\Pi^{1/2}\mathcal{P}\Pi^{-1/2} + \Pi^{-1/2}\mathcal{P}^T\Pi^{1/2}}{2}. \tag{4.41}$$

and from,

$$\mathcal{L} = I - \Pi^{-1/2}R\Pi^{1/2}, \tag{4.42}$$

where $R = \frac{1}{2}(P + \tilde{P})$, and $\tilde{P}_{i,j} = p_j P_{j,i}/p_i$ is the time reversed Markov chain so R is the reversiblization, also called detail balance, [8]. So analogous to the symmetric ncut problem, Eq. 4.33, a relaxed, not symmetric ncut problem, can be written, min $\frac{z^t \mathcal{L} z}{z^t z}$ subject to, $\sum z_i p_i^{\frac{1}{2}} = 0$. The solution of the optimization can be shown by the Courant-Fischer theorem,

$$\lambda_2 = \min_{z^t p^{1/2} = 0, z \neq 0} \frac{z^t \mathcal{L} z}{z^t z} = \min_{z^t p = 0, z \neq 0} \frac{\sum_{i,j}(y_i - y_j)^2 p_i P_{i,j}}{\sum_i y_i^2 p_i}, \tag{4.43}$$

attained by the eigenvector $z = v_2$, corresponding to λ_2 of $\mathcal{L}$, and where $y = \Pi^{-1/2}v_2$, and $\Pi = diag(p)$ in terms of the dominant eigenvector of P. And thus the spectral partitioning problem is translated to a min-max optimization problem, and for the not symmetric problem, this symmetrization allows the use of the main theoretical tool, the Courant-Fischer theorem that requires a symmetric matrix, as developed in [14], and reviewed in [8].

4.3 Regarding Goal 1: Part 2: Stretching and Folding in Finite Time

Material here associated with PI's publication, Tian Ma, Nicholas Ouellette, Erik M Bollt, "Stretching and Folding in Finite Time,'" Chaos 26, 023112 (2016).

4.3.1 Introduction

It is almost definitional to say that chaotic and turbulent flows mix efficiently: by a repeated process of stretching and folding material volumes [45], chaotic advection [3] leads to mixing that is significantly enhanced relative to diffusion alone [51]. Historically, this rapid mixing has been quantified and modeled using statistical metrics such as the relative dispersion of fluid elements [48] or effective diffusivities [52]. These kinds of tools capture the global effects of the fluid advection on mixing and can provide a simple, high-level parameterization of its effects that is useful for modeling.

But even though transport and mixing are sped up by chaotic advection on average, this enhancement is rarely uniform; rather, it mirrors the complex spatiotemporal heterogeneity of the flow itself [26]. Even in a flow that is in general strongly mixing, there may be regions that are separated from the rest of the flow by dynamical barriers, and which mix only weakly. In the terminology of dynamical-systems theory, these weakly mixing regions may be described as being elliptic, while those parts of the flow that mix strongly and transport material rapidly are hyperbolic. Since chaotic and turbulent flows are typically highly unsteady and aperiodic, however, the instantaneous elliptic or hyperbolic character of a particular region of the flow may change rapidly in time. Thus, any useful partitioning of the flow field into strongly and weakly mixing regions must identify not only places that are elliptic or hyperbolic but those that maintain their character over some macroscopic time window. Such regions may be said to be *coherent* [39], and are often referred to as coherent structures.

Analyzing flows from the standpoint of coherent structures may allow us to move past global descriptions of transport and mixing, such as overall effective diffusivities or purely statistical models, to local characterizations. Such local information is often a very important complement to its global counterpart; in an oil spill in the ocean, for example, one is more interested in knowing which areas of the coastline will be affected than in the net spreading rate of the oil [37, 6]. Due to this goal, then, a vast set of methods has been devised to detect and characterize coherent structures in general unsteady flows. We focus here on methods that use primarily Lagrangian information; that is, techniques that are based on the information contained in the trajectories of fluid elements, typically computed over some finite time that specifies the window in which we desire coherence. Broadly, one can classify these methods into those that attempt to find the boundaries that separate coherent structures from the rest of the flow [20, 43, 23, 2, 4, 24, 19] and those that try to find the structures themselves by set oriented methods, [15, 17, 7, 34, 6, 56, 16].

We recently developed a Lagrangian diagnostic that we term the finite-time curvature (FTC) [33], and that was designed to study shape-coherent sets (that is, those that retain their shape as they are advected) [32]. Here, we study the FTC field in more detail, and show that it contains significant further information beyond shape coherence that can be used to quantify mixing. We compare the FTC to the more commonly known finite-time Lyapunov exponents (FTLEs) often used to identify Lagrangian coherent structures (LCSs), and show that the two contain similar and complementary, though not identical, information: FTLE can be thought of as the local propensity for stretching in a dynamical system, while FTC (appropriately extended to account for finite scale) additionally captures the local propensity for *folding*. Although stretching and folding often take place nearby each other in similar places, their spatial distribution is not identical [26]; folding, as a nonlinear process, often occurs in areas where stretching is large, although the converse is not true.But since it is axiomatic that both stretching and folding are necessary for chaos [10, 47, 1, 7], both kinds of information are needed to characterize chaotic mixing fully, and it is natural to ask about folding as well as stretching. Thus, the finite-scale FTC we introduce here usefully extends the toolbox that can be used to study mixing in complex flows from the Lagrangian standpoint. We note that a clear mathematical definition of folding is lacking in

this context. Here, we argue that since the FTC defined in Eq. (4.49) describes the propensity for a material curve to develop curvature as time evolves, it is a sensible indicator of what is meant by folding in general terms. In dynamical systems, "folding" is often associated with the Smale horseshoe [54], which may be taken as a complete fold of a region doubly across itself when a full shift symbolic dynamics results [27, 46], and "less folding" or double covering results in a subshift [27, 8, 7]. In the context of fluid dynamics, Theodorsen's horseshoe vortex paradigm follows from constructing a model of individual hairpins [49, 9, 53, 57, 58]. Even a bending in two directions, such as an inflection point, can eventually lead to horseshoes, as depicted in Fig. 1 of Ref. [5].

We begin below by reviewing the FTLE in Section 4.3.2. In Section 4.3.3, we review the definition of the FTC, and generalize our previous methods to account for the effects of finite scale and resolution. We then illustrate our methods using experimental data from a quasi-two-dimensional laboratory flow in Section 4.3.7. Finally, we summarize our results in Section 4.3.8.

4.3.2 Review of Finite Time Lyapunov Exponents

The finite-time Lyapunov exponent (FTLE) is a measure of local stretching in the flow. FTLEs are commonly used as indicators for hyperbolic LCSs [41], since in hyperbolic regions of the flow, stretching is locally maximal. Even though FTLEs have shortcomings for the detection of LCSs (primarily that they also detect regions of high shear [19]), they remain perhaps the mostly widely used tool for studying Lagrangian coherence, particularly in experiments where more detailed information may not be available. Here, we briefly review the definition and properties of FTLEs, so that we may later compare and contrast them with FTC.

Given a velocity field $\mathbf{u}(\mathbf{x},t)$ on a manifold $M \subset \mathbb{R}^d$, the trajectories $\mathbf{x}(t)$ of fluid elements are solutions of

$$\frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x},t), \tag{4.44}$$

where $\mathbf{x} \in M$ and $\mathbf{u}(\mathbf{x},t)$ is at least $C^2(M)$. Integrating this equation yields the flow map $\phi_t^{t+\tau}: \mathbf{x}(t) \mapsto \mathbf{x}(t+\tau)$. The finite-time Cauchy–Green strain tensor of the velocity field along the trajectory $\mathbf{x}(t)$ is given by the the symmetric, time-dependent, $d \times d$ matrix

$$J_\tau = \left[\frac{d\phi_t^{t+\tau}\mathbf{x}(t)}{d\mathbf{x}}\right]^* \frac{d\phi_t^{t+\tau}\mathbf{x}(t)}{d\mathbf{x}}, \tag{4.45}$$

where A^* denotes the adjoint of A. In the following, we assume that $d = 2$, but this assumption is not mathematically necessary.

If over a finite time interval $[t, t+\tau]$, the minimum and maximum eigenvalues $\lambda_{\min}(\tau)$ and $\lambda_{\max}(\tau)$ of J_τ satisfy the condition

$$\ln \lambda_{\min}(\tau) < 0 < \ln \lambda_{\max}(\tau), \tag{4.46}$$

then the canonical local material advective behavior is described by the evolution of circles into ellipses, where the major axis of the ellipse lies along the direction of instability. This condition implies that there is compression in one direction and expansion in the other along the trajectory. Such a trajectory in a time-dependent velocity field is referred to as a *hyperbolic trajectory*.

Recall that the spectral norm of the Jacobian $\frac{d\phi_t^{t+\tau}\mathbf{x}(t)}{d\mathbf{x}}$ is given by

$$\left\| \frac{d\phi_t^{t+\tau}\mathbf{x}(t)}{d\mathbf{x}} \right\|^2 = \lambda_{\max}(\tau).\phi_t^{t+\tau} \tag{4.47}$$

Then the FTLE, which represents the maximum stretching at the point $\mathbf{x}(t)$ along the trajectory over a time τ, is given by

$$\sigma_\tau(\mathbf{x}(t)) = \frac{1}{|\tau|} \ln \left\| \frac{d\phi_t^{t+\tau}\mathbf{x}(t)}{d\mathbf{x}} \right\| = \frac{1}{|\tau|} \ln \sqrt{\lambda_{\max}(\tau)}. \tag{4.48}$$

It is common to define the repelling and attracting LCSs to be the ridges of the FTLE that are local maxima in space when computed in forward time ($\tau > 0$) and backward time ($\tau < 0$), respectively [21, 50]. Stated in terms relevant to our discussion here, the intensity of the FTLE field in a flow describes the propensity of ensembles of fluid elements that are nearby at time t to separate as the flow evolves.

Recall that, generally, a linear transformation of a circle yields an ellipse [18, 1]. See Fig. 12. Then the positive semi-definite matrix J_τ in Eq. (4.48), as the strain formed from the variation of the flow $\frac{d\phi_\tau \mathbf{x}(t)}{d\mathbf{x}}$ along the orbit $\mathbf{x}(t)$, can be shown [18] to have eigenvectors pointing along the major axis of the ellipse, and correspondingly the singular values are the lengths of the major axis corresponding to the growth rate of errors. We will derive further interpretation of this stretching picture of infinitesimal action along orbits below, as related to infinitesimal folding. Note that the finite-size Lyapunov exponent (FSLE) [12, 35] is conceptually similar to the FTLE, but is computed slightly differently by measuring the time it takes for nearby particles to separate by some fixed amount. This computation is popular experimentally, especially in oceanography, since one can use discrete tracers such as buoys rather than densely seeding the system with particles on a fixed grid. However, it has been shown that, although related, the FSLE can be somewhat different from FTLE, particularly depending on details such as the numerical implementation and the particle seeding [42].

4.3.3 Review and Computation of Finite Time Curvature

In our recent work, [33] we developed a direct Lagrangian measure of "curvature propensity," which we termed the maximum finite-time curvature (maxFTC, or simply FTC), and which we used to infer the location of shape-coherent sets [32]. Here, we show that there is further structure to be found in the same computed field.

The idea of a shape-coherent set is that the general concept of "coherence" can be interpreted independently of the more commonly discussed direct notions of transport. That is, a set that maintains its shape to a high degree along its orbit can in a visceral way be considered to be coherent (specifically shape-coherent, in our terminology). We defined a measurement of this concept as $\sup_{S(A)} \frac{m(S(A) \cap \phi_T(A))}{m(B)}$ [32], where S is the set of all rigid body motions (that is, translations and rotations). Thus, for a given set A, we measured how closely a simple translation and rotation of A could be best matched to the true nonlinear flow $\phi_t^{t+\tau}$ of A. When the measured relative overlap can be made significantly close to unity, the set A is called "shape coherent". With this notion in hand, simple geometric reasoning allows one to observe that the way to maintain shape is to maintain the curvature of the boundary of the set. Thus, investigating regions of space that have a very low propensity to change curvature may reveal such sets. This idea is closely related to the classical concept of curve congruence by matching curvature [11]. Thus, a shape-coherent set defined in this way is meant to correspond to a set that mostly holds together under the flow. In Fig. 11, we have illustrated such a set (labeled a, in the figure) in contrast to two other sets that have changed their shape. In [32], we recalled a tenet of differential geometry, that the fundamental theorem of curvature indicates that the boundary curve can remain the same in time if and only if the curvature does not change. We showed that this idea can be made into a regularity theory that small changes in curvature correspond to small changes in shape coherence, and, conversely, that large changes in curvature are required to cause a set A to lose its shape coherence significantly. Note that other recent work studying other concepts of coherence have included figures comparable to Fig. 11 but with different back stories; in [22] Fig. 1, two advected curves are described as remaining "coherent if an initially uniform material belt ... around it shows no leading-order variations in stretching after advection," whereas relaxing a bit in Fig. 1 in Ref. [14] sets are described as coherent if they have boundaries that are small relative to their volume and that stay small. Both of these are stronger requirements as shape coherence allows arbitrary shapes as long as they have slowly changing boundary curvature.

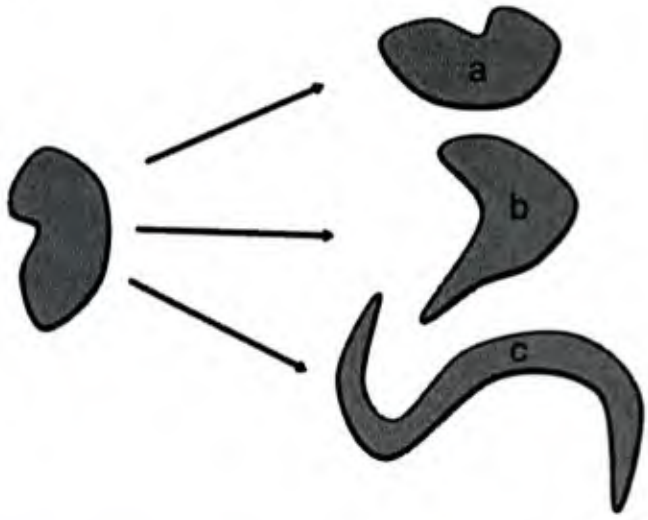


Figure 11: A schematic of shape coherence. The idea of shape coherence for a given set (left) is that it should evolve under the nonlinear flow $\phi_t^{t+\tau}$ in a manner that is nearly a rigid body motion, namely one composed of translation and rotation. Case a is such a scenario, where as cases (b) and (c) both develop significant stretching and folding. In these later two cases, the curvature of the boundaries has changed from the original, most significantly at the points near where the arrows point.

We therefore defined the maximum finite-time curvature (maxFTC) $K_{t_0}^{t_0+\tau}(\mathbf{z})$ for a point $\mathbf{z}$ in a plane $M \subset \mathbb{R}^2$ under a flow $\phi_{t_0}^{t_0+\tau}$ over the time interval $[t_0, t_0+\tau]$ to be

$$K_{t_0}^{t_0+\tau}(\mathbf{z}) = \lim_{\varepsilon \to 0} \sup_{\|\mathbf{v}\|=1} \kappa(\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))), \tag{4.49}$$

where

$$l_{\varepsilon,\mathbf{v}}(\mathbf{z}) := \{\hat{\mathbf{z}} = \mathbf{z} + \varepsilon s\mathbf{v}, |s| < 1\}. \tag{4.50}$$

Here, $\mathbf{v}$ is a unit vector and κ is the curvature. See Eq. (4.61). So, $l_{\varepsilon,\mathbf{v}}(z))$ is a small line segment passing through the point $\mathbf{z} = (x, y)$, when $\varepsilon \ll 1$ (although note that the limit as $\epsilon \to 0$ and the computation of curvature cannot in general be interchanged). The set of points $\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))$ yields a set of curvatures $\kappa(\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z})))$ that is part of the supremum in Eq. 4.49. Interpretations of different estimates of this set are detailed below.

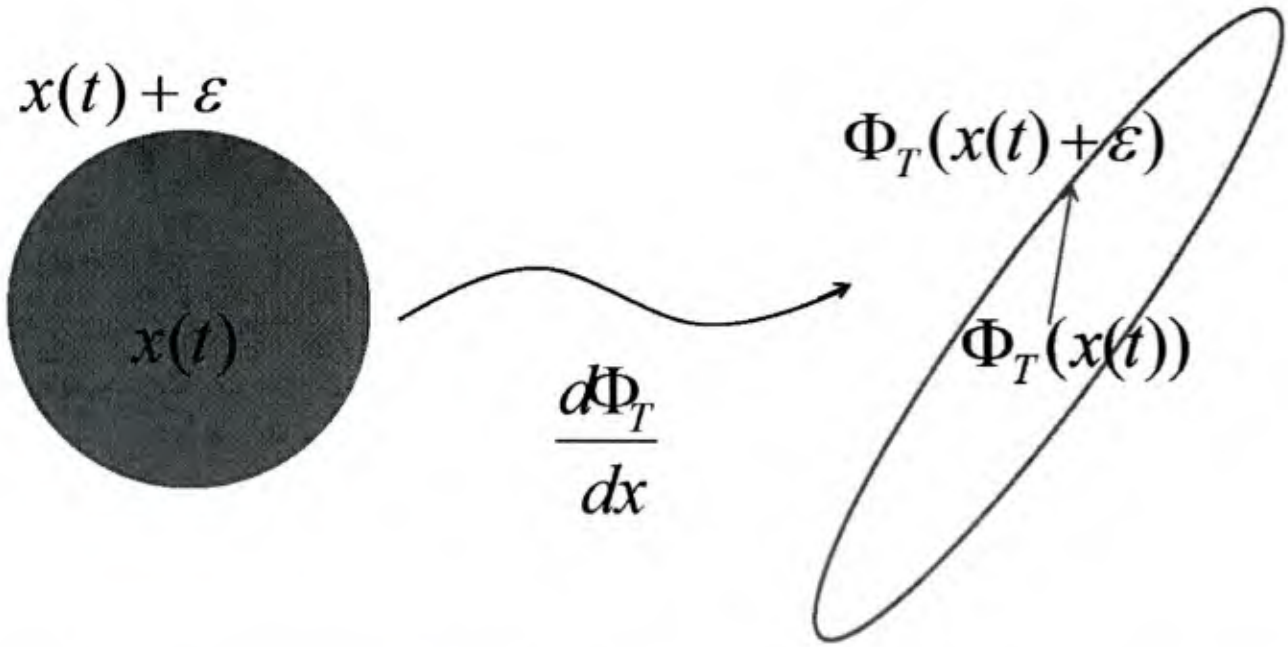


Figure 12: On the finest scale, a small material circle of radius ε along the orbit of a point $\mathbf{x}(t)$ evolves to an ellipse, and this estimate describes that the linearity dominates. However the geometry of the intermediate scale where this estimate differs from observation is described in Fig. 14, and the suggestion of Fig. 22 is that this transition between fine and intermediate scale may occur at smaller length scales than suggested by the common practice of studying only the linear terms.

