International Journal of Bifurcation and Chaos, Vol. 23, No. 7 (2013) 1330026 (17 pages)

DOI: 10.1142/S0218127413300267

RELATIVELY COHERENT SETS AS A HIERARCHICAL PARTITION METHOD

TIAN MA* and ERIK M. BOLLT†
Department of Mathematics and Computer Science,
Clarkson University, P. O. Box 5815,
Potsdam, NY 13699, USA
**mat@clarkson.edu*
†*bolltem@clarkson.edu*

Received February 21, 2013

Finite time coherent sets [Froyland *et al.*, 2010] have recently been defined by a measure-based objective function describing the degree that sets hold together, along with a Frobenius–Perron transfer operator method to produce optimally coherent sets. Here, we present an extension to generalize the concept to hierarchically define relatively coherent sets based on adjusting the finite time coherent sets to use relative measures restricted to sets which are developed iteratively and hierarchically in a tree of partitions. Several examples help clarify the meaning and expectation of the techniques, as they are the nonautonomous double gyre, the standard map, an idealized stratospheric flow, and empirical data from the Mexico Gulf during the 2010 oil spill. Also for the sake of analysis of computational complexity, we include an Appendix concerning the computational complexity of developing the Ulam–Galerkin matrix estimates of the Frobenius–Perron operator centrally used here.

Keywords: Coherent pairs; relative measures; relatively coherent pairs; Frobenius–Perron operator; subdivision flow chart.

1. Introduction

Central to understanding mixing and transport mechanisms is the related question of defining partitions relative to which the transport can be discussed. To this end, the concept of almost invariant sets defined for autonomous systems [Dellnitz & Junge, 2000; Bollt & Santitissadeekorn, 2012; Froyland & Padberg, 2009] and coherent sets for nonautonomous systems [Froyland *et al.*, 2010] are central since transport may be defined as the measure of the set that leaves the partition element corresponding to the almost invariant set (or finite time coherent set) in a given time epoch. See also [Froyland & Padberg, 2009]. Transfer operator methods are proven to be computationally effective for use in identifying almost invariant sets and finite time coherent sets. See [Bollt *et al.*, 2010; Froyland & Padberg, 2009; Shadden *et al.*, 2005]. The method to identify coherent pairs used here is based on the Frobenius–Perron operator, the Ulam–Galerkin method and the thresholding method. See [Froyland & Padberg, 2009].

In this work, we extend the concept of finite time coherent pairs to incorporate relative measure, and we call this relatively coherent pairs. This extension provides the theoretical framework to simply apply the definition of finite time coherence iteratively and at each stage we hierarchically

†Author for correspondence

define the relative measure on each element of the subpartitions which are developed. The results can be collected in a tree structure to emphasize the hierarchical nested nature of such partitions.

Also we present an adapted thresholding method that can identify the finite-time relatively coherent structures in successive scales of time-dependent systems. We study four examples in this paper, for both closed systems and relatively open systems. In the first and second examples, we show our method identifies the fine scaled relatively coherent structures in a nonautonomous double gyre and a standard map. In the third example, an idealized zonal stratospheric flow, our method gives fine scaled details with respect to relative coherence. In our fourth example, we show the method is also efficient with an open system, corresponding to oceanographic flows in the Mexico Gulf.

This paper is organized as follows. In Sec. 2, we define relative coherent structures. Then we briefly describe the Frobenius–Perron operator, and the Ulam–Galerkin matrix estimate and the thresholding method. In Sec. 3, we give the details of the algorithm and a successive flow chart for explanation and we define the theoretical term restricted Frobenius–Perron operator. In Sec. 4, we apply the method to four examples. Conclusions are given in Sec. 5. In Appendix A, we analyze the computational complexity.

2. Relatively Coherent Structures

We define relatively coherent structures by specializing the definition of coherent pairs, which we now review [Froyland *et al.*, 2010]. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, where $\mathcal{A}$ is a σ-algebra and μ is a normalized measure that is not necessarily invariant. The key to the specialization to relatively coherent pairs is the use of an iteratively defined relative measure, refined from the initial global measure μ. From these, Frobenius–Perron operators follow as does the computation.