In practice, it is a computational challenge as to how we implement both the optimization step and the limit step ($\varepsilon \to 0$) in Eq. (4.49). As we discuss further below, on the finest spatial scales there is a strong correlation between FTC and FTLE; on intermediate scales, however, they differ, as the FTC can uncover the nonlinearities in the flow. To distinguish these "finest" scales and "intermediate" scales precisely, we choose an $\varepsilon > 0$ and define a Finite-Scale Finite-Time Curvature (fsFTC) as

$$K_{t_0}^{t_0+\tau}(\mathbf{z},\varepsilon) = \sup_{\|\mathbf{v}\|=1} \kappa(\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))). \tag{4.51}$$

In contrast to the FTC in Eq. (4.49), notice that the limit is omitted, and the argument explicitly includes the independent variable $\varepsilon > 0$. Equations (4.49) and (4.51) both contain the expression $\kappa(\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))$, which corresponds to a curvature computation on each of the set of points on the line segment(s) $l_{\varepsilon,\mathbf{v}}(\mathbf{z})$. That is, $\kappa(\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z})) := \{\kappa\mathbf{w} : \mathbf{w} \in l_{\varepsilon,\mathbf{v}}(\mathbf{z})\}$ is a set of curvature values on which the sup is posed, across all line segments $l_{\varepsilon,\mathbf{v}}(\mathbf{z})$ in all orientations $\mathbf{v}$. Since the limit in Eq. (4.49) is taken after the curvature computation, this effectively selects the curvature at $\mathbf{z}$ in the orientation in which it is maximal. We have written it in this manner with the limit so as to reflect the computational method for calculating the fsFTC described below in terms of estimates by small line segments. We emphasize here that we are most interested in small but not infinitesimal line segments; however, the full definition of the FTC in Eq. (4.49) is equivalent to measuring the maximal curvature of all advected line segments through the base point $\mathbf{z}$ at $\phi_{t_0}^{t_0+\tau}(\mathbf{z})$. Furthermore and generally, for smooth flows the sup should be realized, and so can be replaced by a max operation. Note that the image of the line segment $l_{\varepsilon,\mathbf{v}}(\mathbf{z})$ is generally a curve $\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))$, and the role of the limit is to isolate the curvature to the locality of the image of the center point, at $\phi_{t_0}^{t_0+\tau}(\mathbf{z})$. In practice, the finite scale implicit in where and how the curvature is estimated in the neighborhood of $\phi_{t_0}^{t_0+\tau}(\mathbf{z})$ makes a difference for what is observed.

4.3.4 FTC on Intermediate Scales

The most straightforward interpretation of the estimation of the FTC by the fsFTC formula in Eq. (4.51) is by what we may refer to as the brute-force computation. First choose n sample vectors $\mathbf{v}$ pointed around a unit circle and uniformly spaced, and scale these by a small but finite distance $\varepsilon > 0$ to form $\varepsilon\mathbf{v}$. The direction of each unit vector $\mathbf{v}$ is specified by an angle γ. Then form n triplets of points

$$A = \mathbf{z} - \varepsilon\mathbf{v}, B = \mathbf{z}, C = \mathbf{z} + \varepsilon\mathbf{v}, \tag{4.52}$$

approximating the line-segment instance of Eq. (4.50). Each can then be mapped forward under the flow to new locations

$$A' = \phi_{t_0}^{t_0+\tau}(\mathbf{z} - \varepsilon\mathbf{v}), B' = \phi_{t_0}^{t_0+\tau}(\mathbf{z}), C' = \phi_{t_0}^{t_0+\tau}(\mathbf{z} + \varepsilon\mathbf{v}) \tag{4.53}$$

after a finite time τ; see Fig. 14 for an illustration. It is a fact of geometry that there is a unique circle passing through any three (non-colinear) points A', B', C' called the Menger circle. Calling the radius of this circle $R_{A',B',C'}$, the Menger curvature [36, 29, 28] is defined as $\kappa_M(A', B', C') = 1/R_{A',B',C'}$. The Menger curvature may be computed by the convenient classical formula,

$$\kappa_M(A', B', C') = \frac{c}{2\sin\gamma}. \tag{4.54}$$

See Fig. 13. Here γ may be chosen as any one of the 3 angles of the triangle $A'B'C'$ described by the three points and c is the length of the corresponding opposite side. Note that order of labeling is not important, as the circle is uniquely defined by the three points.

The unit vectors $\mathbf{v} = (\cos(s), \sin(s))$ for each $s \in [0, 2\pi)$ specify points (A, B, C) and their images (A', B', C') in Eqs. (4.52)-(4.53) for each angle. We may then write $\kappa_M(s, \varepsilon) := \kappa_M(A', B', C')$. Referring to Fig. 14 for labeling, each of the n uniformly sampled points around a circle of radius ε, $s_i = ih$, $h = 2\pi/n$, $i = 0, 1, ...(n-1)$, yields a sample of the function by points $(s_i, \kappa_M(s_i, \varepsilon))$. See Fig. 14. From this finite sample we may estimate

$$K_{t_0}^{t_0+\tau}(\mathbf{z}, \varepsilon) \approx \max_i\{\kappa_M(s_i, \varepsilon)\}, \tag{4.55}$$

for fixed ε.

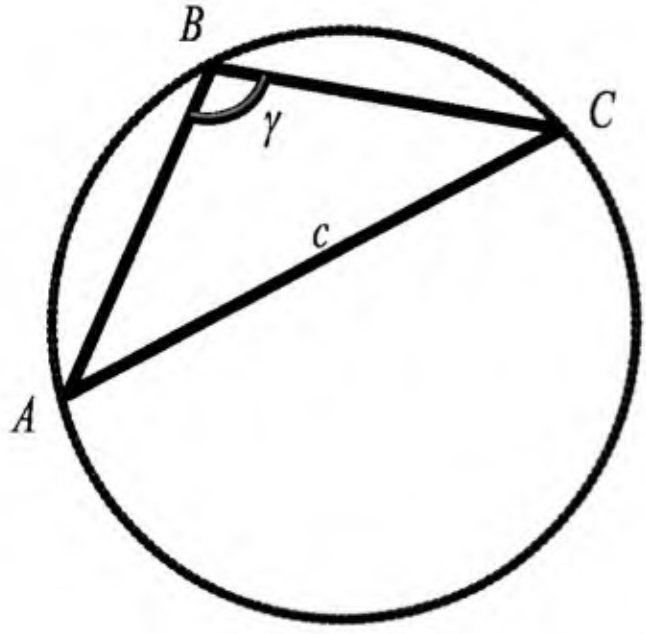
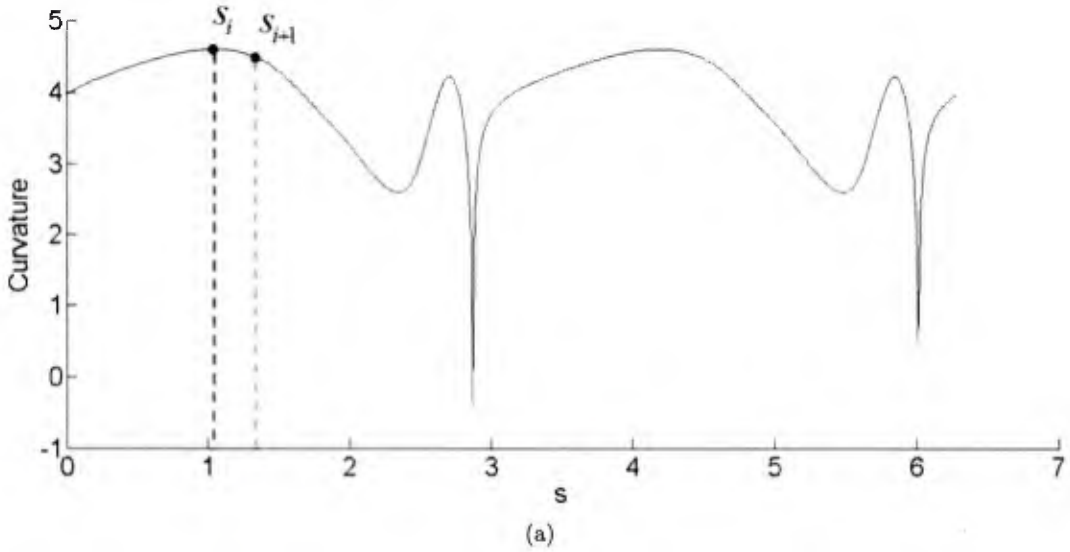
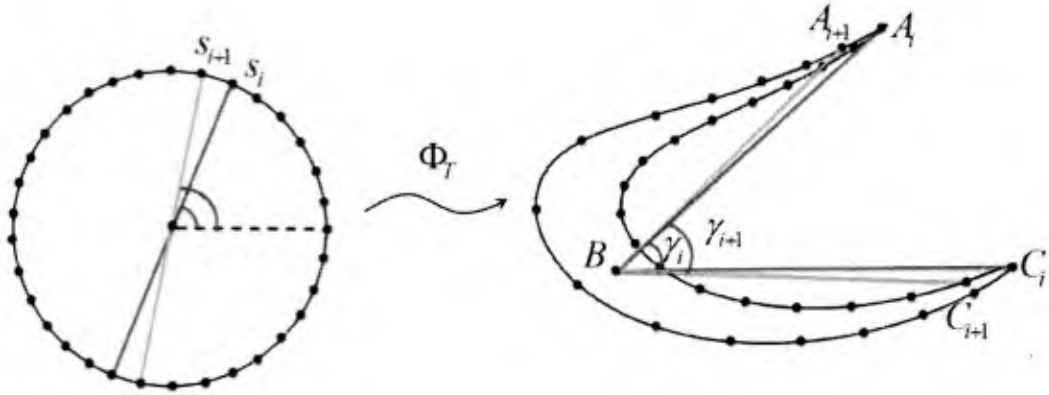


Figure 13: Menger curvature of three points A, B, C is defined in terms of the radius of the unique circle passing through these points by formulas, Eqs. (4.52) and (4.54).



(a)



(b)

Figure 14: Curvature due to sampling 3 opposing points along a line segment $l_{\varepsilon,\mathbf{v}}$ for $n = 33$ line segments with angular positions s_i uniformly around a small circle, according to Eqs. (4.52)-(4.53), and estimating curvature of the image by the Menger curvature by Eq. (4.54). See also Fig. 22. The intermediate scale gives rise to (a) curvature as a function of s, angle in radians around the circle, that differs significantly from the fine scale seen by the limit curve in Fig. 22, related to the prediction Eq. (4.61). (b) In this intermediate scale, the image of a circle is not simply an ellipse as suggested by Fig. 12, but rather there may be significant folding nonlinearities as shown here. In the regions of phase space where folding is prominent, this effect can be significant.

Proceeding more carefully, but still on a finite intermediate scale with $\varepsilon > 0$, we can estimate the fsFTC, which then will give us an estimate of the FTC. Unlike the above procedure, where we simply selected the maximal value from a large sampling, a more efficient and accurate way to optimize any function is to use a standardized optimization algorithm such as the gradient descent method [44]. This approach will work if the flow $\phi_{t_0}^{t_0+\tau}$ is sufficiently smooth at $\mathbf{z}$. It then follows that the fsFTC $K_{t_0}^{t_0+\tau}(\mathbf{z}, \varepsilon)$ can be estimated by the optimal Menger curvature $\mathbf{z}$ for a given small ε, so that

$$K_{t_0}^{t_0+\tau}(\mathbf{z}, \varepsilon) \approx \max_{s\in[0,2\pi)} \{\kappa_M(s, \varepsilon)\}. \tag{4.56}$$

where we We already noted that the image of the line segment $l_{\varepsilon,\mathbf{v}}(\mathbf{z})$ is the curve $\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))$; the Menger curvature $\kappa_M(A', B', C')$ is used here as an estimator of the curvatures in the neighborhood of $\phi_{t_0}^{t_0+\tau}(l_{\varepsilon,\mathbf{v}}(\mathbf{z}))$, assuming small ε and smooth $\phi_{t_0}^{t_0+\tau}$. See Fig. 14. The function $\kappa_M(s,\varepsilon)$ shown in Fig. 14a shows complex structure that is due to the finite, nonzero $\varepsilon > 0$. In contrast, as shown below, for the finest scales where $\varepsilon \ll 1$, $\kappa_M(s,\varepsilon)$ displays only simple periodic behavior of the ellipse with respect to s. We emphasize that this observation is not simply an issue with estimation for nonzero ε; as we demonstrate below, the transition between intermediate and fine scales may occur at extremely small ε compared with observable scales in experiments. More interestingly, however, we also show that we can use the inherent finite-scale nature of the fsFTC to infer information about the flow nonlinearities.

Finally in this section, we make a remark for clarity and contrast. There are different kinds of circles that are drawn here, with different meanings. Fig. 13 describes a general circle, relating to the inverse of the radius of an osculating circle (originally named the "circulus osculans," or kissing circle, by Leibniz) that touches the curve at a single point [13, 55], which is one classical way to begin the discussion of curvature (alternatively to beginning the discussion with the rate of change of the tangent vector with respect to arc length). In Fig. 13, we show how the Menger curvature [36, 29] is an estimator of curvature, particularly if A and C are close to B. Separately, the concept of FTC and fsFTC at a point $\mathbf{z}$ is described in terms of line segment through a point $\mathbf{z}$ such that when a material curve flows forward in time, this segment develops maximal curvature. Since the set of all such line segments of radius ϵ describes a disc, we draw the images of discs in Fig. 14. Since the curvature is taken before the limit in Eq. (4.49), we are indeed defining the curvature of the image of a material line, and to estimate this, we use an intermediate scaled triplet of points and the Menger formula. The circle in Fig. 14 is the most salient for our purposes here; the rest are drawn for geometric discussion and computation. Finally, there is an estimate of the curvature in Appendix 4.3.9 in terms of a small line segment through a point that may be at the edge of a circle centered on a nearby point, and from this we get an estimate on the fine scales.

4.3.5 FTC on the Finest Scales

For fine scales with $\varepsilon \ll 1$, small enough that the image of a circle of radius ε about $\mathbf{z}$ is essentially an ellipse with minimal non-affine distortion, the computation of FTC simplifies to an analytic expression we describe here. In this case, the smallness of the scale effectively linearizes the action of the flow, meaning that the FTC can no longer capture the propensity of the flow to fold curves. In fact, under the action of a linear flow, the image of a line is exactly a line. Hence, one may deduce that following small line segments by the above definitions would suggest there is zero propensity for the flow to develop curvature. However, by the discussion in the Appendix, and referring to Fig. 17, we show that under finite precision computation, where three points on a straight line are represented by three points that are not quite collinear, then some curvature is nonetheless measured. However, the curvature measured is in fact related to the linear flow of $D = \frac{d\phi_t^{t+\tau}\mathbf{x}(t)}{d\mathbf{x}}$.

Write correspondingly the strain as $J_\tau = D^*D$. We can apply a singular value decomposition to D, obtaining $D = U\Sigma V^*$. $\Sigma = diag(\sigma_1, \sigma_2)$ is the diagonal matrix whose nonzero entries are the major and minor axis lengths of the image ellipse of the unit circle shown; equivalently, they are eigenvalues of J_τ. $V = [v_1|v_2]$ is the orthogonal matrix whose orthogonal column vectors are oriented along the major and minor axes of the image ellipse, and are also the right eigenvectors of J_τ, which, without loss of generality, we have illustrated for convenience to be oriented along the x and y axis and we take $\mathbf{z} = 0$. We have included a point $\mathbf{w} = (\sigma_1, 0)$, without loss of generality at angle $s = 0$, along the major axis v_1. $\overline{\mathbf{w}} = (r\sigma_1, 0)$ is the corresponding point on the concentric ellipse shown. We also illustrate in Fig. 17 a red line segment of length 2δ that gives the Menger image with sides $\bar{\delta}$ through $\mathbf{w}$.

Let us now consider the development of curvature by the local dynamics as suggested by Fig. 17. The Menger curvature through the points on the ellipse shown in Fig. 17, $A' = (\sigma_1 r\cos(s), \sigma_2 r\sin(s))$, $B' = \overline{\mathbf{w}} = (r\sigma_1, 0)$, and $C' = (\sigma_1 r\cos(s), -\sigma_2 r\sin(s))$, estimates the curvature of the ellipse at $\overline{\mathbf{w}}$.

Specifically,

$$\lim_{s\to 0} \kappa_M(A', \overline{\mathbf{w}}, C') = \kappa(\overline{\mathbf{w}}) = \frac{\sigma_1}{r{\sigma_2}^2}. \tag{4.57}$$

The first part of this equality follows from the continuity property of Menger curvature: for smooth curves, if the three points $A', \overline{\mathbf{w}}, C' \to \overline{\mathbf{w}}$, then the Menger curvature limits to the curvature at the point. The computation of the second part of the equality is detailed in Appendix 4.3.9. In stating the next limit, notice that Eq. (4.58) differs from Eq. (4.57) in that $\overline{\mathbf{w}}$ and $\mathbf{w}$ are similarly positioned, but on concentric ellipses. Considering concentric circles of radius $r > 1$, the Menger curvature of the points shown in Fig. 17, $A' = (\sigma_1 r\cos(s), \sigma_2 r\sin(s))$, $B' = \mathbf{w} = (\sigma_1, 0)$, and $C' = (\sigma_1 r\cos(s), -\sigma_2 r\sin(s))$, are estimated by the curvature of the ellipse at $\overline{\mathbf{w}}$, since $\overline{\mathbf{w}} \to \mathbf{w}$ as $r \to 1$, so that

$$\lim_{\overline{\mathbf{w}}\to\mathbf{w}} \kappa_M(A', \mathbf{w}, C') = \kappa(\mathbf{w}) = \frac{\sigma_1}{{\sigma_2}^2}. \tag{4.58}$$

It then follows that if the flow $\phi_{t_0}^{t_0+\tau}$ at $\mathbf{z}$ is continuously differentiable, the FTC is estimated by the limit of curvatures as

$$K_{t_0}^{t_0+\tau}(\mathbf{z}) = \lim_{\varepsilon\to 0} K_{t_0}^{t_0+\tau}(\mathbf{z}, \varepsilon) \sim \max_{s\in[0,2\pi)} \{\kappa_M(s, \varepsilon)\} \sim \max\{\frac{\sigma_1}{\sigma_2^2}, \frac{\sigma_2}{\sigma_1^2}\}. \tag{4.59}$$

where the terms are, in order, the FTC, the fsFTC, the maximum Menger curvature evolved from center of the circle according to Eq. (4.56), and the maximum ratio of the singular values. If furthermore the flow is area-preserving, then $\sigma_1 = 1/\sigma_2$. Hence, in this special case, $\frac{\sigma_1}{\sigma_2^2} = \sigma_1^3$, $\frac{\sigma_2}{\sigma_1^2} = \sigma_2^3$, and so $K_{t_0}^{t_0+\tau}(\mathbf{z}) = \sigma_1^3$. The arguments supporting these statements in Appendix 4.3.9 are straightforward. We note that one can construct some special cases where the estimate in Eq. (4.59) does not perform well; however, such cases are not generic.