Generally, we assume that $\Omega \subset \mathbb{R}^d$. Given a time-dependent flow $\Phi(z, t; \tau) : \Omega \times \mathbb{R} \times \mathbb{R} \to \Omega$, through the time epoch τ of an initial point z at time t, a coherent pair $(A_t, A_{t+\tau})$ can be considered as a pair of subsets of Ω such that,

$$\Phi(A_t, t; \tau) \approx A_{t+\tau}.$$

Definition 2.1 [Froyland *et al.*, 2010]. $(A_t, A_{t+\tau})$ is a (ρ_0, t, τ)-coherent pair if

$$\begin{aligned}\rho_\mu(A_t, A_{t+\tau}) &:= \frac{\mu(A_t \cap \Phi(A_{t+\tau}, t+\tau, -\tau))}{\mu(A_t)} \\ &\geq \rho_0, \end{aligned} \tag{1}$$

where the pair $(A_t, A_{t+\tau})$ are "robust" to small perturbation and $\mu(A_t) = \mu(A_{t+\tau})$.

Note that the definition centrally depends on the full measure μ on Ω, and we will substitute successive relative measures on refinements. Now we consider a relative measure on K induced by μ, where K is a nonempty measurable subset of Ω. In this way enter refinements of the initial partition on successive scales. A **relative measure** of K to Ω is

$$\mu_K(A) := \frac{\mu(A \cap K)}{\mu(K)} \tag{2}$$

for all $A \in \mathcal{A}$.

From the above definition, it follows that the space $(K, \mathcal{A}|_K, \mu_K)$ is also a measure space, where $\mathcal{A}|_K$ is the restriction of M to $\mathcal{A}$ and μ_K is a normalized measure on K. We call the space $(K, \mathcal{A}|_K, \mu_K)$, the *relative measure space*. Now, we define the *relatively coherent pairs*.

Definition 2.2. Relatively coherent structures are those (ρ_0, t, τ)-coherent pairs defined by Definition 2.1, with respect to given relative measures on a subset $K \subset \Omega$, of a given scale.

To find relatively coherent structures in time-dependent dynamical systems, the basic tool is the Frobenius–Perron operator. Let $(\Omega, \mathcal{A}, \mu)$ be a measure space and μ is a normalized Lebesgue measure. If $S : \Omega \to \Omega$ is a nonsingular transformation such that $\mu(S^{-1}(A)) = 0$ for all $A \in \mathcal{A}$ satisfying $\mu(A) = 0$, the unique operator $P : L^1(\Omega) \to L^1(\Omega)$ defined by,

$$\int_A Pf(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx) \tag{3}$$

for all $A \in \mathcal{A}$ is called the Frobenius–Perron operator corresponding to S, where $f(x) \in L^1(\Omega)$. See [Lasota & Mackey, 1994]. In our case, S can be considered as the flow map Φ and the formula above can be written as

$$\begin{aligned} P_{t,\tau} f(z) &:= f(S^{-1}(z)) \cdot |\det D(S^{-1}(z))| \\ &= f(\Phi(z, t+\tau; -\tau)) \\ &\quad \cdot |\det D\Phi(z, t+\tau; -\tau)|. \end{aligned} \tag{4}$$

Suppose X is a subset of M, let Y be a set that includes $S(X)$. We develop partitions for X and Y respectively. In other words, let $\{B_i\}_{i=1}^m$ be a partition for X and $\{C_j\}_{j=1}^n$ be a partition for Y. The Ulam–Galerkin matrix follows a well-known finite-rank approximation of the Frobenius–Perron operator, the entry of which is of the form

$$\hat{P}_{i,j} = \frac{\mu(B_i \cap S^{-1}(C_j))}{\mu(B_i)}, \tag{5}$$

where μ is the normalized Lebesgue measure on Ω. As usual, we numerically approximate $\hat{P}_{i,j}$ by,

$$P_{i,j} = \frac{\#\{x_k : x_k \in B_i \,\&\, S(x_k) \in C_j\}}{\#\{x_k : x_k \in B_i\}}, \tag{6}$$

where the sequence $\{x_k\}$ is a set of test points (passive tracers). See [Ding *et al.*, 2002].