In the next section we interpret consequences both for the relationship of the study of curvature evolution to coherence, hyperbolicity, and the contrast of scales. Below, we will discuss then how it is often the intermediate-scaled folding that presents the interesting features.

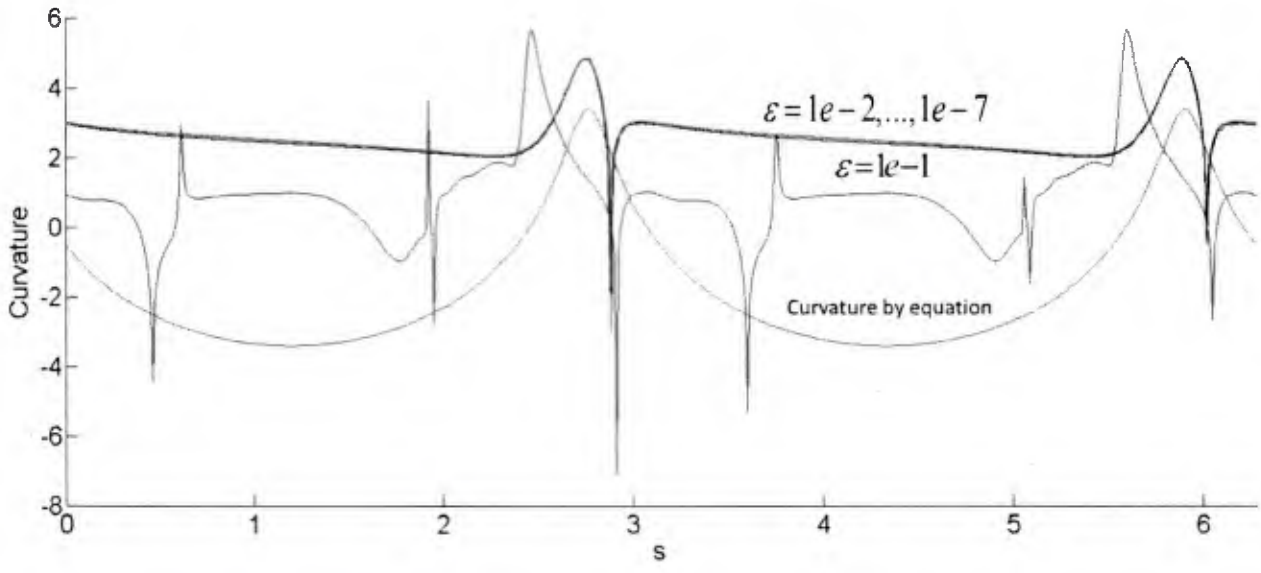


Figure 15: Curvature as a function of angle $0 \leq s < 2\pi$ around the circle, with ε decreasing in decades. Compare this picture to Fig. 14, where we show how the details of the fsFTC (Eq. (4.51)) involve estimates of the way in which small line segments $l_{\varepsilon,\mathbf{v}}$ (Eq. (4.50)) through the center point $\mathbf{z}$ each yield an estimated curvature by the Menger formula (Eq. (4.54)) in Eq. (4.55). Considering each line segment through the center point yields a curve of curvature values, one for each possible angle s specifying the rotational orientation of the initial line segment $l_{\varepsilon,\mathbf{v}}$. Here, $\varepsilon \approx 10^{-2}$ roughly corresponds to the transition between the intermediate and finest scales. The system used is the standard, non-autonomous double gyre, given by $\dot{x} = -\pi A \sin(\pi f(x,t)) \cos(\pi y), \dot{y} = \pi A \cos(\pi f(x,t)) \sin(\pi y) \frac{df}{dx}$, $f(x,t) = \epsilon \sin(\omega t) x^2 + (1 - 2\epsilon \sin(\omega t))x$ with parameters $A = 0.1$, $\epsilon = 0.1$, and $w = 2\pi/10$. The gyre was run for a time of $T = 20$, and the curves here correspond to an initial condition of $(0.1240, 0.1200)$. Note that we see convergence of the curvature estimates—the blue curves accumulating on the green—and that for maximal values, the red curve as described by values around a circle occur at maximal position suggested by the discussion in the Appendix.

4.3.6 Contrast of Scales

As noted above, the FTC should be interpreted in different ways depending on the spatial scale on which it is applied. Specifically, consider Fig. 22, where we plot profiles of the curvatures around the circle for decreasing $\varepsilon = 10^{-1}, 10^{-2}, .., 10^{-7}$. We indicate the transition that separates the intermediate scale from the fine scale at roughly $\varepsilon = 10^{-2}$, where the local circle of line segments exhibits significant folding, from the finest scales, where it shows the simpler curvature of ellipses behavior predicted by Eqs. (4.57)-(4.59),(4.61). Figure 22 illustrates the two somewhat different types of information contained in the FTC. To restate these two aspects more descriptively, *minimal* values of the FTC mark regions in the flow where material curves change spatiotemporally slowly. This feature of the FTC was the original reason we developed the analysis [33], as it can be used to locate shape-coherent sets [32]. The converse, however, also holds: *maximal* values of the FTC indicate regions where material curves most rapidly develop curvature, and thus the places in the flow where the shape of a set changes most rapidly.

How can we understand these opposite aspects in more familiar terms? The description in terms of shape coherence suggests that it is the elliptic-like behavior of the flow that causes the low troughs of the FTC field. Hyperbolic material curves, however, will cause any transverse curve of material to deform rapidly, suggesting that *ridges* of the FTC field likely indicate hyperbolic behavior. Thus, the single FTC computation can be used to locate both hyperbolic (ridges) and elliptic (valleys) regions of the flow field.

The analysis above in Eqs. (4.57)-(4.59) affirms that there is a strong relationship between FTC and

FTLE on fine scales, since both are related to the singular values of the Jacobian matrices along the flow. But, on the finest spatial scale, Eq. (4.59) shows that FTC is related to the *cube* of the FTLE. This feature has an important consequence: one would expect that ridges of the FTC field will be more prominent and will appear for shorter advection times than ridges of the FTLE field. We expect, heuristically even if not in detail, that this feature will be retained even on coarser scales since the general notion that hyperbolicity emphasizes deformation of material curves is generally understood by consideration of a transverse curve that samples the flow and becomes exponentially deformed. The more rapid appearance of structure in the FTC field makes it a promising candidate for forecasting applications, since less future-time information is needed to locate hyperbolic regions. Both complimentary aspects of coherence are revealed in a FTC field.

4.3.7 Experimental Demonstration

To illustrate the similarities and differences between the FTC and the FTLE in a practical example and to demonstrate that they are computable using real data, we measured both using experimental data from a quasi-two-dimensional laboratory flow. The details of this experiment have been described previously [25, 30, 31]. Briefly, we used a thin electromagnetically driven layer of an electrolytic fluid to produce nearly two-dimensional flow. The working fluid was a layer of salt water (16% NaCl by mass in deionized water) measuring $86\times86\times0.5$ cm^3 that rested on a glass plate coated with a hydrophobic wax. A square array of permanent neodymium-iron-boron magnets with vertical dipole moments and a lateral spacing of $L_m = 2.54$ cm lies beneath the glass; here, the magnets were arranged in stripes of alternating polarity. By running an electric current (of 1.25 A for the experiments analyzed here) through the salt water, we generated a Lorentz body force on the fluid that set it into motion. The Reynolds number $\mathrm{Re} = UL_m/\nu$, based on the in-plane root-mean-square velocity U, the magnet spacing, and the kinematic viscosity ν was 270.

We measured the velocity field using particle tracking velocimetry (PTV). The electrolyte was seeded with 50-μm-diameter fluorescent polystyrene microspheres; as the spheres are somewhat less dense than the electrolyte, they rise to its surface. To eliminate long-range surface-tension-driven forces between the floating particles, we floated a layer of pure water 5 mm deep on top of the electrolyte. We imaged the motion of the particles in the central 31.7×23.6 cm^2 (roughly 12.5×9 L_m) of the apparatus using a 4 megapixel IDT MotionPro M5 camera at a rate of 60 frames per second. We then tracked about 30 000 particles per frame using a multi-frame predictive tracking algorithm [40], and computed time-resolved velocities from the trajectories by convolving them with a smoothing and differentiating kernel [38]. Finally, we used the instantaneous velocities of all the tracked particles to create velocity fields by projecting the data onto the eigenmodes of a streamfunction [25], removing noise from the data and ensuring that the measured velocity fields are reliably two-dimensional.

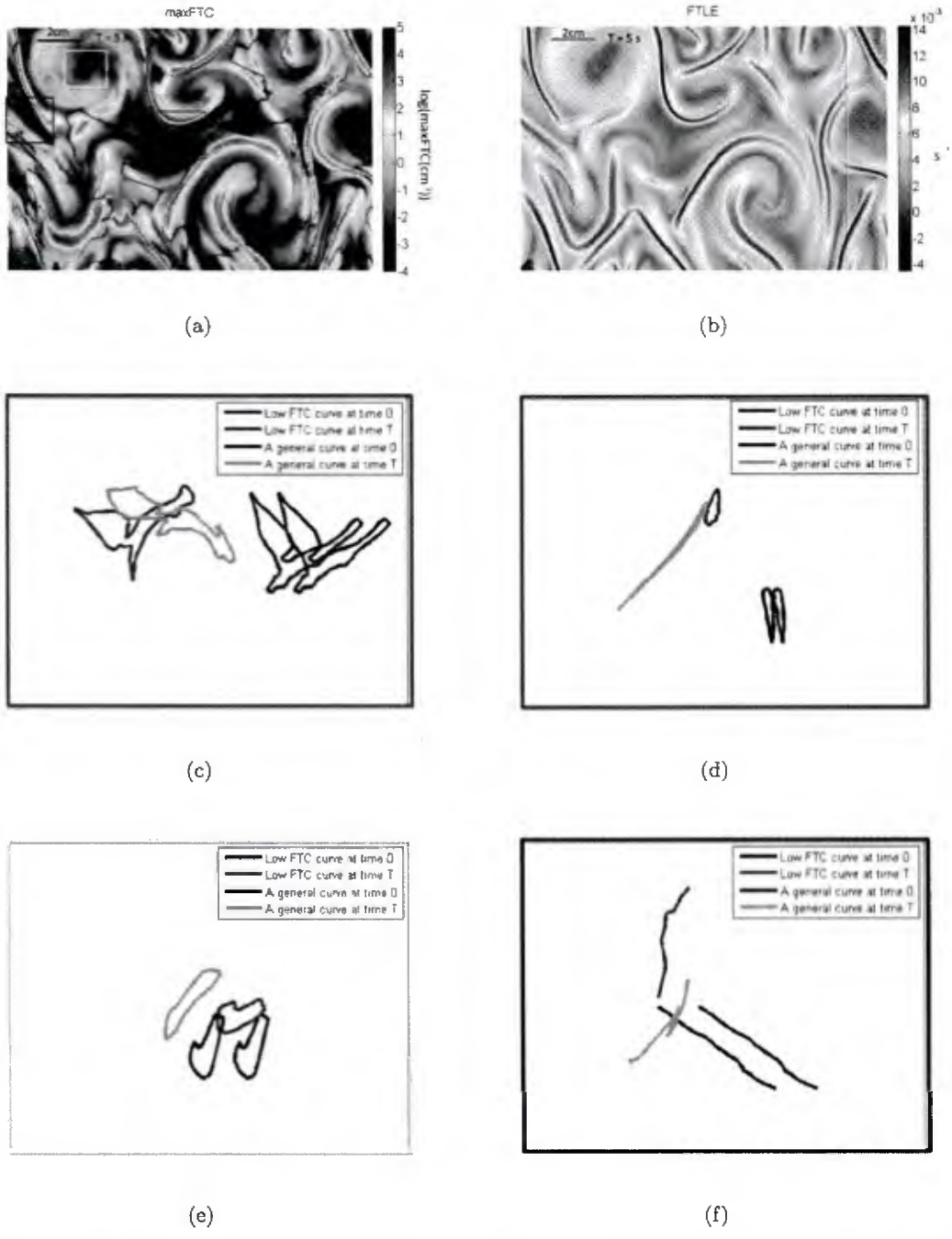


Figure 16: A (a) FTC field and a (b) FTLE field computed for the same data from the quasi-two-dimensional laboratory flow. Integration time for both fields is $5s \approx 3T_L$, where $T_L = L_m/U$ is the eddy turnover time. Note that in the FTLE and FTC many of the same features are seen in both methods of analysis, (cohabitating of the high FTLE ridges and FTC ridges), but if we look closely, there are many places where the two are telling different stories. Separately, since the FTC was designed to reveal shape coherence in the troughs of the FTC, then this figure is highlighting several such regions - boxed and blown up. So as initial, and after evolution, (blue then red), those troughs clearly reveal that the shape is mostly the same. This is a sensitive property in that even very nearby same-shaped starting curves (black) become highly deformed after the same time epoch shown (green). (c)-(f) demonstrate that low FTC troughs maintain high shape coherence but not general curves.

In Fig. 16, we show both the FTC and the FTLE computed over the same time window for the experiment. Even though it is high, the experimental resolution is still finite; thus, the FTC plotted here should be interpreted as the fsFTC, and we may expect it to display finite-scale features—most importantly, the folding propensity of the flow.

The FTLE (Fig. 16(b)) contains the features that are typically seen in this kind of flow. Most of the field is filled with relatively small, though still positive, values, indicating that the full field mixes chaotically, but relatively weakly. The FTLE also reveals very strongly stretching regions that are nearly co-dimension one; these line-like regions should approximately correspond to the Lagrangian coherent structures in the flow. In the cores of the areas in between the FTLE ridges, the FTLE values go negative, marking regions that are elliptic in character.

Comparing the FTLE with the fsFTC shown in Fig. 16(a) and (b), it is clear that the fsFTC reveals some of the same features, even though the FTLE captures only stretching while the fsFTC additionally captures folding. This rough spatial correspondence between strong stretching and strong folding is not surprising, and has been demonstrated before [26]. Here, however, for the same integration time (which was fixed for both panels in Fig. 16), these features are sharper for the fsFTC: the colormap for the fsFTC in Fig. 16(b) is on a logarithmic scale, while it is on a linear scale for the FTLE in Fig. 16(a). This kind of rapid convergence is very useful in experimental or observational applications, where long records of future flow-field information may not be available. But more than this simple difference, the fsFTC reveals additional structure. Regions where the fsFTC is small are not simply not stretching; they are also, by construction, *not folding*, and are thus evolving as semi-rigid bodies over the time window used to calculate the fsFTC. This kind of behavior should be contrasted with structures such as the Great Red Spot on Jupiter, where a macroscopic region of the flow is elliptic in that it is separated from the rest of the flow by a persistent transport barrier, but where the flow inside the barrier is still highly turbulent. For macroscopic areas where the fsFTC is small throughout, we expect that the internal flow is not only distinguished from the external flow, but is also evolving simply and not in a turbulent or chaotic fashion. Of additional note are the thin lines of low fsFTC values; these minimal curves of the fsFTC behave as is they are nearly rigid, and so they can rotate as time evolves but cannot bend.

4.3.8 Concluding Remarks of Goal 1, Part 2

Chaotic mixing can be schematically as a repeated process of the stretching and folding of material volumes; both of these processes are required to produce efficient mixing in a volume-preserving flow. By explicitly considering the effects of finite resolution on the finite-time curvature, we have shown here that this fsFTC naturally contains information about both stretching and folding. Thus, we have demonstrated that the FTC can be used to indicate more than the shape coherence it was designed for. We illustrated the utility of the fsFTC by computing both it and the more common FTLE on experimental data from a laboratory flow, showing that the fsFTC is simply computable and that it tends to reveal sharp features more rapidly than the FTLE. Thus, the fsFTC we have introduced is a valuable addition to the growing collection of Lagrangian methods that can be used to explore and characterize mixing and transport in complex flows.

4.3.9 Appendix for Goal 1 Part 2

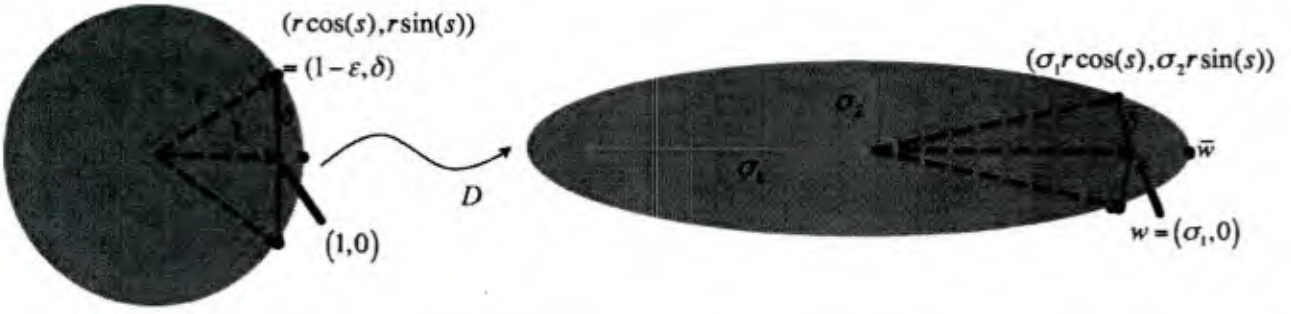


Figure 17: Two concentric circles of radius 1 and $r > 1$, each parameterized by angle s, and correspondingly the right triangle shown with sides, 1, δ and r. The center is shown as $(0, 0)$. In the appendix we follow the curvature growth of a "slightly perturbed" straight line (left shown in red) chosen perpendicular to the pre image of the dominant singular vector as shown, with coordinates $(1-\epsilon, \delta), (1, 0)$, and $(1-\epsilon, -\delta)$ and the image under a linear transformation $D = diag(\sigma_1, \sigma_2)$.

Here we prove the statement in Eqs. (4.57)-(4.58) regarding the ratio of singular values. It is well known that the image of a circle (that is, the set of all unit vectors) is an ellipse under the action of a general 2 by 2 matrix D. Curvature at a point is defined as the inverse of the radius of the osculating (tangent) circle to a curve at the point. Thus, a unit circle has curvature $\kappa = 1$ by definition. Now we will show that the growth of curvature of the circle will be related to the ratio of singular values according to Eqs. (4.57)-(4.58) and that also, the computed Menger curvature of "line segments" will give the same under the condition of 1) small line segments, and 2) small enough that in general configuration that a small amount of numerical imprecision will creep into the estimate. This second part is both always present in real computations, and necessary for the result, since in exact arithmetic, the image of a line is a line under a linear transformation.

Let the image of the circles shown in Fig. 17 be the ellipses shown, where without loss of generality the body axes are aligned with the coordinate axes for convenience as shown, and the major axes v_1, v_2 have lengths σ_1, σ_2 from the SVD $D = U\Sigma V^*$. The inner ellipse shown can be written implicitly as $\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} = 1$, or parametrically as

$$\gamma(s) =< \sigma_1 \cos s, \sigma_2 \sin s > . \tag{4.60}$$

By a standard computation of the curvature in terms of a parametrically represented curve, it follows that

$$\kappa(s) = \frac{|\gamma'(s) \times \gamma''(s)|}{|\gamma'(s)|} = \frac{\sigma_1 \sigma_2}{|\sigma_1^2 \sin^2(s) + \sigma_2^2 \cos^2(s)|^{3/2}}. \tag{4.61}$$

Solving $\kappa'(\hat{s}) = 0$ for the critical points $\hat{s}$ yields $\hat{s} = 0, \frac{\pi}{2}$, and hence the extrema of curvature are found by substitution into Eq. (4.61) to be

$$\kappa(\hat{s}) = \{\frac{\sigma_1}{\sigma_2^2}, \frac{\sigma_2}{\sigma_1^2}\}. \tag{4.62}$$

Likewise, the curvatures of the inner and outer ellipses at $\mathbf{w}$ and $\overline{\mathbf{w}}$, respectively, are $\kappa(\mathbf{w}) = \frac{\sigma_1}{\sigma_2^2}$ and $\kappa(\overline{\mathbf{w}}) = \frac{\sigma_1}{r\sigma_2^2}$. The limit in Eq. (4.57) follows immediately since the cosine and sine functions in the parametric equation of an ellipse (Eq. (4.60)) are continuous.

Alternatively, now consider the computation of the growth ratio of Menger curvature of the slightly bent line segment shown in Fig. 8. Let us label the three points as $P_1 = (1, 0), P_2 = (1-\epsilon, \delta)$, and $P_3 = (1-\epsilon, -\delta)$. For an exactly straight line, $\epsilon = 0$; however, due to finite precision arithmetic, we are considering $0 < \epsilon << 1$, and further, discussing a small line segment, we choose $\delta \ll 1$, but not nearly as

small as the arithmetic precision so that $\epsilon \ll \delta$. First we compute the Menger curvature of the (left) line segment shown by Eq. (4.54). Let us define the vectors

$$\begin{aligned} \mathbf{a} &= P_1 \vec{-} P_2 =< P_1.x - P_2.x, P_1.y - P_2.y >=< \epsilon, -\delta >, \\ \mathbf{b} &= P_1 \vec{-} P_3 =< P_1.x - P_3.x, P_1.y - P_3.y >=< \epsilon, \delta >, \\ \mathbf{c} &= P_2 \vec{-} P_3 =< P_2.x - P_3.x, P_2.y - P_3.y >=< 0, -2\delta > . \end{aligned} \tag{4.63}$$

Note that for simplicity of presentation we have chosen the same computational "error" in both the points P_2 and P_3, rather than separate errors ϵ_1 and ϵ_2. Our conclusions and results do not change, however, if $\epsilon_1 = \epsilon_2 = \epsilon$, as can be readily checked. Now, let $\gamma = \angle P_2 P_1 P_3$. By the standard definition of the cross product, $|(\sin(\gamma)| = \|\mathbf{a} \times \mathbf{b}/\|\mathbf{a}\|\|\mathbf{b}\|)\|$ (interpreting $\mathbf{a}$ and $\mathbf{b}$ as three-dimensional vectors with zero in the third component, as usual), and

$$\|\mathbf{a} \times \mathbf{b}\| = |\mathbf{a}; \mathbf{b}| = \begin{vmatrix} \epsilon & -\delta \\ \epsilon & \delta \end{vmatrix} = 2\epsilon\delta, \|\mathbf{a}\|^2 = \|\mathbf{b}\|^2 = \epsilon^2 + \delta^2, \tag{4.64}$$

we have

$$\sin(\gamma) = \frac{2\epsilon\delta}{\epsilon^2 + \delta^2}. \tag{4.65}$$

Then from the Menger formula (Eq. (4.54)) and substitution we have

$$\kappa_M(P_2, P_1, P_3) = \frac{\mathbf{c}}{2\sin(\gamma)} = \frac{\|\mathbf{a}\|\|\mathbf{b}\|\|\mathbf{c}\|}{2|\mathbf{a}; \mathbf{b}|} = \frac{(\epsilon^2 + \delta^2)}{2\epsilon\delta}\frac{2\delta}{2} = \frac{\epsilon^2 + \delta^2}{2\epsilon} \tag{4.66}$$

Remembering our assumptions that $\epsilon \ll \delta \ll 1$, $\kappa_M(P_2, P_1, P_3) \sim 0$, consistent with a small (almost) straight line. Under the linear transformation $D = diag(\sigma_1, \sigma_2)$ shown in Fig. 17, the three the points in Eq. (4.63) map to

$$\begin{aligned} \mathbf{a}' &= P_1' \vec{-} P_2' =< P_1.x' - P_2.x', P_1'.y - P_2.y' >=< \sigma_1\epsilon, -\sigma_2\delta >, \\ \mathbf{b}' &= P_1' \vec{-} P_3' =< P_1.x' - P_3.x', P_1.y' - P_3.y' >=< \sigma_1\epsilon, \sigma_2\delta >, \\ \mathbf{c}' &= P_2' \vec{-} P_3' =< P_2.x' - P_3.x', P_2.y' - P_3.y' >=< 0, -2\sigma_2\delta > \end{aligned} \tag{4.67}$$

and the angle at the image is $\gamma' = \angle P_2' P_1' P_3'$. Correspondingly,

$$\|\mathbf{a}' \times \mathbf{b}'\| = |\mathbf{a}'; \mathbf{b}'| = \begin{vmatrix} \sigma_1\epsilon & -\sigma_2\delta \\ \sigma_1\epsilon & \sigma_2\delta \end{vmatrix} = 2\sigma_1\sigma_2\epsilon\delta, \|\mathbf{a}'\|^2 = \|\mathbf{b}'\|^2 = \sigma_1^2\epsilon^2 + \sigma_2^2\delta^2, \sin(\gamma') = \frac{2\sigma_1\sigma_2\epsilon\delta}{\sigma_1^2\epsilon^2 + \sigma_2^2\delta^2}, \tag{4.68}$$

from which the Menger curvature of the image points becomes

$$\kappa_M(P_2', P_1', P_3') = \frac{\mathbf{c}'}{2\sin(\gamma')} = \frac{\|\mathbf{a}'\|\|\mathbf{b}'\|\|\mathbf{c}'\|}{2|\mathbf{a}'; \mathbf{b}'|} = \frac{(\sigma_1^2\epsilon^2 + \sigma_2^2\delta^2)}{2\sigma_1\sigma_2\epsilon\delta}\frac{2\delta}{2} = \frac{\sigma_1^2\epsilon^2 + \sigma_2^2\delta^2}{2\sigma_1\epsilon} \tag{4.69}$$

Finally, we compute the ratio of the growth of the Menger curvature using Eqs. (4.66) and (4.69):

$$\frac{\kappa_M(P_2, P_1, P_3)}{\kappa_M(P_2', P_1', P_3')} = \frac{\epsilon^2 + \delta^2}{2\epsilon} / \frac{\sigma_1^2\epsilon^2 + \sigma_2^2\delta^2}{2\sigma_1\epsilon} \sim \frac{\delta^2}{2\epsilon} / \frac{\sigma_2^2\delta^2}{2\epsilon\sigma_1} = \frac{\sigma_1}{\sigma_2^2}. \tag{4.70}$$

This last estimate in Eq. (4.70) again follows from the assumption $\epsilon \ll \delta \ll 1$, and hence we see that for small line segments, under finite precision, the limiting behavior of true curvatures of the circle evolving into the ellipse Eq. (4.62) is related to ratios of singular values.

In the case that $\epsilon_1 \neq \epsilon_2$, it is straightforward to check, modifying the position of the points in Eqs. (3pts)-(4.67) slightly, that the resulting curvature of the almost collinear points and its image are

$$\begin{aligned} \kappa_M(P_2, P_1, P_3) &= \frac{\sqrt{(\epsilon_1^2+\delta^2)(\epsilon_2^2+\delta^2)((\epsilon_1-\epsilon_2)^2+\delta^2))}}{2(\epsilon_1+\epsilon_2)\delta} \\ \kappa_M(P_2', P_1', P_3') &= \frac{\sqrt{(\sigma_1^2\epsilon_1^2+\sigma_2^2\delta^2)(\sigma_1^2\epsilon_2^2+\sigma_2^2\delta^2)((\sigma_1\epsilon_1-\sigma_1\epsilon_2)^2+\sigma_2^2\delta^2))}}{2\sigma_1\sigma_2(\epsilon_1+\epsilon_2)\delta} \end{aligned} \tag{4.71}$$

and that therefore the above computation, based on assuming $\epsilon = \epsilon_1 = \epsilon_2$, does not change the asymptotic result Eq. (4.70).

4.4 Regarding Goal 2: Questions Regarding Coherence and Transport in Turbulent Dynamical Systems

joe Material here associated with PI's publication, Jie Sun, Fernando Quevedo, Erik Bollt, "Statistical Inverse Formulation of Optical Flow with Uncertainty Quantification," Inverse Problems 34 105008 (2018)

4.4.1 Introduction

Optical flow reflects the visual motion between consecutive images. Determination of optical flow is important for applications ranging from machine learning and computer vision [94], artificial intelligence and robotics [76, 93], to scientific applications from oceanography to weather forecasting [78, 77, 80, 88, 89], to name a few. In classical approaches optical flow is determined by solving a variational optimization problem which requires the judicious choice of a regularization parameter for the problem to be well defined [73, 85, 101]. The choice of regularization parameter turns out critical for the satisfactory inference of optical flow. Despite the many (competing) methods of selecting the regularization parameter, none seems to be most "natural" comparing to the others [84, 96, 100]. Another important feature of classical optical flow approaches is that they produce a single solution (by design) as a "point estimate". In practice, the magnitude of the flow as well as measurement error and noise can vary significantly from one part of the images to another. Although not captured in classical optical flow, such "uncertainty" would provide valuable information if attainable as part of the solution.

In this work we develop a statistical Bayesian inversion approach for the inference of optical flow. In this framework the optical flow problem is reformulated from the Bayesian perspective as a statistical inversion problem. From this new perspective, different types of information can be naturally fused into a single posterior distribution, which is sampled using an appropriately designed Markov chain Monte Carlo scheme. Unlike the point estimate of optical flow obtained by classical variational approaches, the proposed statistical inversion approach is a methodological way of uncertainty propagation to produce a distribution of candidate optical flow fields from which various statistical properties can be extracted, including ensemble average computation and uncertainty quantification.

The rest of the paper is organized as follows. In Section 2 we develop a discrete approximation of the variational optical flow problem, giving rise to a finite-dimensional linear inverse problem. In Section 3 we reformulate such inverse problem from a Bayesian perspective as a statistical inversion problem and present an efficient algorithm to sample the resulting posterior distribution which is the key to determining the statistical properties of the underlying optical flow fields. In Section 4 we showcase the utility of our Bayesian approach on several benchmark examples and different noise levels. Finally, discussion and conclusion are presented in Section 5.

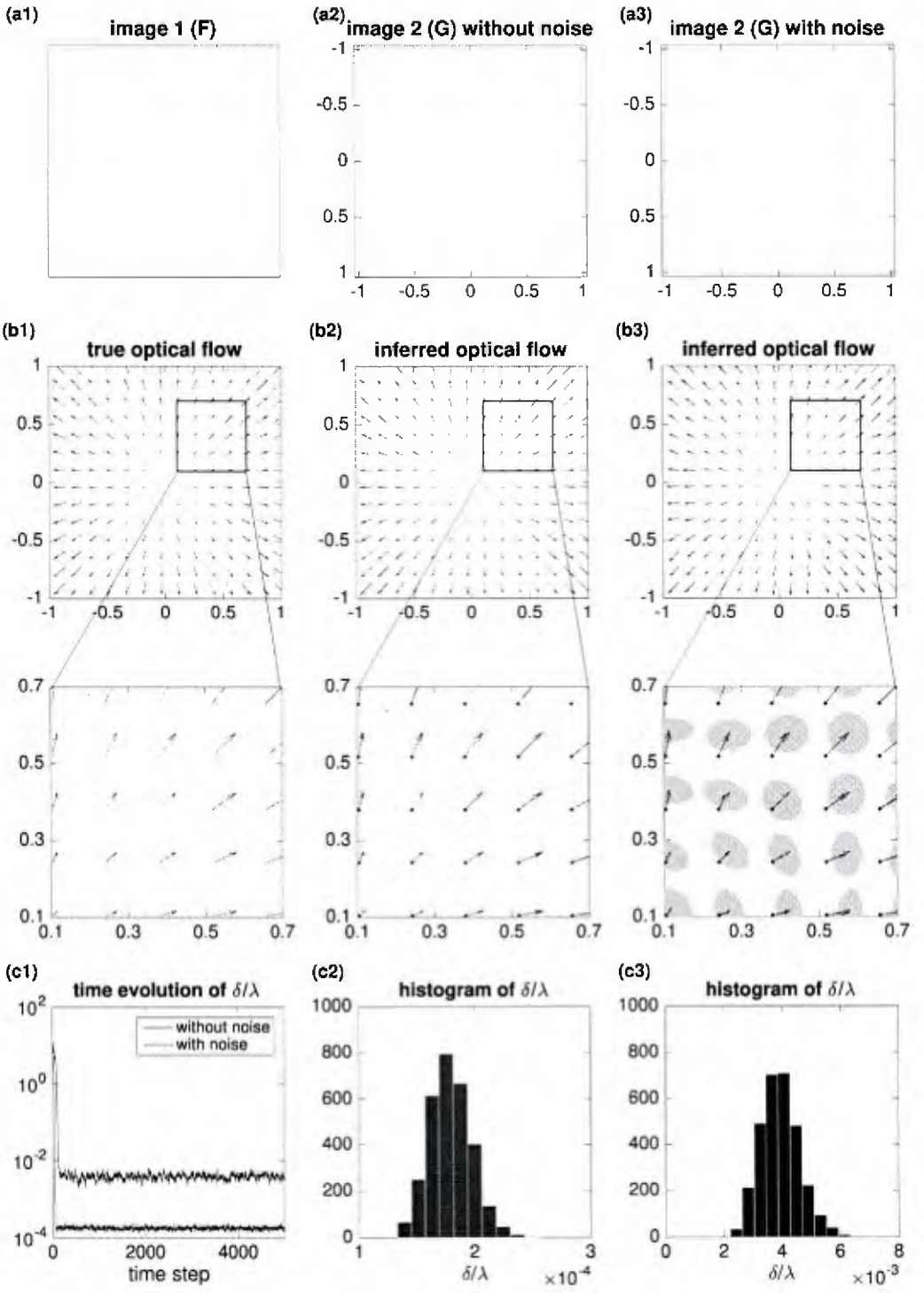


Figure 18: Statistical inversion based optical flow example 1: $\langle U(x,y), V(x,y)\rangle = \langle x, y\rangle$. Top row (a1-a3): image data generated according to Eq. (4.108) for F (a1) and using Eq. (4.109) for G either without noise (a2) or with noise (a3). Middle rows (b1-b3): the true optical flow field (b1) and the inferred optical flow field from the MCMC samples. In all panels corresponding to the inferred optical flow, the mean vector at each point is plotted as an arrow; for the zoomed-in views, we show the computed confidence region defined by Eq. (4.110) as shaded ellipses around the tip of the arrows. Bottom row (c1-c3): time evolution (c1) as well as the distribution (c2-c3) of the effective regularization parameter σ/λ, where the distributions are obtained after discarding the initial transient in the MCMC sampling process.

4.4.2 Statistical Inversion Formulation of Optical Flow

Given a sequence of consecutively captured images, the visual relative motion between them is referred to as optical flow, which can often provide insight about the actual physical motion. The inference of optical flow is an outstanding scientific question, which requires making assumptions about the underlying motion as well as the measurement process.

4.4.3 Problem Setup

Consider two single-channeled (typically grayscale) digital images ("pictures") taken from the same scene at two nearby time instances. The image data are represented by two matrices $F = [F_{ij}]_{n_x \times n_y}$ and $G = [G_{ij}]_{n_x \times n_y}$. Thus each image contains $n_x \times n_y$ pixels, defined on a common two-dimensional subspace Ω. The goal is to find matrices $U = [U_{ij}]_{n_x \times n_y}$ and $V = [V_{ij}]_{n_x \times n_y}$ where $\langle U_{ij}, V_{ij} \rangle$ represents the optical "velocity" occurring at the (i,j)-th pixel inferred from the two images. The image data F and G are often regarded as sampled data from smooth functions $F(x,y)$ and $G(x,y)$, with $F_{ij} = F(x_i, y_j)$ and $G_{ij} = G(x_i, y_j)$ where $\{(x_i, y_j)\}_{(i=1,\ldots,n_x; j=1,\ldots,n_y)}$ are grid points from a spatial domain Ω. Thus U and V can be viewed as discrete spatial samples of a smooth velocity field $\vec{W}(x,y) = \langle U(x,y), V(x,y) \rangle$ defined on Ω that captures the visual optical motion occurring between the two observed images.

4.4.4 Variational Approach of Inferring Optical Flow

The classical variational approach of optical flow starts by defining an "energy" functional whose minimization yields an estimation of the optical flow field [73]. One of the most widely used functional was proposed by Horn and Schunck in 1981 [85], given by

$$E(U,V) = \iint_\Omega (F_x U + F_y V + F_t)^2 dxdy + \alpha \iint_\Omega (\|\nabla U\|^2 + \|\nabla V\|^2) dxdy, \tag{4.72}$$

where $U(x,y)$ and $V(x,y)$ are smooth functions defined over Ω which represent a candidate flow field. In the Horn-Schunck functional, the first term is often referred to as *data fidelity* as it measures the deviation of the total image intensity from being conservative, that is, the condition

$$dF/dt = F_x U + F_y V + F_t = 0. \tag{4.73}$$

The second term measures the *solution regularity* by penalizing solutions that have large spatial gradients and is called the regularization term. The relative emphasis of smoothness as compared to "fitting" the total image intensity conservation equation (4.73) is controlled by the positive scalar α which is called a *regularization parameter*. The main role of the regularization term is to ensure that the minimization of the functional is a well posed problem. Without the regularization term there can typically exists an infinite number of solutions all of which trivially satisfies the conservation equation (4.73).

Given α, the functions U and V that minimize the Horn-Schunck functional (4.72) satisfy the Euler-Lagrange equations

$$\begin{cases} F_x(F_x U + F_y V + F_t) = \alpha(U_{xx} + U_{yy}), \\ F_y(F_x U + F_y V + F_t) = \alpha(V_{xx} + V_{yy}), \end{cases} \tag{4.74}$$

which are typically solved by some iterative scheme over a finite set of spatial grid points [85, 95] to produce an estimation of the optical flow. Alternatively, one could also discretize the functional (4.72) itself to yield a finite-dimensional inverse problem as discussed in Section 2.3 below with solution strategy reviewed in Section 4.5.