The following thresholding method from [Froyland *et al.*, 2010] finds optimally coherent pairs in a time-dependent dynamical system, with respect to the chosen measure. In the following section, we will iteratively *adapt the method to the relative measure*. This algorithm thresholds to the singular values and singular vectors of the matrix P obtained by the Ulam–Galerkin method:

(1) Calculate the second singular value and corresponding to the left and right singular vectors $\{x_i\}$ and $\{y_j\}$ of the normalized Ulam–Galerkin matrix.
(2) Find values $\{(b_k, c_k)\}$ as pairs such that,

$$\rho(X(b_k), Y(c_k)) = \frac{\displaystyle\sum_{i:x_i > b_k \,\&\, j:y_j > c_k} p_i P_{ij}}{\displaystyle\sum_{i:x_i > b_k} p_i} \tag{7}$$

by thresholding, where

$$\begin{aligned} p_i &= \mu(B_i), \\ X(b_k) &= \bigcup_{i:x_i > b_k} B_i \quad \text{and} \\ Y(c_k) &= \bigcup_{j:y_j > c_k} C_j. \end{aligned} \tag{8}$$

(3) Choose a partition related to a pair (b^*, c^*) of (b_k, c_k) such that

$$\rho^* = \max_k \{\rho(X(b_k), Y(c_k))\}. \tag{9}$$

The partition is maximally coherent with respect to μ on Ω and the test set $\{(b_k, c_k)\}$.

3. Algorithm

We now describe how to find relatively coherent pairs, making use of relative measure and correspondingly restricted Frobenius–Perron operators. By the thresholding Eqs. (7)–(9), we have obtained optimal coherent pairs, which are defined as (X_1, Y_1) and (X_2, Y_2). Y_1 can be considered as the image of X_1 under a flow Φ in time-τ. In order to find relatively coherent structures in X_1, Y_1, X_2 and Y_2, we define relative measures on each of these sets. Define relative measures $\mu_{X_1}(S)$ and $v_{Y_1}(T)$ on the coherent pair (X_1, Y_1), according to Eq. (2). Then we have the measure spaces $(X_1, \mathcal{A}|_{X_1}, \mu_{X_1})$ and $(Y_1, \mathcal{A}|_{Y_1}, v_{Y_1})$ with μ_{X_1} and v_{Y_1} the normalized measures. The probability measure v can be considered as the discretized image of μ. The detail of construction of v can be found in [Froyland *et al.*, 2010].

The relative measures for both X_1 and Y_1 allow the adaptation of the thresholding methods on $(X_1, \mathcal{A}|_{X_1}, \mu_{X_1})$ and $(Y_1, \mathcal{A}|_{Y_1}, v_{Y_1})$ under the relative measures. Then follows two relatively coherent pairs in the previous coherent pair (X_1, Y_1), which are now named as (X_{11}, Y_{11}) and (X_{12}, Y_{12}). Also, the coherent pair (X_2, Y_2) can be divided to two relatively coherent pairs, (X_{21}, Y_{21}) and (X_{22}, Y_{22}).

Next, we repeat the building process above, but we will also require a way to decide when to stop. That is, to define eight normalized measures on each of X_{11}, Y_{11}, X_{12}, Y_{12}, X_{21}, Y_{21}, X_{22} and Y_{22}, respectively, such that they become new spaces with corresponding relative measures, we then apply the adapted thresholding method on these new spaces to get more relatively coherent structures.

Now, we state our hierarchical method as an algorithm with a stopping criterion, and for convenience, we use (X_i, Y_i) and (X_j, Y_j) to denote two coherent pairs, i and j which can be stated to emphasize the hierarchy tree.