4.4.5 Discrete Approximation of the Variational Optical Flow Functional

As discussed in the previous section, the classical variational approach of optical flow works by first formulating and minimizing a functional over smooth vector fields, and then evaluating the obtained vector field at the grid points on which the original image data are given. Here we take a different route, by first discretizing the functional (4.72) to convert the functional minimization (an infinite-dimensional problem) into a finite-dimensional linear inverse problem, and then solving the inverse problem to yield solutions which give the values of a vector field defined over the same grid points as the image data.

The remainder of this section will be focused on the conversion of the functional (4.72) into a finite-dimensional function defined on a set of uniformly distributed grid points

$$\{(x_i, y_j)\}_{i=1,2,\ldots,n_x; j=1,2,\ldots,n_y}, \tag{4.75}$$

where $x_{i+1}-x_i = \Delta x$ and $y_{j+1}-y_j = \Delta y$ are the spacing in the x direction and the y direction, respectively. The conversion will be achieved by approximating the integrals in Eq. (4.72) with appropriately derived summations over the grid points. For notational convenience, we use a bold lowercase variable to denote the vectorization of a matrix. For example, the boldface vector $\boldsymbol{q}$ denotes the column vector obtained by "vertically stacking" the columns of a matrix $Q = [\vec{Q}_1, \vec{Q}_2, \ldots, \vec{Q}_n]$ in order [82], where $\vec{Q}_i$ denotes the i-th column of Q. That is,

$$\boldsymbol{q} = \mathrm{vec}(Q) = \begin{pmatrix} \vec{Q}_1 \\ \vec{Q}_2 \\ \vdots \\ \vec{Q}_n \end{pmatrix}. \tag{4.76}$$

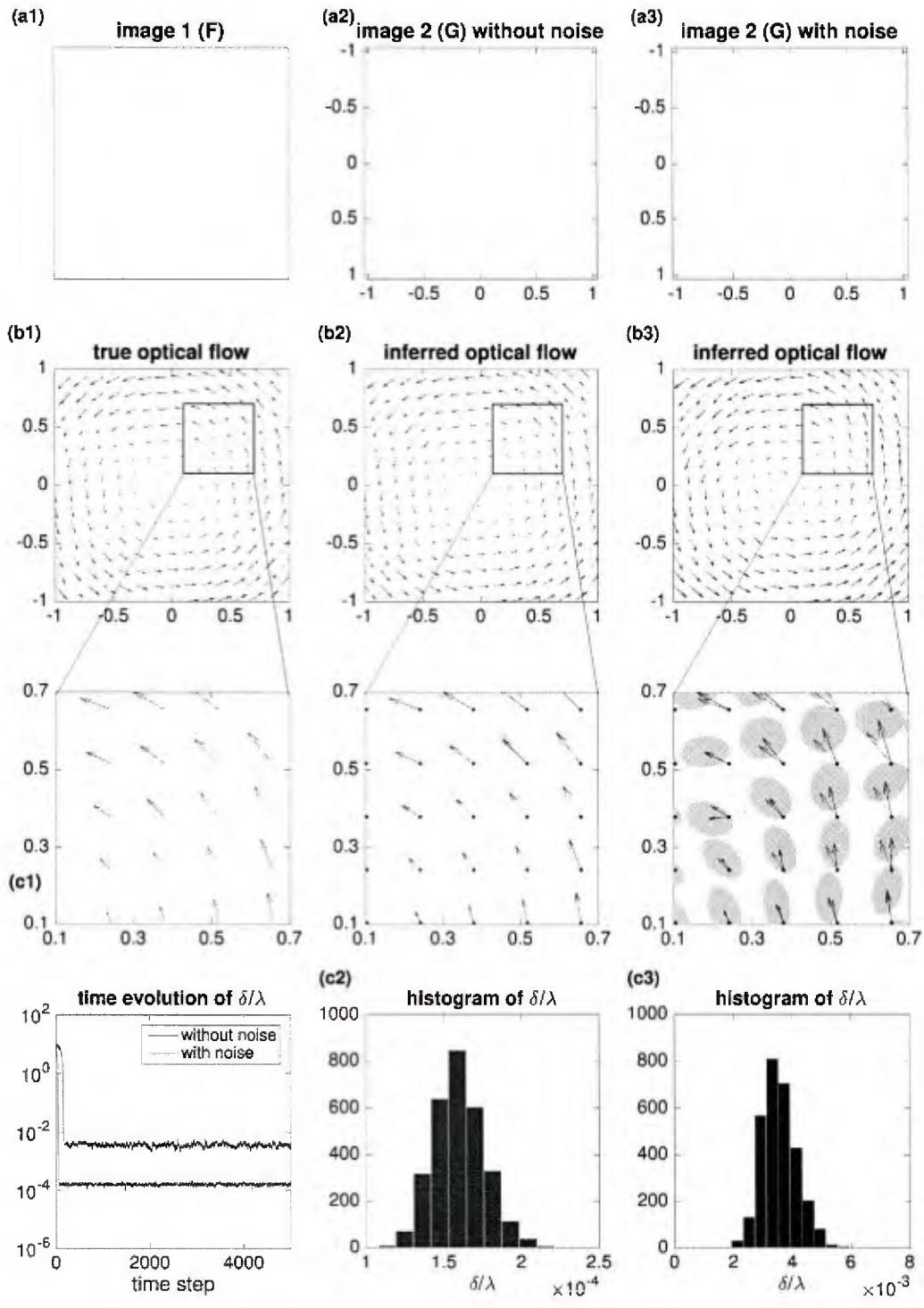


Figure 19: Statistical inversion based optical flow example 1: $\langle U(x,y), V(x,y)\rangle = \langle x, y\rangle$. Top row (a1-a3): image data generated according to Eq. (4.108) for F (a1) and using Eq. (4.109) for G either without noise (a2) or with noise (a3). Middle rows (b1-b3): the true optical flow field (b1) and the inferred optical flow field from the MCMC samples. In all panels corresponding to the inferred optical flow, the mean vector at each point is plotted as an arrow; for the zoomed-in views, we show the computed confidence region defined by Eq. (4.110) as shaded ellipses around the tip of the arrows. Bottom row (c1-c3): time evolution (c1) as well as the distribution (c2-c3) of the effective regularization parameter σ/λ, where the distributions are obtained after discarding the initial transient in the MCMC sampling process.

First let us consider the data fidelity term: $\iint_\Omega (F_x U + F_y V + F_t)^2 dxdy$. The spatial derivatives $F_x(x,y)$ and $F_y(x,y)$ can be approximated by a finite difference scheme. For example, the simple *forward difference* yields the approximations:

$$\begin{cases} F_x(x,y) \approx \frac{1}{\Delta x}\left[F(x+\Delta x, y) - F(x,y)\right], \\ F_y(x,y) \approx \frac{1}{\Delta y}\left[F(x, y+\Delta y) - F(x,y)\right]. \end{cases} \tag{4.77}$$

We next express these derivatives as operations on the column vector $\boldsymbol{f}$. To do this, we define matrix $S_k = [S_k^{(ij)}]_{k\times k}$ as

$$S_k^{(ij)} = \begin{cases} -\delta_{ij} + \delta_{i+1,j}, & \text{if } i < k; \\ -\delta_{i,j-1} + \delta_{i,j}, & \text{if } i = k. \end{cases} \tag{4.78}$$

The forward difference applied to $\boldsymbol{f}$ can be represented as

$$\begin{cases} \boldsymbol{f}_x \approx Q_x \boldsymbol{f}, \\ \boldsymbol{f}_y \approx Q_y \boldsymbol{f}, \end{cases} \tag{4.79}$$

where

$$\begin{cases} Q_x \equiv \frac{1}{\Delta x}[I_n \otimes S_m], \\ Q_y \equiv \frac{1}{\Delta y}[S_n \otimes I_m]. \end{cases} \tag{4.80}$$

The temporal derivative can be estimated from the data by the difference between the two images, to yield

$$\boldsymbol{f}_t \approx \boldsymbol{g} - \boldsymbol{f}. \tag{4.81}$$

With these definitions, we obtain a discretized version of the conservation equation (4.73) expressed as a linear system:

$$A\boldsymbol{x} = \boldsymbol{b}, \tag{4.82}$$

where

$$\begin{cases} A = [\mathrm{diag}(\boldsymbol{f}_x), \mathrm{diag}(\boldsymbol{f}_y)], \\ \boldsymbol{x} = [\boldsymbol{u}^\top, \boldsymbol{v}^\top]^\top, \\ \boldsymbol{b} = -\boldsymbol{f}_t. \end{cases} \tag{4.83}$$

Here $\mathrm{diag}(\boldsymbol{f})$ represents a diagonal matrix whose diagonal elements are given by the entries of the vector $\boldsymbol{f}$. From this connection we approximate the first integral in the functional (4.72) as $\|A\boldsymbol{x} - \boldsymbol{b}\|^2$ where $\|\cdot\|$ denotes the standard Euclidean norm.

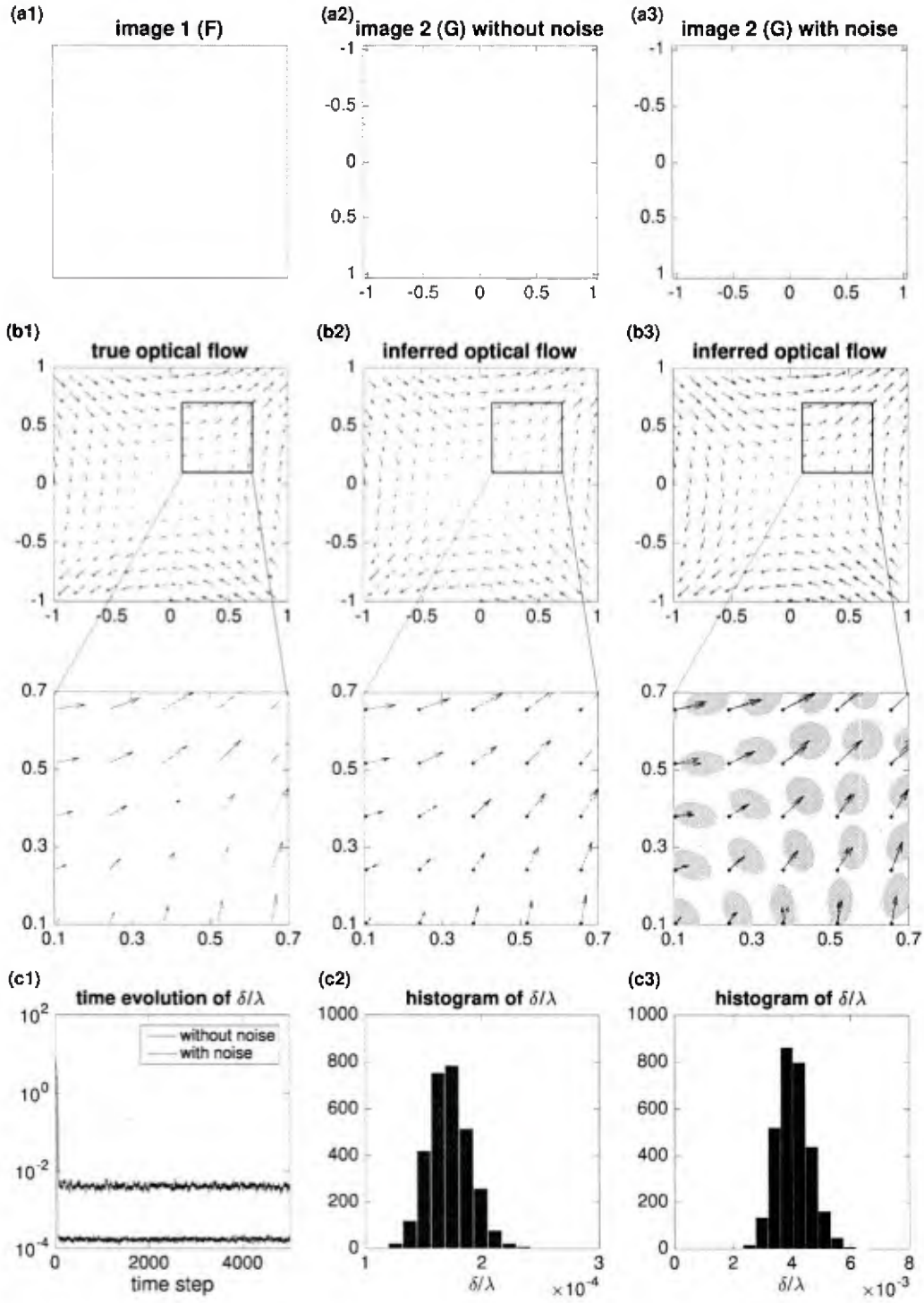


Figure 20: Statistical inversion based optical flow example 1: $\langle U(x,y), V(x,y)\rangle = \langle x, y\rangle$. Top row (a1-a3): image data generated according to Eq. (4.108) for F (a1) and using Eq. (4.109) for G either without noise (a2) or with noise (a3). Middle rows (b1-b3): the true optical flow field (b1) and the inferred optical flow field from the MCMC samples. In all panels corresponding to the inferred optical flow, the mean vector at each point is plotted as an arrow; for the zoomed-in views, we show the computed confidence region defined by Eq. (4.110) as shaded ellipses around the tip of the arrows. Bottom row (c1-c3): time evolution (c1) as well as the distribution (c2-c3) of the effective regularization parameter σ/λ, where the distributions are obtained after discarding the initial transient in the MCMC sampling process.

Next we develop a finite-dimensional approximation of the regularization term in the functional (4.72). This requries discretization of ∇U and ∇V. Using a similar forward difference to approximate the partial derivatives, we obtain

$$\begin{cases} \nabla U \approx \frac{1}{\Delta x}\left[U(x+\Delta x, y) - U(x,y)\right] + \frac{1}{\Delta y}\left[U(x, y+\Delta y) - U(x,y)\right], \\ \nabla V \approx \frac{1}{\Delta x}\left[V(x+\Delta x, y) - V(x,y)\right] + \frac{1}{\Delta y}\left[V(x, y+\Delta y) - V(x,y)\right]. \end{cases} \tag{4.84}$$

For the vectorized variables $\boldsymbol{u}$ and $\boldsymbol{v}$, we have

$$\begin{cases} \nabla \boldsymbol{u} = \boldsymbol{u}_x + \boldsymbol{u}_y \approx (Q_x + Q_y)\boldsymbol{u}, \\ \nabla \boldsymbol{v} = \boldsymbol{v}_x + \boldsymbol{v}_y \approx (Q_x + Q_y)\boldsymbol{v}, \end{cases} \tag{4.85}$$

where Q_x and Q_y are defined in Eq. (4.80). Consequently, we obtain the approximation of the second integral in the Horn-Schunck functional (4.72) as

$$\begin{cases} \iint_\Omega \|\nabla U\|^2 dxdy \approx \boldsymbol{u}_x^\top \boldsymbol{u}_x + \boldsymbol{u}_y^\top \boldsymbol{u}_y \approx \boldsymbol{u}^\top [Q_x^\top Q_x + Q_y^\top Q_y]\boldsymbol{u}, \\ \iint_\Omega \|\nabla V\|^2 dxdy \approx \boldsymbol{v}_x^\top \boldsymbol{u}_y + \boldsymbol{v}_y^\top \boldsymbol{v}_y \approx \boldsymbol{v}^\top [Q_x^\top Q_x + Q_y^\top Q_y]\boldsymbol{v}, \end{cases} \tag{4.86}$$

which then gives

$$\iint_\Omega (\|\nabla u\|^2 + \|\nabla v\|^2) dxdy \approx \boldsymbol{x}^\top Q \boldsymbol{x}, \tag{4.87}$$

where the matrix

$$Q = I_2 \otimes [Q_x^\top Q_x + Q_y^\top Q_y]. \tag{4.88}$$

Therefore, the variational optical flow formulation (4.72) can be reformulated at a finite spatial resolution as the following regularized optimization problem:

$$\min_{\boldsymbol{x}} \left(\|A\boldsymbol{x} - \boldsymbol{b}\|^2 + \alpha \boldsymbol{x}^\top Q \boldsymbol{x} \right), \tag{4.89}$$

which is a standard linear least squares problem with Tikhonov regularization, with more details to be presented in the next section.

4.5 Inverse Problem in Finite Dimensions

Although there has been a great deal of progress on the mathematical characterization of inverse problems in the terms of functional analysis, a practical problem often concerns finding a solution in a finite-dimensional space. At a fundamental level, the most common inverse problem stems from a linear model [84, 86, 100]

$$\boldsymbol{b} = A\boldsymbol{x} + \boldsymbol{\eta}, \tag{4.90}$$

where $\boldsymbol{b} \in \mathbb{R}^m$ is a column vector of *observed data*, $A = [a_{ij}]_{m\times n} \in \mathbb{R}^{m\times n}$ is a (known) matrix representing the underlying model, the column vector $\boldsymbol{\eta} \in \mathbb{R}^n$ denotes (additive) noise, and $\boldsymbol{x} \in \mathbb{R}^n$ is the vector of *unknowns* to be inferred.

Given A and $\boldsymbol{b}$, the problem of inferring $\boldsymbol{x}$ in Eq. (4.90) is called an *inverse problem* because rather than direct "forward" computation from the model, it requires a set of indirect, "backward", or "inverse" operations to determine the unknowns [100]. Depending on the rank and conditioning of the matrix A, the problem may be ill-posed or ill-conditioned. In classical approaches, these issues are dealt with by adjusting the original problem to a (slightly) modified optimization problem as discussed in Section 4.5.1 whose solution is meant to represent the original, as discussed in Section 4.5.2.

We note that in the classical setting a solution to the inverse problem is a vector $\boldsymbol{x}$ as a result of solving an optimization problem. Such a solution is referred to as an *point estimate* because it gives one solution vector without providing any information about how reliable (or uncertain) the solution is [84, 100]. On the other hand, the statistical inversion approach to inverse problems provides an *ensemble* of solutions together with a distribution from which not only point estimates can be made but also their uncertainty quantification [81, 86].

4.5.1 Least squares solution

The classical least squares solution to the inverse problem is given by [82]

$$\boldsymbol{x}_{\ell_2} = A^{\dagger}\boldsymbol{b}, \tag{4.91}$$

where $A^{\dagger}$ denotes the pseudo-inverse of A which can be obtained from the singular value decomposition of A [82]. Depending on the rank of A, the least squares solution $\boldsymbol{x}_{\ell_2}$ is associated with one of the minimization problems:

$$\begin{cases} \min_{A\boldsymbol{x}=\boldsymbol{b}} \|\boldsymbol{x}\|_2, & \text{if } \operatorname{rank}(A) < n \text{ (which is necessarily the case if } m < n); \\ \min_{\boldsymbol{x}} \|A\boldsymbol{x} - \boldsymbol{b}\|_2, & \text{if } \operatorname{rank}(A) = n. \end{cases} \tag{4.92}$$

Here $\|\cdot\|_2$ denotes the ℓ_2 (Euclidean) norm. Let the true solution to Eq. (4.90) be $\boldsymbol{x}^*$, that is, $\boldsymbol{b} = A\boldsymbol{x}^* + \boldsymbol{\eta}$. It follows that

$$\boldsymbol{x}_{\ell_2} - \boldsymbol{x}^* = A^{\dagger}\boldsymbol{\eta}. \tag{4.93}$$

In practice, even when the matrix A has full column rank ($\operatorname{rank}(A) = n$), the discrepancy between the true and least squares solutions is typically dominated by noise when some singular values of A are close to zero, rendering A an ill-conditioned matrix and the solution $\boldsymbol{x}^*$ unstable [100].

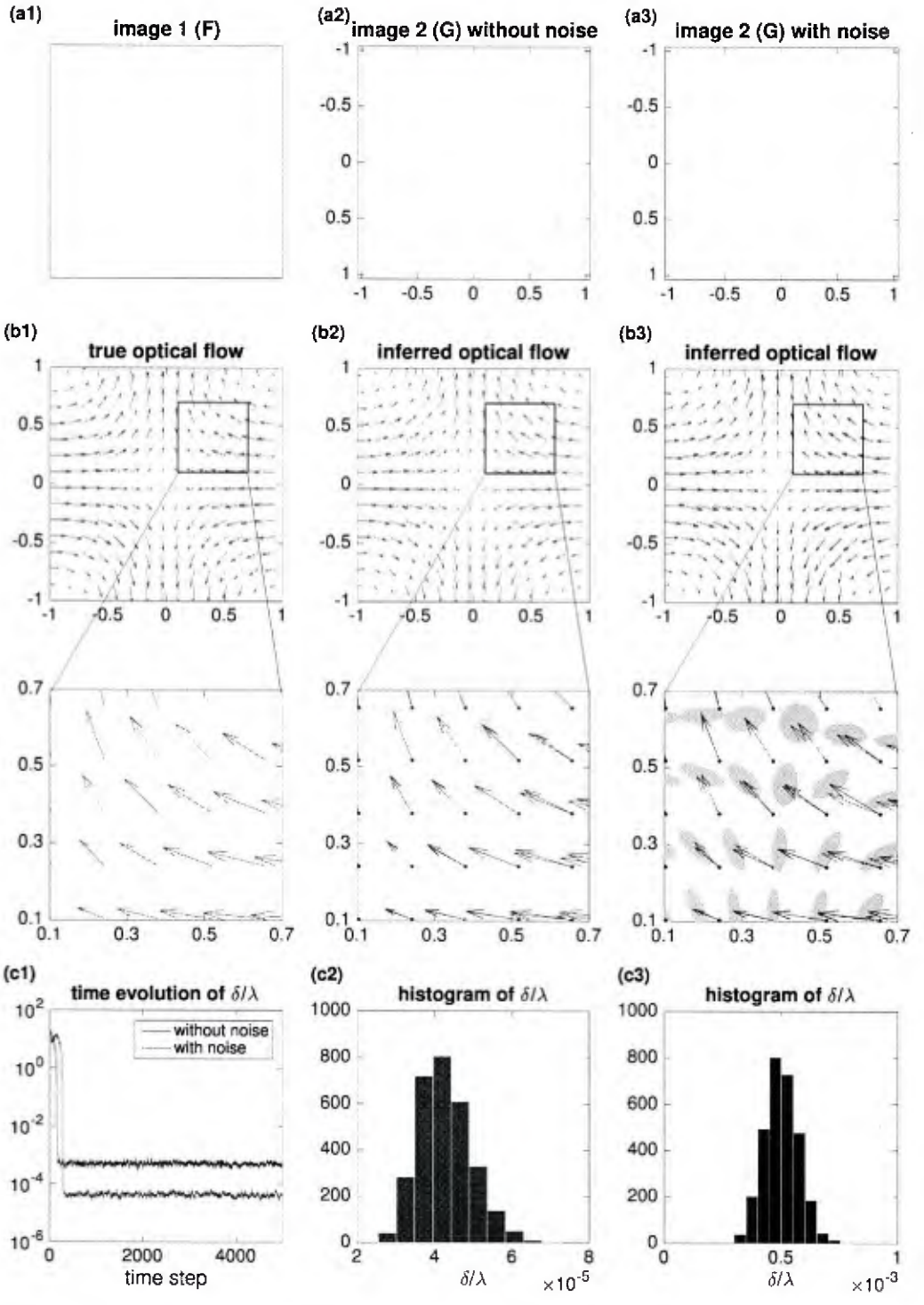


Figure 21: Statistical inversion based optical flow example 1: $\langle U(x,y), V(x,y)\rangle = \langle x, y\rangle$. Top row (a1-a3): image data generated according to Eq. (4.108) for F (a1) and using Eq. (4.109) for G either without noise (a2) or with noise (a3). Middle rows (b1-b3): the true optical flow field (b1) and the inferred optical flow field from the MCMC samples. In all panels corresponding to the inferred optical flow, the mean vector at each point is plotted as an arrow; for the zoomed-in views, we show the computed confidence region defined by Eq. (4.110) as shaded ellipses around the tip of the arrows. Bottom row (c1-c3): time evolution (c1) as well as the distribution (c2-c3) of the effective regularization parameter σ/λ, where the distributions are obtained after discarding the initial transient in the MCMC sampling process.

4.5.2 Tikhonov Regularization

A powerful approach to resolve the instabilities due to noise and the near-singularity of A is to *regularize* the problem. In the classical Tikhonov regularization one adds a quadratic regularization term $\alpha \boldsymbol{x}^T L \boldsymbol{x}$ to penalize non-smoothness, giving rise to a regularized problem [97, 98, 99, 100]:

$$\min_{\boldsymbol{x}} \left(\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \alpha \boldsymbol{x}^T L \boldsymbol{x} \right). \tag{4.94}$$

In the regularized problem the positive scalar parameter α controls the weight of regularization and L is typically a symmetric positive definite matrix, both of which need to be chosen appropriately for the problem to be uniquely defined [97, 98, 99, 100]. We refer to the two terms $\|A\boldsymbol{x} - \boldsymbol{b}\|_2^2$ and $\alpha \boldsymbol{x}^\top L \boldsymbol{x}$ in (4.94) as *data fidelity* and *solution regularity*, respectively. In a simplistic description, they can be described as "selecting" a solution $\boldsymbol{x}_\alpha$ that balances the desire to "solve" $A\boldsymbol{x} = \boldsymbol{b}$ and to be "regular" as measured by $\boldsymbol{x}^\top L \boldsymbol{x}$. The regularization parameter α therefore dictates the extent to which the compromise is made between the two.

For a fixed α, we denote the corresponding regularized solution by

$$\boldsymbol{x}_\alpha = argmin_{\boldsymbol{x}} \{ \|A\boldsymbol{x} - \boldsymbol{b}\|^2 + \alpha \boldsymbol{x}^\top L \boldsymbol{x} \}. \tag{4.95}$$

By standard vector calculus, it can be shown that $\boldsymbol{x}_\alpha$ is in fact a solution to the modified linear system

$$(A^\top A + \alpha L)\boldsymbol{x}_\alpha = A^\top \boldsymbol{b}, \tag{4.96}$$

which is typically well-posed for appropriate choices of L and α. When the matrices are large and sparse, Eq. (4.96) can be solved by iterative methods rather than a direct matrix inversion which tends to be numerically costly and unstable [82].

The key remains how to appropriately choose the regularization parameter α. Despite the existence and ongoing development of many competing methods for selecting α most of which focus on asymptotical optimality as the number of data points approach infinity, none of them stands out as a "natural" choice unless specific priori information about the noise in the data are available (see Chapter 7 of Ref. [100]). In the following section we show that this problem of selecting an appropriate regulation parameter can be effectively bypassed under a statistical inversion framework from a Bayesian perspective.

4.5.3 Statistical Inversion Approach

The statistical inversion approach to an inverse problem starts with using the Bayes rule to express the *posterior* distribution $p(\boldsymbol{x}|\boldsymbol{b})$, which is the conditional distribution of the "solution vector" $\boldsymbol{x}$ given the observed data $\boldsymbol{b}$, as [81, 86]

$$p(\boldsymbol{x}|\boldsymbol{b}) = \frac{1}{p(\boldsymbol{b})} p(\boldsymbol{b}|\boldsymbol{x}) \cdot p(\boldsymbol{x}). \tag{4.97}$$

Here the *likelihood* function $p(\boldsymbol{b}|\boldsymbol{x})$ is the probability density function (pdf) of the random variable $\boldsymbol{b}$ given $\boldsymbol{x}$ which is determined by the underlying model; $p(\boldsymbol{x})$ is the *priori* distribution of $\boldsymbol{x}$; and $p(\boldsymbol{b}) > 0$ acts as a normalization constant which does not affect the solution procedure or the final solution itself.

The key of the statistical inversion formulation is that each candidate solution $\boldsymbol{x}$ is associated with the probability $p(\boldsymbol{x}|\boldsymbol{b})$ that is determined (up to a normalization constant $1/p(\boldsymbol{b})$) once the likelihood function and the prior distribution are given. For a given inverse problem, the likelihood function is determined by the underlying model such as Eq. (4.90) including the noise distribution. On the other hand, the prior distribution $p(\boldsymbol{x})$ is typically constructed according to some prior knowledge of the solution. To obtain desired solutions as well as their statistical properties such as uncertainty quantification, one typically needs to *sample* from the posterior distribution. Efficient sampling methods will be reviewed toward the end of this section.

The unique feature of enabling information fusion and uncertainty quantification has made the statistical inversion approach to inverse problems an attractive venue for the development of new theory and applications. In image processing applications, it has been utilized for many problems such as image denoising and deblurring [74, 87], sparse signal reconstruction [91], and more recently for optical flow computation [79]. In particular, we note that our approach, although different in many of the technical aspects, shares a similar statistical inversion perspective as Ref. [79].

4.5.4 Tikhonov regularization and MAP solution

Under the statistical inversion framework, a particularly popular solution is one that maximizes the posteriori probability given the measured data. Such a solution $\boldsymbol{x}$, called a *maximum a posteriori* (MAP) estimator, is defined as

$$\begin{aligned} \boldsymbol{x} &= argmax_{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{p}) = argmin_{\boldsymbol{x}}\{-\ln p(\boldsymbol{x}|\boldsymbol{b})\} && (4.98) \\ &= argmin_{\boldsymbol{x}}\{-\ln p(\boldsymbol{b}|\boldsymbol{x}) - \ln p(\boldsymbol{x})\}. && (4.99) \end{aligned}$$

As noted in Refs. [74, 86], the MAP estimation given by Eq. (4.99) is in fact equivalent to the Tikhonov regularization specified in Eq. (4.95) upon appropriate choice of the model and prior pdf. In particular, consider the model given by Eq. (4.90) with independent and identically distributed (iid) Gaussian noise of variance λ^{-1}. It follows that

$$p(\boldsymbol{b}|\boldsymbol{x}) = p(\boldsymbol{\eta}) \propto \exp\left(-\frac{\lambda}{2}\|A\boldsymbol{x} - \boldsymbol{b}\|^2\right), \tag{4.100}$$

where the symbol "$\propto$" means "proportional to". If the prior is also taken to be a Gaussian distribution as

$$p(\boldsymbol{x}) \propto \exp\left(-\frac{\delta}{2}\boldsymbol{x}^\top L\boldsymbol{x}\right), \tag{4.101}$$

then the term $-\ln p(\boldsymbol{x}|\boldsymbol{b})$ in the MAP estimator becomes

$$-\ln p(\boldsymbol{x}|\boldsymbol{b}) \propto \|A\boldsymbol{x} - \boldsymbol{b}\|^2 + (\delta/\lambda)\boldsymbol{x}^\top L\boldsymbol{x}. \tag{4.102}$$

The choice of $\delta/\lambda = \alpha$ then yields the same MAP estimator $\boldsymbol{x}$ as the regularization solution $\boldsymbol{x}_\alpha$ given by Eq. (4.95). Thus, with these assumptions of the form of the noise, the distribution of the prior, there is a logical bridge between two different philosophies for inverse problems, since the MAP estimator from the posterior distribution under the statistical inversion framework is equivalent to the result of a specific Tikhonov regularization solution.

Furthermore, and crucially, as shown in the next section, more important information exists in the statistical inversion framework. Specifically, by sampling from the posterior distribution the statistical inversion approach allows for not only a point estimate but also other statistical properties associated with the solution, in particular uncertainty quantification.

4.5.5 Computational Aspects

In the statistical inversion formalism, once the form of the posterior distribution is derived, the remaining part of the work is devoted to efficient sampling from the posterior distribution. Typically a Markov chain Monte Carlo (MCMC) sampling approach is adopted. The main idea is to generate a sequence of samples according to a prescribed Markov chain whose unique stationary distribution is the desired posterior distribution.

In this paper we will follow the work by [74] to consider a specific class of the noise and prior distributions:

$$\begin{cases} \text{likelihood function: } p(\boldsymbol{b}|\boldsymbol{x},\lambda) \propto \lambda^{m/2}\exp\left(-\frac{\lambda}{2}\|A\boldsymbol{x} - \boldsymbol{b}\|^2\right), \\ \text{prior distribution: } p(\boldsymbol{x}|\delta) \propto \delta^{n/2}\exp\left(-\frac{\delta}{2}\boldsymbol{x}^\top L\boldsymbol{x}\right), \end{cases} \tag{4.103}$$

Here the noise is assumed to be additive, Gaussian, and independent of the measured data, with variance λ^{-1}, giving rise to the form of the likelihood function. On the other hand, the prior distribution is considered to be Gaussian with covariance matrix $(\delta L)^{-1}$ (matrix δL is referred to as the *precision* matrix). For optical flow applications, we will choose $L = Q$ where Q is given by Eq. (4.88) which corresponds to a spatial regularization measure. As it turns out, this choice of L is closely related to the selection of prior according to a spatial Gaussian Markov random field which is common in tackling spatial inverse problems [75, 83].

To completely specify the posterior distribution, we also need to choose prior distributions for the parameters λ and δ. These are often called *hyperpriors*. Following Ref. [74], we choose the priors for $p(\lambda)$ and $p(\delta)$ to be Gamma distributions, as

$$\begin{cases} p(\lambda) \propto \lambda^{\alpha_\lambda - 1} \exp(-\beta_\lambda \lambda), \\ p(\delta) \propto \delta^{\alpha_\delta - 1} \exp(-\beta_\delta \delta). \end{cases} \tag{4.104}$$

Such choice ensures that $p(\lambda)$ and $p(\delta)$ are conjugate hyper-priors. Typically, without much prior knowledge of the values of λ and δ one would choose the values of α_λ, β_λ, α_δ, and β_δ to ensure the distributions $p(\lambda)$ and $p(\delta)$ to be "wide", allowing the Markov chain to have the opportunity to explore a large part of the parameter space. Following the remedy suggested in Ref. [74, 75, 83], we set $\alpha_\lambda = \alpha_\delta = 1$ and $\beta_\lambda = \beta_\delta = 10^{-4}$ unless otherwise noted. We tested other choice of parameters as well and they mainly affect the length of the transient in the MCMC sampling process and do not seem to have a strong influence on the asymptotic outcome.

Consequently, the full conditional distributions that relate to the posterior distribution are given by

$$\begin{cases} p(\boldsymbol{x}|\lambda, \delta, \boldsymbol{b}) \propto \exp\left(-\frac{\lambda}{2}\|A\boldsymbol{x} - \boldsymbol{b}\|^2 - \frac{\delta}{2}\boldsymbol{x}^\top L\boldsymbol{x}\right), \\ p(\lambda|\boldsymbol{x}, \delta, \boldsymbol{b}) \propto \lambda^{m/2+\alpha_\lambda - 1} \exp\left(\lambda\left[-\frac{1}{2}\|A\boldsymbol{x} - \boldsymbol{b}\|^2 + \beta_\lambda\right]\right), \\ p(\delta|\boldsymbol{x}, \lambda, \boldsymbol{b}) \propto \delta^{n/2+\alpha_\delta - 1} \exp\left(\delta\left[-\frac{1}{2}\boldsymbol{x}^\top L\boldsymbol{x} - \beta_\delta\right]\right). \end{cases} \tag{4.105}$$

In other words,

$$\begin{cases} \boldsymbol{x}|\lambda, \delta, \boldsymbol{b} \sim \mathcal{N}\left((\lambda A^\top A + \delta L)^{-1}\lambda A^\top \boldsymbol{b}, (\lambda A^\top A + \delta L)^{-1}\right), \\ \lambda|\boldsymbol{x}, \delta, \boldsymbol{b} \sim \Gamma\left(m/2 + \alpha_\lambda, \frac{1}{2}\|A\boldsymbol{x} - \boldsymbol{b}\|^2 + \beta_\lambda\right), \\ \delta|\boldsymbol{x}, \lambda, \boldsymbol{b} \sim \Gamma\left(n/2 + \alpha_\delta, \frac{1}{2}\boldsymbol{x}^\top L\boldsymbol{x} + \beta_\delta\right). \end{cases} \tag{4.106}$$

Here we emphasize that for applications with non-square matrix $A_{m\times n}$ (such as our optical flow application), it is important to have m and n in the right places to ensure an appropriately defined MCMC.

4.5.6 A Gibbs sampler of the posterior distribution

We adopt the block Gibbs sampler developed in Refs. [74, 86] as a specific MCMC procedure to sample the posterior distribution. In theory the sample distribution asymptotically converges to the true posterior distribution. The approach contains the following steps.

0. Initialize δ_0 and λ_0, and set $k = 0$.
1. Sample $\boldsymbol{x}^k \sim \mathcal{N}\left((\lambda A^\top A + \delta L)^{-1}\lambda A^\top \boldsymbol{b}, (\lambda A^\top A + \delta L)^{-1}\right)$.
2. Sample $\lambda_{k+1} \sim \Gamma\left(m/2 + \alpha_\lambda, \frac{1}{2}\|A\boldsymbol{x}^k - \boldsymbol{b}\|^2 + \beta_\lambda\right)$.
3. Sample $\delta_{k+1} \sim \Gamma\left(n/2 + \alpha_\delta, \frac{1}{2}(\boldsymbol{x}^k)^\top L(\boldsymbol{x}^k) + \beta_\delta\right)$.
4. Set $k \leftarrow k + 1$ and return to Step 1.