Algorithm 1

(1) Define relative measures μ_{X_i} and v_{Y_i} and relative measure spaces $(X_i, \mathcal{A}|_{X_i}, \mu_{X_i})$ and $(Y_i, \mathcal{A}|_{Y_i}, v_{Y_i})$ for (X_i, Y_i), where

$$\mu_{X_i}(S) = \frac{\mu(S)}{\mu(X_i)} \quad \text{and} \quad v_{Y_i}(T) = \frac{v(T)}{v(Y_i)} \tag{10}$$

for all $S \subset X_i$ and for all $T \subset Y_i$.
(2) Apply the adapted thresholding method on $(X_i, \mathcal{A}|_{X_i}, \mu_{X_i})$ and $(Y_i, \mathcal{A}|_{Y_i}, v_{Y_i})$ for (X_i, Y_i)

to develop refined coherent pairs (X_{ii}, Y_{ii}) and (X_{ij}, Y_{ij}).

(3) Repeat the above two steps for the pair (X_j, Y_j) to obtain further refined coherent pairs (X_{ji}, Y_{ji}) and (X_{jj}, Y_{jj}).

According to the first three steps, a relatively coherent structure can be denoted by

$$(X_{k_1k_2\ldots k_q}, Y_{k_1k_2\ldots k_q}) \tag{11}$$

and $k_p \in \{i, j\}, p = 1, 2, \ldots, q$, through q-steps of the algorithm, $q+1$ levels deep into the tree. See Fig. 1. That is, the subscript $k_1k_2\cdots k_q$ can be any possible finite q permutations of i and j. We usually choose $i = 1$ and $j = 2$. Figure 1 is a flow chart depicting four levels, which describes the steps to find relatively coherent pairs in finer "scales". The chart emphasizes a "tree" structure. Besides, we only compute the Ulam matrix once at the beginning, then we mask the matrix. For more detail of masking matrix, see [Bollt & Santitissadeekorn, 2012].

However, we cannot repeat the algorithm forever, so we must decide a stopping criterion. The following step is as a completion for the algorithm.

(4) Stop a given branch if in Eq. (9),

$$\rho^* = \max_k \{\rho(X_{k_1k_2\ldots k_q}(b_k), Y_{k_1k_2\ldots k_q}(c_k))\}, \tag{12}$$

is such that,

$$\rho^* < \rho_0, \tag{13}$$

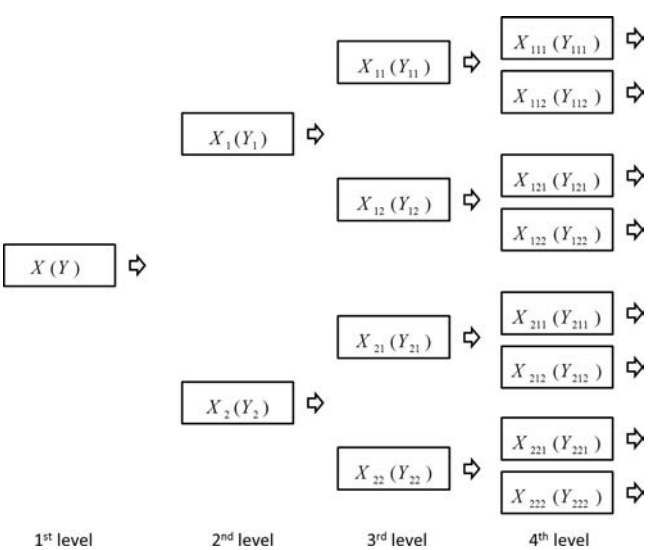


Fig. 1. Successive relatively coherent sets tree as per Algorithm 1.

where $\rho_0 \in (0, 1)$ is a threshold from Eq. (1), descriptive of optimal coherence which is not very coherent.

The Gulf Example in Fig. 7 shows how such stopping criterion leads to the number of coherent pairs $\neq 2^q$.

4. Examples

Before showing our examples, we introduce a theorem and its corollary on computational complexity, the detail is in Appendix A.

The question is how many sample initial points should we use for a given Ulam grid. In other words, for a given grid, how many initial points is enough to well represent the whole domain under a flow so that we can accurately build the Ulam–Galerkin matrix. Intuitively, we may wish to add "as many as possible", however, it will lead to an expensive computation. Generally, finer grids require exponentially more points, depending on the dimensionality, but also on the local stretching of the map. Let $f : X \times [t_0, t] \to X$ be a flow and X be compact. We use $X(t_0)$ and $X(t)$ to denote the status of X at time t_0 and t under the flow f, respectively. For an arbitrary point $x \in X$, we set $x(t_0)$ and $x(t)$ as the positions of x at times t_0 and t. Consider a square box with length q, we use such size q box to make partitions for $X(t_0)$ and $X(t)$, which are $B = \{B_i\}_{i=1}^m$ and $C = \{C_j\}_{j=1}^n$, m and n are positive integers.

Theorem 1. *For a Lipschitz flow f in the plane $X \in \mathbb{R}^2$ and partitions $B = \{B_i\}_{i=1}^m$ and $C = \{C_j\}_{j=1}^n$ consist of identical boxes with length $q > 0$ for $X(t_0)$ and $X(t)$, where $t_0 < t$. Let $P_{i,j}$ be from Eq. (6) and $\hat{P}_{i,j}$ be from Eq. (5). For each $P_{i,j}$ built by N uniformly sampled points from a B_i-q-box, $\exists$ is a confidence coefficient $r \in (0, 1)$ which depends on the Lipschitz constant M of f, and the density δ of sampling points N. According to Eqs. (A.7) and (A.13), it follows that $r \to 0$ as $N \to \infty$.*

Note that the confidence coefficient r may vary for different B_i. If we use local Lipschitz constants for each box B_i and this discussion could lead to adaptive grids but here we simply use a uniform Lipschitz constant from a global grid.

As a direct conclusion, we have the following corollary.

Corollary 4.1. *Alternatively, choosing $r \in (0, 1)$ for a B_i box, the number of uniform sampling*

points $N(r, B_i)$ of B_i box follows from Eqs. (A.7) and (A.13), such that we can control the error $|P_{i,j} - \hat{P}_{i,j}|$ term by term as desired.

4.1. *The nonautonomous double gyre*

Consider the nonautonomous double gyre system,

$$\begin{aligned} \dot{x} &= -\pi A \sin(\pi f(x,t)) \cos(\pi y), \\ \dot{y} &= \pi A \cos(\pi f(x,t)) \sin(\pi y), \end{aligned} \tag{14}$$

where $f(x,t) = \epsilon \sin(\omega t)x^2 + (1 - 2\epsilon \sin(\omega t))x$, $\epsilon = 0.25$, $\omega = 2\pi$ and $A = 0.25$. See [Froyland & Padberg, 2009].

Let the initial time be $t = 0$ and the final time be $t = 10$. We use 24 200 identical triangles $\{B_i\}_{i=1}^{24\,200}$ to cover the region $[0, 2] \times [0, 1]$ by Delaunay triangulation. For good sampling, the relationships between the given grid and the necessary number of points we choose is discussed on computational complexity. The Lipschitz constant M of this model is 1.3999. Therefore, by Theorem 1, for a 24 200 by 24 200 matrix and a confidence threshold $r = 0.5\%$, we have all δs is around $8.9997e{-}04$, so we need about at least 9 000 000 points. Then randomly and uniformly we choose 10 000 000 points in the region as our initial conditions.

We numerically calculate the final status of these points by the Runge–Kutta method to estimate the flow. Because the double gyre model is an area-preserving system, the same triangulation can be used as the image partition $\{C_j\}_{j=1}^{24\,200}$. The Ulam–Galerkin's transition matrix estimates the Frobenius–Perron operator that has the size 24 200 by 24 200.