Here the computational burden is mainly due to Step 1, which requires drawing samples from a multivariate Gaussian variable, which is equivalent to solving the following linear system at each iteration for $\boldsymbol{x}^k$:

$$(\lambda_k A^\top A + \delta_k L)\boldsymbol{x}^k = \lambda_k A^\top \boldsymbol{b} + \boldsymbol{w}, \text{ where } \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \lambda_k A^\top A + \delta_k L). \tag{4.107}$$

For large matrices, instead of a direct solve using Gauss elimination, an iterative method is usually preferred. Among the various notable iterative methods such as Jacobi, Gauss-Seidel (G-S), and conjugate gradient (CG) [90], we adopted the CG for all the numerical experiments as reported in this paper, with a starting vector of all zeros, maximum of 500 iterations, and error tolerance of 10^{-6}

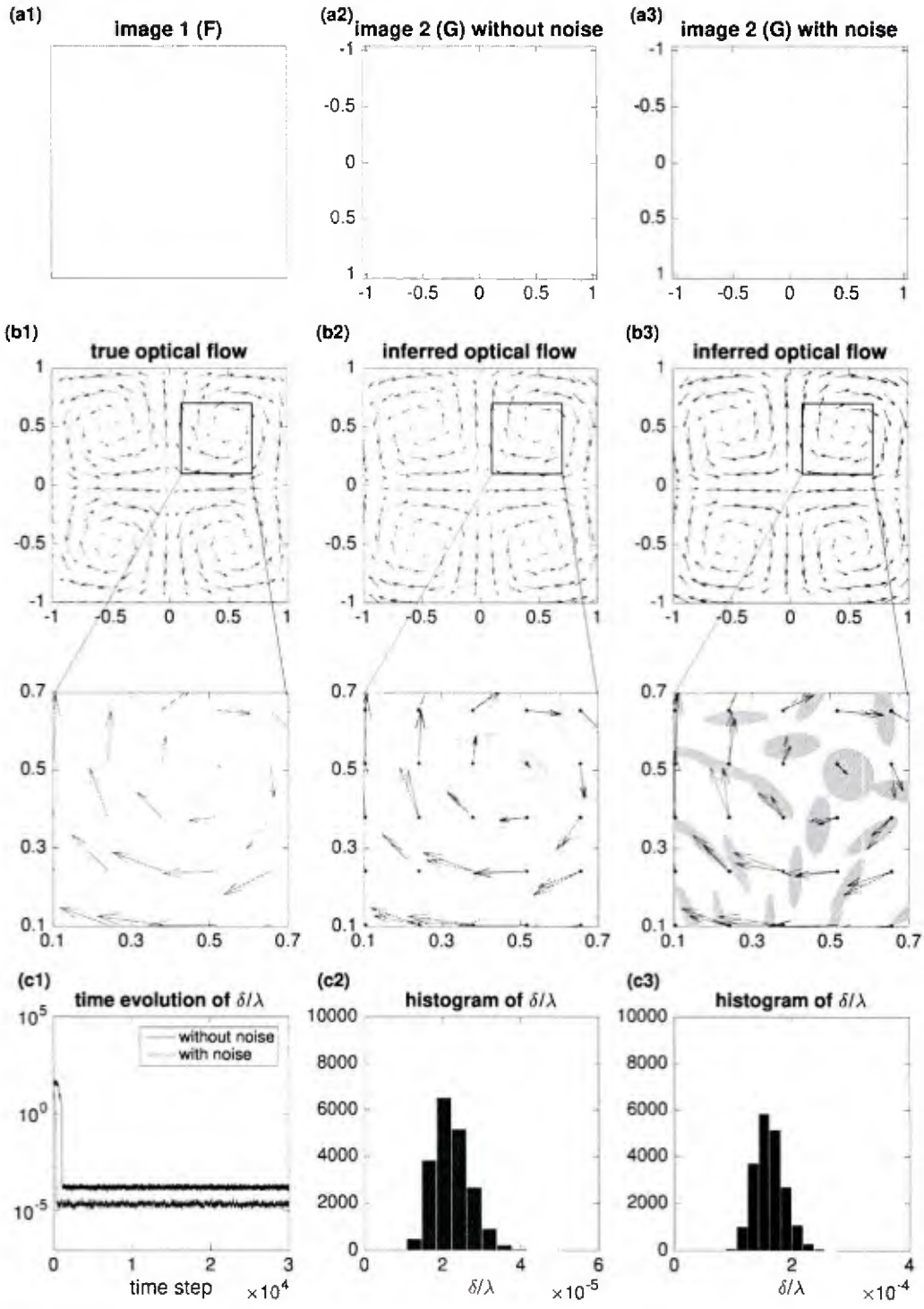


Figure 22: Statistical inversion based optical flow example 1: $\langle U(x,y), V(x,y)\rangle = \langle x, y\rangle$. Top row (a1-a3): image data generated according to Eq. (4.108) for F (a1) and using Eq. (4.109) for G either without noise (a2) or with noise (a3). Middle rows (b1-b3): the true optical flow field (b1) and the inferred optical flow field from the MCMC samples. In all panels corresponding to the inferred optical flow, the mean vector at each point is plotted as an arrow; for the zoomed-in views, we show the computed confidence region defined by Eq. (4.110) as shaded ellipses around the tip of the arrows. Bottom row (c1-c3): time evolution (c1) as well as the distribution (c2-c3) of the effective regularization parameter σ/λ, where the distributions are obtained after discarding the initial transient in the MCMC sampling process.

4.5.7 Examples of Statistical Inversion Based Optical Flow

To test our proposed statistical inversion approach to optical flow, we consider several synthetic image pairs for which the ground truth optical flow (U, V) is known. In each example, we generate two images according to the following equation

$$g = f - f_x u - f_y v + \eta, \tag{4.108}$$

where f is the first image, g is the second image, and η denotes (multivariate) noise whose individual components are independently drawn from a Gaussian distribution with zero mean and fixed standard deviation σ.

We consider 5 qualitatively different optical flow fields, all defined on the same normalized spatial domain $[-1, 1] \times [-1, 1]$, with 30 uniform grid points in each direction, resulting in image matrices of size 30-by-30. The first image in all the examples are identical, generated by the equation

$$F(x, y) = \frac{1}{2}\left[\cos(\pi x)\cos(\pi y) + 1\right], \tag{4.109}$$

as shown in panel (a) of Figures 18 to 22. The second image G is generated according to Eq. (4.108) using various optical flow fields and noise levels as described below, with the spatial derivatives numerically implemented using the forward difference scheme.

Example 1: $\langle U(x, y), V(x, y)\rangle = \langle x, y\rangle$.

Example 2: $\langle U(x, y), V(x, y)\rangle = \langle -y, x\rangle$.

Example 3: $\langle U(x, y), V(x, y)\rangle = \langle y, \sin(x)\rangle$.

Example 4: $\langle U(x, y), V(x, y)\rangle = \langle -\pi\sin(0.5\pi x)\cos(0.5\pi y), \pi\cos(0.5\pi x)\sin(0.5\pi y)\rangle$.

Example 5: $\langle U(x, y), V(x, y)\rangle = \langle -\pi\sin(\pi x)\cos(\pi y), \pi\cos(\pi x)\sin(\pi y)\rangle$.

For each optical flow field, we generate the second image G either without noise or with noise (standard deviation $\sigma = 0.02$). Then, we adopt MCMC-Gibbs sampling procedure and corresponding choice of prior pdf and hyerpriors presented in Section 4 to produce a posterior distribution $p(U, V)$ of the optical flow field. Having a full distribution (the posterior distribution) rather than a single solution is a unique feature of the Bayesian approach to optical flow, which enables one to enumerate and quantify uncertainty as desired.

Figures 18~22 show the results of the statistical inversion based optical flow for the 5 test example, respectively. In each figure, the top row (a1-a3) shows the image data of the first image F (a1), and the second image G generated from Eq. (4.108) with no noise (a2) and with noise under standard deviation $\sigma = 0.02$ (a3), respectively. The middle rows (b1-b3) show the true optical flow field (b1) compared with the inferred "mean" optical flow fields together with uncertainty quantification from the MCMC samples (b2-b3). In particular, at each point $z = (x, y)$, we construct a 2d normal pdf $\mathcal{N}(\mu, \Sigma)$ by using the sample mean μ and sample covariance Σ estimated from the MCMC samples after discarding the initial transients. This allows us to obtain a "mean" optical flow at point z defined as $\langle u(z), v(z)\rangle = \langle \mu_1, \mu_2\rangle$. Uncertainty is quantified by computing a confidence region that contains q probability mass of the fitted multivariate normal distribution given by [92]

$$(z - \mu)^\top \Sigma^{-1}(z - \mu) \le \chi_2^2(q). \tag{4.110}$$

Here $\chi_2^2(q)$ denotes the q-th quantile of the Chi-squared distribution with two degrees of freedom, that is, $\chi_2^2(q) = K^{-1}(q)$ where K is the cdf of χ_2^2. These confidence regions (shaded ellipses) are shown for the zoomed-in plots for the inferred optical flow fields. Finally, the last row (c1-c3) in each figure shows how the MCMC procedure produces a distribution of the effective regularization parameter σ/λ. Panel (c1) shows the change of σ/λ over time in the MCMC sampling procedure, indicating convergence to a stationary distribution typically after a quick initial transient. The remaining panels (c2) and (c3) show the distribution of σ/λ after discarding the initial transient, for both the case of no noise (c2) and the case with noise (c3).

We point out a few observations from the numerical experiments. First, the estimated optical flow compare reasonably well with the true flow in all examples of the qualitatively different optical flow fields, supporting the utility of the proposed statistical inversion approach [see panels (b1-b3) in all figures]. Secondly, we again point out that the MCMC procedure used in our statistical inversion approach to optical flow does not require an active prior choice of the regularization parameter. The MCMC samples seems to quickly converge to a stationary distribution for the effective parameter [panel (c1) in all figures], from which the distributions of parameters and solutions can be determined. Finally, comparing to the noise free images, the estimation of optical flow becomes less accurate when noise is added. It is worth mentioning that the statistical inversion approach in fact allows us to "predict" this difference without knowing the ground-truth optical flow, by quantifying and comparing the uncertainty of solutions [panels (b2) and (b3) in all figures].

4.6 Discussions and Conclusion Remarks for Goal 2

In this paper we take a statistical inversion approach to the optical flow inference problem. The key step is to formulate a linear inverse problem from which the posterior distribution can be expressed by utilizing knowledge about the form of model, noise, and other prior information. From a Bayesian perspective, this information is combined to produce a posterior distribution describing the propagation of prior information in context of the problem. We have shown that traditional variational approaches such as the seminal work developed by Horn and Schunck [85] can in fact be regarded as a special case within the statistical inversion framework by making specific assumptions about the model, noise, and prior distribution. Thus we recap that there are major advantages over the classical variational calculus approach to inverse problems where by necessity the ill-posedness is dealt with by adding an ad hoc regularity term that hopefully agrees with expected physical interpretation. From a Bayesian perspective, the ill-posedness is dealt with naturally under the statistical inversion framework by restating as a well-posed extended problem in a larger space of probability distributions [81, 86]. This therefore naturally removes a key difficulty of choosing an appropriate regularity parameter encountered in classical methods. Instead, in contrast to classical optical flow methods which only yield single solutions as "point estimates", the statistical inversion approach produces a distribution of solutions which can be sampled in terms of most appropriate estimators and also for uncertainty quantification. Specifically in the context of an optic flow problem, we expect a distribution of regularity parameters, and correspondingly a distribution of optical flow vectors at each point.

In this paper we focused on a statistical inversion formulation that is based on the classical Horn-Schunck functional, other functionals can be similarly coped with so long as a linear inverse problem can be formulated. The Horn-Schunck framework assumed rigid body motion and conservation of brightness, giving rise to a divergence-free functional. Other data fidelity terms [77, 88, 78] can be formulated to correspond to the physics of the underlying application, for example for fluid and oceanographic problems where a stream function, or even a quasi-static approximation to assume such physics as corriolis can be used. Likewise, regularity in time and multiple time shots may be appropriate [77], as these correspondingly more complex formulations nonetheless come back to a linear inverse problem tenable in the framework of this paper. Specifically within the statistical inversion framework developed in the current paper, all of these could be recast so that the data fidelity term may be written into the basic form here to allow the statistical inverse problems framework, which we plan in future work. Likewise, other numerical differentiation and integration schemes can be used as well in place of the simple forward difference used here.

References

[1] C. Antoni, B. Vasconcelos, Modeling, clustering, and segmenting video with mixtures of dynamic textures, IEEE Transactions on Pattern Analysis and Machine Intelligence, (2008), 30(5).

[2] S. Arya and D. M. Mount, Approximate Range Searching, Computational Geometry: Theory and Applications, vol. 17, 2000, pp. 135-163.

[3] S. Bagheri, "Koopman-mode decomposition of the cylinder wake". J. Fluid Mech. (2013), vol. 726, pp. 596-623.

[4] R. Banisch, P. Koltai, "Understanding the geometry of transport: diffusion maps for Lagrangian trajectory data unravel coherent sets," preprint, arxiv, 1603.04709, 2016.

[5] JRM, D, Bargteil, M. Kingsbury, K.A. Mitchell, T. Solomon, "Invariant barriers to reactive front propagation in fluid flows". EPL 98 ,2012, 44005.

[6] R. Basnayake and E.M. Bollt, A Multi-Time Step Method to Compute Optical Flow with Scientific Priors for Observations of a Fluidic System, BIRS Book Chapter, Springer Proceedings in Mathematics and Statistics, Vol 70, April, (2014) pp 59-79.

[7] M. Belkin and P. Niyogi. "Laplacian eigenmaps and spectral techniques for embedding and clustering", NIPS 2002, Vol. 14.

[8] E.M. Bollt and N. Santitissadeekorn, "Applied and Computational Measurable Dynamics", Book Publisher: Society for Industrial and Applied Mathematics, (2013).

[9] O. Boiman, E. Shechtman and M. Irani, "In Defense of Nearest-Neighbor Based Image Classification. CVPR, pp.1-8, 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008.

[10] L. Bovens, and S. Hartmann (2003), "Solving the Riddle of Coherence", Mind 112(448): 601-633.

[11] J. Li, C. Boyer, T. Solomon, K.A. Mitchell, "Frozen reaction fronts in steady ows: a burning-invariant-manifold perspective", PRE 92 063005 (2015)

[12] M. Budisic, R. M. Mohr, I. Mezic, Applied Koopmanism. Chaos: An Interdisciplinary Journal of Nonlinear Science. 22, 047510133 (2012).

[13] A. Bugeau, P. Prez. Detection and segmentation of moving objects in complex scenes. ELSEVIER, Computer Vision and Image Understanding, Vol. 113, Issue 4, 459-476, (2009).

[14] F. Chung, "Laplacians and the Cheeger inequality for directed graphs," Annals of Combinatorics, 9 (2005), 1-19.

[15] R.R. Coifman, S. Lafon. "Diffusion maps". Appl. Comput. Harmon. Anal. 21 (2006) 5 30.

[16] Cooperative Institute for Meteorological Satellite Studies, http://cimss.ssec.wisc.edu/.

[17] A. Denner, O. Junge, D. Matthes. Computing Coherent Sets using the Fokker-Planck Equation. Preprint.

[18] I.S. Dhillon, Y. Guan, and B. Kulis, (2004). "Kernel k-means: spectral clustering and normalized cuts". Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 551-556.

[19] A. Fabregat, A.C. Poje, I. Mezic, "Finite-time Partitions for Lagrangian Structure Identification in Gulf Stream Eddy Transport," arxiv, 1606.07382, (2016).

[20] M. Falorni, "The discovery of the Great Red SPOT of Jupiter," Journal of the British Astronomical Association, vol.97, no.4, p.215-219 (1987).

[21] M. Fernandez, D. Stabafeb, "8 Days, 10 Feet and the Snow Isnt Done Yet," The New York Times, Feb 12, (2007).

[22] B. Fitelson, (2003), A Probabilistic Theory of Coherence?, Analysis 63: 194-199.

[23] J.J. Fortney, M.S. Marley, D. Saumon, K. Lodders. Synthetic Spectra and Colors of Young Giant Planet Atmospheres: Effects of Initial Conditions and Atmospheric Metallicity. arXiv:0805.1066v1 [astro-ph].

[24] G. Froyland, S. Lloyd, and N. Santitissadeekorn. Coherent sets for nonautonomous dynamical systems. Physica D, 239:1527?1541, 2010.

[25] G. Froyland and K. Padberg. Almost-invariant sets and invariant manifolds ? connecting probabilistic and geometric descriptions of coherent structures in flows. Physica D, 238:1507?1523, 2009.

[26] G. Froyland, "An analytic framework for identifying finite-time coherent sets in time- dependent dynamical systems," Physica D: Nonlinear Phenomena, 250:11719, 2013.

[27] G. Froyland, "Dynamic isoperimetry and the geometry of Lagrangian coherent structures," Nonlinearity, 28:3587-3622, 2015.

[28] G. Froyland, N. Santitissadeekorn, and A. Monahan. "Transport in time-dependent dynamical systems: Finite-time coherent sets," Chaos, 20:043116, 2010.

[29] G. Froyland and K. Padberg-Gehle. "Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion," In Wael Bahsoun, Chris Bose, Gary Froyland, editors, Ergodic Theory, Open Dynamics, and Coherent Structures. Proceedings in Mathematics and Statistics, volume 70, pages 171-216, Springer, 2014

[30] G. Froyland and K. Padberg-Gehle, "A rough-and-ready cluster-based approach for extracting finite-time coherent sets from sparse and incomplete trajectory data," Chaos, 25:087406, 2015

[31] D. Giannakis, J. Slawinska and Z. Zhao. Spatiotemporal feature extraction with data-driven Koopman operators. (2015), J. Mach. Learn. Res. Proceedings, 44, 103-115

[32] D. Giannakis. Data-driven spectral decomposition and forecasting of ergodic dynamical systems (2015), Appl. Comput. Harmon. Anal., in review.

[33] A. Hadjighasem, D, Karrasch, H. Teramoto, and G. Haller. A Spectral Clustering Approach to Lagrangian Vortex Detection. Preprint. arXiv:1506.02258v3.

[34] A. Hadjighasem, D. Karrasch, H. Teramoto, G. Haller, Spectral clustering approach to Lagrangian vortex detection. Phys. Rev. E 93, 063107.

[35] A. Hadjighasem and G. Haller, Geodesic transport barriers in Jupiter's atmosphere: a video-based analysis, SIAM Review 58 (2016) 69-89

[36] G. Haller. Finding finite-time invariant manifolds in two-dimensional velocity fields. Chaos, Vol. 10, No. 1, (2000).

[37] G. Haller, F.J. Beron-Vera, Geodesic theory of transport barriers in two-dimensional flows Physica D, 241 (2012) 1680-1702

[38] G. Haller, Lagrangian Coherent Structures. Annual Rev. Fluid. Mech, 47 (2015) 137-162.

[39] R. Kannan, S. Vempala, A. Vetta, "On clusterings - good, bad and spectral," Proceedings of the 41st Annual Symposium on Foundations of Computer Science, (2000).

[40] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu. "The efficient K-means clustering algorithm: analysis and implementation". IEEE Trans. Pattern Analysis Mach. Intell. 2002, 24(7), 8817892.