We apply the thresholding method on the matrix to find the first two coherent pairs in the initial status and final status. See Fig. 2. The first two coherent pairs are colored blue and red in the first level of both the upper and lower charts. We define the left part of the initial status as X_1 and the left part of the final status as Y_1, both of which are filled with blue as halves of the initial status and final status. Thus, two relative measures can be defined on each of these two parts separately,

$$\mu_{X_1}(S) = \frac{\mu(S)}{\mu(X_1)}, \quad v_{Y_1}(T) = \frac{v(T)}{v(Y_1)}. \tag{15}$$

On the other hand, we can do the same in the right red-filled regions which we call the initial status X_2 and the final status Y_2 to develop another two relative measures as follow,

$$\mu_{X_2}(S) = \frac{\mu(S)}{\mu(X_2)}, \quad v_{Y_2}(T) = \frac{v(T)}{v(Y_2)}. \tag{16}$$

Now X_1 and Y_1 can be considered as the initial status and the final status of a refined relative sub-"system". Likewise for X_2 and Y_2. By the same process as with the whole double gyres system, we can get some new coherent structures in the new system consisting of X_1 and Y_1. In Fig. 2, following the first blue arrow between first and second levels of both flow charts, X_1, the blue half on the first level of the upper chart is divided by blue and light blue, we define the blue part as X_{11} and the light blue part as X_{12} in the second level. Correspondingly, we have Y_{11} which is blue and Y_{12} which is light blue in the second level of the lower chart of Fig. 2 as the outcome states of X_{11} and X_{12}. X_{11} and Y_{11} are relatively coherent structures, and so are X_{12} and Y_{12}.

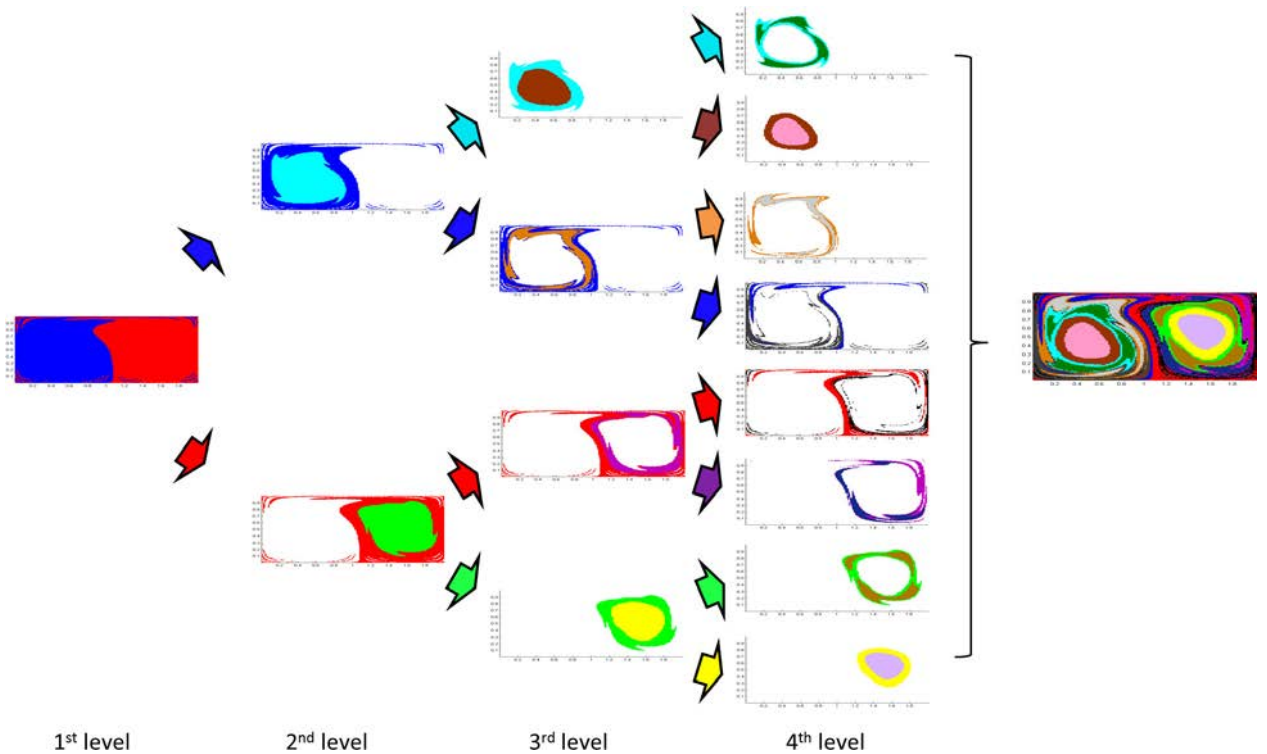
As above, we can develop an X_{21} that is red and an X_{22} that is light green from X_2 in the second level of the upper chart of Fig. 2; and Y_{21} that is red and Y_{22} that is light green from Y_2 in the second level of the lower chart of Fig. 2. The same subscript means X_{21} and Y_{21} are a relatively coherent pair, so are X_{22} and Y_{22}. Now we have four relatively coherent pairs in the second level.

We can eventually get the tree structures in Fig. 2 by repetition of the process. There are eight relatively coherent structures shown in the third level with different colors and 16 relatively coherent structures in the fourth level with different colors. In Fig. 2, we can see the egg-shaped relatively coherent structures which are four resonance "islands" as expected in such Hamiltonian twist maps [Meiss, 1992]. Even finer structures will be revealed by further refinement and sufficient sampling to allow appropriate resolution. Appropriate sampling in a given refinement scale is discussed in terms of computational complexity in developing a given Ulam–Galerkin's matrix for a given fine grid. Note that while closed for initial measure μ, it is open in all subsequent measures.

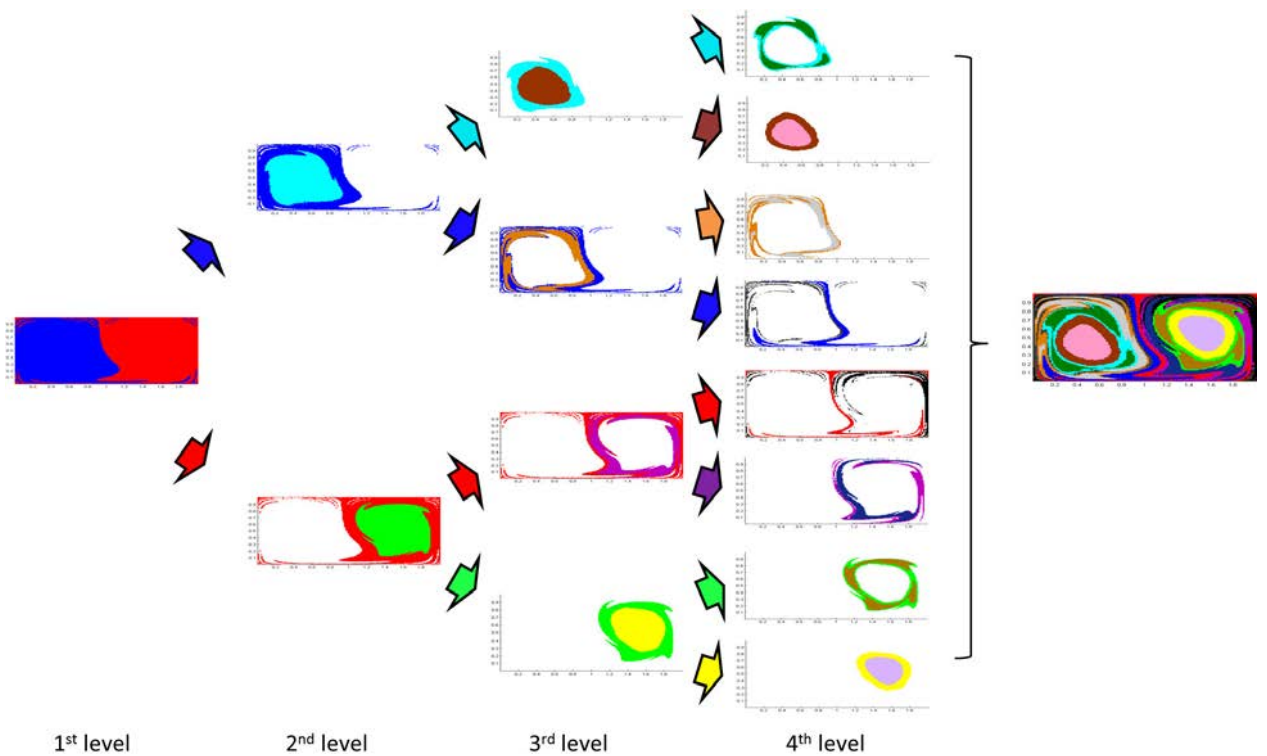
4.2. *An idealized stratospheric flow*

Next, consider the Hamiltonian system

$$\frac{dx}{dt} = -\frac{\partial \Phi}{\partial y}, \quad \frac{dy}{dt} = \frac{\partial \Phi}{\partial x}, \tag{17}$$



(a)



(b)

Fig. 2. The upper figure is the initial status of the double gyre and the lower figure is the image under the time-τ flow, where $\tau = 10$. The same color areas between the two associates one relatively coherent pair. In this case, we have 16 different relative coherent structures with different colors. By following the colored arrows, we can see the relative coherent structures through four levels of refinement. (a) $t = 0$ and (b) $t = 10$.

where

$$\begin{aligned}\Phi(x,y,t) &= c_3 y - U_0 L \tanh\left(\frac{y}{L}\right) \\ &\quad + A_3 U_0 L \operatorname{sech}^2\left(\frac{y}{L}\right)\cos(k_1 x) \\ &\quad + A_2 U_0 L \operatorname{sech}^2\left(\frac{y}{L}\right)\cos(k_2 x - \sigma_2 t) \\ &\quad + A_1 U_0 L \operatorname{sech}^2\left(\frac{y}{L}\right)\cos(k_1 x - \sigma_1 t). \end{aligned} \tag{18}$$