[41] P. Klein, and T. Warfield (1994), "What Price Coherence?", Analysis 54: 129-132.

[42] Y. Lan and I. Mezic. "Linearization in the large of nonlinear systems and Koopman operator spectrum". Physica D, 242:42-53, (2013).

[43] A.Lasota and J.A.Yorke, Exact dynamical systems and the Frobenius-Perron operator, Trans.Amer.Math.Soc. 273, (1982), 375-384.

[44] E. Lawler (2001). "4.5. Combinatorial Implications of Max-Flow Min-Cut Theorem, 4.6. Linear Programming Interpretation of Max-Flow Min-Cut Theorem". Combinatorial Optimization: Networks and Matroids. Dover. pp. 117-120. ISBN 0-486-41453-1.

[45] A. Luttman, E.M. Bollt, R. Basnayake, and S. Kramer, "A Stream Function Approach to Optical Flow with Applications to Fluid Transport Dynamics," Proceedings in Applied Mathematics and Mechanics 11 1 855 - 856 (2012).

[46] A. Luttman, E.M. Bollt, R. Basnayake, S. Kramer, N.B. Tullaro, "A Stream Function Framework for Estimating Fluid Flow from Digital Imagery," CHAOS 23, 033134 (2013).

[47] T. Ma, E.M. Bollt, "Relatively Coherent Sets as a Hierarchical Partition Method , " International Journal of Bifurcations and Chaos, 23 7 1330026 (2013)

[48] T. Ma, E.M. Bollt, "Shape Coherence and Finite-Time Curvature Evolution," International Journal of Bifurcation and Chaos, Vol. 25, No. 5 (2015) 1550076.

[49] T. Ma, E.M. Bollt, "Differential Geometry Perspective of Shape Coherence and Curvature Evolution by Finite-Time Nonhyperbolic Splitting," SIAM Journal on Applied Dynamical Systems (SIADS), Vol. 13 No. 3 Pg. 1106-1136 2014

[50] T. Ma, N. Ouellette, E.M. Bollt, "Stretching and Folding in Finite Time , " Chaos 26, 023112 (2016)

[51] M. Meila, J. Shi, "A random walks view of spectral segmentation," AISTATS 2001.

[52] M. Meila, J. Shi, "Learning segmentation by random walks," Neural Information Processing Systems, 13 (2001).

[53] J.D. Meiss, Symplectic Maps, Variational Principles, and Transport, Reviews of Modern Physics 64 795-848 (1992)

[54] T. Merricks, (1995), On Behalf of the Coherentist?, Analysis 55: 306-309.

[55] K.A. Mitchell, "Finite-time barriers to front propagation in two-dimensional uid flows", Chaos 25 : 087404 (2015).

[56] M.Mori, On the convergence of the spectrum of Perron-Frobenius operators, Tokyo J. Math. 17 (1994), 1-19.

[57] N.M. Murty, S. Devi, "Pattern Recognition: An Algorithmic Approach". V. (2011). ISBN 0857294946.

[58] B. Nadler, S. Lafon, R.R. Coifman; I.G. Kevrekidis. "Diffusion Maps, Spectral Clustering and Eigenfunctions of Fokker?Planck Operators". in Advances in Neural Information Processing Systems 18. (2005).

[59] NASA , http://www.nasa.gov/mission_pages/cassini/multimedia/pia04866.html

[60] NASA , http://saturn.jpl.nasa.gov/mission/quickfacts/

[61] A.Y. Ng, M.I. Jordan, Y. Weiss: "On spectral clustering: Analysis and an algorithm". In: Advances in neural information processing systems. 2, (2002), S. 849-856.

[62] K. Onu, F. Huhn and G. Haller. LCS Tool : A Computational Platform for Lagrangian Coherent Structures. arXiv:1406.3527v1.

[63] C.H. Papadimitriou, K. Steiglitz (1998). "6.1 The Max-Flow, Min-Cut Theorem". Combinatorial Optimization: Algorithms and Complexity. Dover. pp. 120?128. ISBN 0-486-40258-4.

[64] P. Perona, W. T. Freeman, "A Factorization Approach to Grouping," Proceedings of the 5th European Conference on Computer Vision Vol I 665-670 (1998).

[65] M. Reed and B. Simon, Methods of Modern Mathematical Physics IV: Analysis of Operators, Academic Press, 1978.

[66] C. W. Rowley, I. Mezic, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," J. Fluid Mechanics 641 : 115-127, December 2009.

[67] S.C. Shadden and F. Lekien and J.E. Marsden, "Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows,"Phys. D, 271-304 212, 2005.

[68] S. J. Shi and J. Malik,"Normalized Cuts and Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, (2000).

[69] T. Shogenji, (2003), A Condition for Transitivity in Probabilistic Support?, British Journal for the Philosophy of Science 54: 613-616.

[70] A. Surana, Koopman operator based nonlinear dynamic textures, 2015 American Control Conference (ACC), 1333-1338

[71] R. Talmon, R.R. Coifman, "Empirical intrinsic geometry for nonlinear modeling and time series filtering," PNAS 2013 110-31125351712540.

[72]] S.M.Ulam and J.von Neumann, On combination of stochastic and deterministic processes, Bull. Am. Math. Soc., 53:1120, (1947).

[73] G. Aubert, R. Deriche, and P. Kornprobst, *Computing optical flow via variational techniques*, SIAM J. Appl. Math. **60**, 156–182 (1999).

[74] J. M. Bardsley, *MCMC-based image reconstruction with uncertainty quantification*, SIAM J. Sci. Comput. **34**, A1316–A1332 (2012).

[75] J. M. Bardsley, *Gaussian Markov random field priors for inverse problems*, Inverse Problems and Imaging **7**, 397–416 (2013).

[76] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, *Performance of optical flow techniques*, International Journal of Computer Vision, **12(1)**, 43–77 (1994).

[77] R. Basnayake and E. M. Bollt, *A Multi-Time Step Method to Compute Optical Flow with Scientific Priors for Observations of a Fluidic System*, BIRS Book Chapter in *Ergodic Theory, Open Dynamics, and Coherent Structures*, Editors: Froyland, Gary; Bahsoun, Wael (Springer 2014).

[78] R. Basnayake, A. Luttman, and E. M. Bollt, *A lagged diffusivity method for computing total variation regularized fluid flow*, Contemp. Math **586**, 57–64 (2013).

[79] G. Chantas, T. Gkamas, and C. Nikou, *Variational-Bayes optical flow*, J. Math. Imaging Vision **50**, 199–213 (2014).

[80] I. Cohen and I. Herlin, *Optical flow and phase portrait methods for environmental satellite image sequences*, in Proc. Europ. Conf. Computer Vision, Cambridge, U.K., Apr. 1996, pp. 141–150.

[81] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis, Third Edition*, Chapman & Hall/CRC, 2014.

[82] G. H. Golub and C. F. Van Loan, *Matrix Computations, 4th ed.*, Johns Hopkins University Press, 2012.

[83] D. Higdon, *A primer on space-time modelling from a Bayesian perspective*, Los Alamos National Laboratory, Statistical Sciences Group, Technical Report, LA-UR-05-3097.

[84] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.

[85] B. K. P. Horn and B. G. Schunck, *Determining optical flow*, Artificial Intelligence **17**, 185–203 (1981).

[86] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer 2005.

[87] M. Lebrun, A. Buades, and J. M. Morel, *A nonlocal Bayesian image denoising algorithm*, SIAM J. Imaging Sci. **6**, 1665–1688 (2013).

[88] A. Luttman, E. Bollt, R. Basnayake, and S. Kramer, *A stream function approach to optical flow with applications to fluid transport dynamics*, Proc. Appl. Math. Mech. **11**, 855–856 (2011).

[89] A. Luttman, E. M. Bollt, R. Basnayake, S. Kramer, and N. B. Tufillaro, *A framework for estimating potential fluid flow from digital imagery*, Chaos **23**, 033134 (2013).

[90] Y. Saad, *Iterative methods for sparse linear systems, 2nd ed.*, SIAM, 2003.

[91] M. W. Seeger and H. Nickisch, *Large scale Bayesian inference and experimental design for sparse linear models*, SIAM J. Imaging Sci. **1**, 166–199 (2011).

[92] M. Siotani, *Tolerance regions for a multivariate normal population*, Annals of the Institute of Statistical Mathematics **16**, 135–153 (1964).

[93] K. Souhila and A. Karim, *Optical flow based robot obstacle avoidance*, International Journal of Advanced Robotic Systems **4(1)**, 13–16 (2007).

[94] R. Szeliski, *Computer Vision*, Springer 2011.

[95] L. Le Tarnec, F. Destrempes, G. Cloutier, and D. Garcia, *A proof of convergence of the Horn-Schunck optical flow algorithm in arbitrary dimension*, SIAM J. Imaging Sci. **7**, 277–293 (2014).

[96] A. M. Thompson, J. C. Brown, J. W. Kay, and D. M. Titterington, *A study of methods of choosing the smoothing parameter in image restoration by regularization*, IEEE Trans. Pattern Anal. Mach. Intell. **13**, 326–339 (1991).

[97] A. N. Tikhonov, *Regularization of incorrectly posed problems*, Soviet Mathematics Doklady **4**, 1624–1627 (1963).

[98] A. N. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*, Wiley, New York, 1977.

[99] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer Academic Publishers, 1990.

[100] C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.

[101] J. Weickert and C. Schnörr, *Variational optic flow computation with a spatio-temporal smoothness constraint*, Journal of Mathematical Imaging and Vision **14**, 245–255 (2001).

[102] M.O. Williams, I.G. Kevrekidis, C.W. Rowley, "A Data-Driven Approximation of the Koopman Operator: Extending Dynamic Mode Decomposition". J Nonlinear Sci (2015) 25:1307-1346.

References

[1] Kathleen T Alligood, Tim D Sauer, and James A Yorke. *Chaos.* Springer, 1996.

[2] M. R. Allshouse and J.-L. Thiffeault. Detecting coherent structures using braids. *Physica D*, 241:95–105, 2012.

[3] H. Aref. Stirring by chaotic advection. *J. Fluid Mech.*, 143:1–21, 1984.

[4] F. J. Beron-Vera, Y. Wang, M. J. Olascoaga, G. J. Goni, and G. Haller. Objective detection of oceanic eddies and the Agulhas leakage. *J. Phys. Oceanogr.*, 43:1426–1438, 2013.

[5] Erik M Bollt. Stability of order: An example of horseshoes" near" a linear map. *International Journal of Bifurcation and Chaos*, 9(10):2081–2090, 1999.

[6] Erik M Bollt, Aaron Luttman, Sean Kramer, and Ranil Basnayake. Measurable dynamics analysis of transport in the Gulf of Mexico during the oil spill. *Intl. J. Bifurcat. Chaos*, 22(03), 2012.

[7] Erik M Bollt and Naratip Santitissadeekorn. *Applied and Computational Measurable Dynamics.* SIAM, 2013.

[8] Erik M Bollt, Theodore Stanford, Ying-Cheng Lai, and Karol Życzkowski. What symbolic dynamics do we get with a misplaced partition?: on the validity of threshold crossings analysis of chaotic time-series. *Physica D: Nonlinear Phenomena*, 154(3):259–286, 2001.

[9] Peter Alan Davidson. *Turbulence: an introduction for scientists and engineers.* Oxford University Press, 2004.

[10] Robert L Devaney, Luke Devaney, and Luke Devaney. *An introduction to chaotic dynamical systems*, volume 6. Addison-Wesley Reading, 1989.

[11] M do Carmo. Differential geometry of curves and surfaces prentice-hall international. *Englewood Cliffs*, 1976.

[12] F. d'Ovidio, V. Fernández, E. Hernández-García, and C. López. Mixing structures in the Mediterranean Sea from finite-size Lyapunov exponents. *Geophys. Res. Lett.*, 31:L17203, 2004.

[13] Luther Pfahler Eisenhart. *A treatise on the differential geometry of curves and surfaces.* Ginn, 1909.

[14] Gary Froyland. Dynamic isoperimetry and the geometry of lagrangian coherent structures. *arxiv.org/pdf/1411.7186*, 2015.

[15] Gary Froyland and Kathrin Padberg. Almost-invariant sets and invariant manifolds–connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D*, 238(16):1507–1523, 2009.

[16] Gary Froyland and Kathrin Padberg-Gehle. Finite-time entropy: A probabilistic approach for measuring nonlinear stretching. *Physica D: Nonlinear Phenomena*, 241(19):1612–1628, 2012.

[17] Gary Froyland, Naratip Santitissadeekorn, and Adam Monahan. Transport in time-dependent dynamical systems: Finite-time coherent sets. *Chaos*, 20(4):043116, 2010.

[18] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.

[19] G. Haller. Lagrangian coherent structures. *Annu. Rev. Fluid Mech.*, 47:137–161, 2015.

[20] G. Haller and A. C. Poje. Finite time transport in aperiodic flows. *Physica D*, 119:352–380, 1998.

[21] George Haller. Lagrangian coherent structures from approximate velocity data. *Physics of Fluids (1994-present)*, 14(6):1851–1861, 2002.

[22] George Haller and FJ Beron-Vera. Coherent lagrangian vortices: The black holes of turbulence. *Journal of Fluid Mechanics*, 731:R4, 2013.

[23] J. A. Jiménez Madrid and A. M. Mancho. Distinguished trajectories in time dependent vector fields. *Chaos*, 19:013111, 2009.

[24] D. H. Kelley, M. R. Allshouse, and N. T. Ouellette. Lagrangian coherent structures separate dynamically distinct regions in fluid flows. *Phys. Rev. E*, 88:013017, 2013.

[25] D. H. Kelley and N. T. Ouellette. Onset of three-dimensionality in electromagnetically forced thin-layer flows. *Phys. Fluids*, 23:045103, 2011.

[26] D. H. Kelley and N. T. Ouellette. Separating stretching from folding in fluid mixing. *Nat. Phys.*, 7:477–480, 2011.

[27] Bruce Kitchens. *Symbolic dynamics: one-sided, two-sided and countable state Markov shifts*. Springer Science & Business Media, 1998.

[28] Susan F. Marseken Lambert M. Surhone, Miriam T. Timpledon. *Menger Curvature*. VDM Publishing, 2010.

[29] J. Leger. Menger curvature and rectifiability. *Annals of Mathematics*, 149(3):831–869, 1999.

[30] Y. Liao, D. H. Kelley, and N. T. Ouellette. Effects of forcing geometry on two-dimensional weak turbulence. *Phys. Rev. E*, 86:036306, 2012.

[31] Y. Liao and N. T. Ouellette. Spatial structure of spectral transport in two-dimensional flow. *J. Fluid Mech.*, 725:281–298, 2013.

[32] Tian Ma and Erik Bollt. Differential geometry perspective of shape coherence and curvature evolution by finite-time nonhyperbolic splitting. *SIAM Journal on Applied Dynamical Systems (SIADS)*, 13(3):1106–1136, 2014.

[33] Tian Ma and Erik Bollt. Differential geometry perspective of shape coherence and curvature evolution by finite-time nonhyperbolic splitting. *to appear International Journal of Bifurcations and Chaos*, 2015.

[34] Tian Ma and Erik M Bollt. Relatively coherent sets as a hierarchical partition method. *Intl. J. Bifurcat. Chaos*, 23(07), 2013.

[35] A. J. Mariano, A. Griffa, T. M. Özgökmen, and E. Zambianchi. Lagrangian analysis and predictability of coastal and ocean dynamics 2000. *J. Atmos. Oceanic Technol.*, 19:1114–1126, 2002.

[36] K. Menger. Untersuchungen iber eine allgemeine metrik. vierte untersuchung. zur metrik der liurven. *Math. Ann.*, 103:467–501, 1932.

[37] I. Mezić, S. Loire, V. A. Fonoberov, and P. Hogan. A new mixing diagnostic and gulf oil spill movement. *Science*, 330:486–489, 2010.

[38] N. Mordant, A. M. Crawford, and E. Bodenschatz. Experimental Lagrangian probability density function measurement. *Physica D*, 193:245–251, 2004.

[39] N. T. Ouellette. On the dynamical role of coherent structures in turbulence. *C. R. Physique*, 13:866–877, 2012.

[40] N. T. Ouellette, H. Xu, and E. Bodenschatz. A quantitative study of three-dimensional Lagrangian particle tracking algorithms. *Exp. Fluids*, 40:301–313, 2006.

[41] Thomas Peacock and John Dabiri. Introduction to focus issue: Lagrangian coherent structures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(1):017501, 2010.

[42] R. Peikert, A. Pobitzer, F. Sadlo, and B. Schindler. A comparison of finite-time and finite-size Lyapunov exponents. In P.-T. Bremer, I. Hotz, V. Pascucci, and R. Peikert, editors, *Topological Methods in Data Analysis and Visualization III*, pages 187–200. Springer International Publishing, 2014.

[43] A. C. Poje, G. Haller, and I. Mezić. The geometry and statistics of mixing in aperiodic flows. *Phys. Fluids*, 11:2963–2968, 1999.

[44] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed.* Cambridge University Press, 1992.

[45] O. Reynolds. Study of fluid motion by means of coloured bands. *Nature*, 50:161–164, 1894.

[46] Clark Robinson. *Dynamical systems: stability, symbolic dynamics, and chaos.* CRC press, 1995.

[47] Rex Clark Robinson. *An introduction to dynamical systems: continuous and discrete*, volume 19. American Mathematical Soc., 2012.

[48] B. Sawford. Turbulent relative dispersion. *Annu. Rev. Fluid Mech.*, 33:289–317, 2001.

[49] RS Scorer. *Natural Aerodynamics: International Series of Monographs on Aeronautical Sciences and Space Flight: Aerodynamics*, volume 1. Elsevier, 2014.

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

[51] B. I. Shraiman and E. D. Siggia. Scalar turbulence. *Nature*, 405:639–646, 2000.

[52] E. Shuckburgh and P. Haynes. Diagnosing transport and mixing using a tracer-based coordinate system. *Phys. Fluids*, 15:3342–3357, 2003.

[53] L Sirovich. Chaotic dynamics of coherent structures. *Physica D: Nonlinear Phenomena*, 37(1):126–145, 1989.

[54] Stephen Smale. Differentiable dynamical systems. *Bulletin of the American mathematical Society*, 73(6):747–817, 1967.

[55] P. G. Tait. Note on the circles of curvature of a plane curve. *Proc. Edinburgh Math. Soc.*, page 403, 1896.

[56] Phanindra Tallapragada and Shane D Ross. A set oriented definition of finite-time lyapunov exponents and coherent sets. *Communications in Nonlinear Science and Numerical Simulation*, 18(5):1106–1126, 2013.

[57] T. Theodorsen. Mechanism of turbulence. *Proceedings of the Midwestern Conference on Fluid Mechanics (Ohio State University, Columbus, OH*, 1952.

[58] T. Theodorsen. The structure of turbulence. *50 Jahre Grenzschictforschung, edited by H. Gortler and W. Tollmein*, 1955.