This is a quasiperiodic system that represents an idealized zonal stratospheric flow [Froyland *et al.*, 2010]. There are two known Rossby wave regimes in this system. Let $U_0 = 63.66, c_2 = 0.205U_0, c_3 = 0.7U_0, A_3 = 0.2, A_2 = 0.4, A_1 = 0.075$ and the other parameters in Eq. (18) be the same as stated in [Rypina *et al.*, 2007].

The Lipschitz constant M here is 0.7854. Equation (1) tells us the necessary number of points for a 32 640 by 39 694 matrix is about 15 000 000. We choose 20 000 000 points in the domain $X = [0, 6.371\pi * 10^6] \times [-2.5 * 10^6, 2.5 * 10^6]$ of the flow and use 32 640 triangles as the partition $\{B_i\}_{i=1}^{32\,640}$ for the initial status points and 39 694 triangles as the partition $\{C_j\}_{j=1}^{39\,694}$ for the final status of the points. Note that this system is "open" relatively to the domain X chosen, though it is a area-preserving flow. The two coherent pairs are colored blue and red which are defined as (X_1, Y_1) and (X_2, Y_2) at the first level of Fig. 3. Again, we now build the relative measures and tree of relatively coherent pairs. By applying the method as we have done with the previous two examples, we develop four and eight different coherent structures for the second and third levels, respectively. In Fig. 4, we repeat the color scheme in both the initial status and final status in each of the second and third levels for the relatively coherent structure. Thus, we now see a much finer scaled relative coherence in the dynamical system than previously seen.

4.3. *The standard map*

Consider the standard map,

$$\begin{aligned} p_{n+1} &= p_n + K\sin(\theta_n) \\ \theta_{n+1} &= \theta_n + p_{n+1} \end{aligned} \tag{19}$$

where p_n and θ_n are taken modulo 2π. See [Meiss, 1992], and this is a map on the torus $\Omega = [0, 1) \times [0, 1)$. We will study the case that $K = 1.2$ in our example, as this is well known to be $K = 1.2 > K_{\text{cr}} = 0.971635$ shortly after the breakup of the last "golden" torus allowing momentum boosting orbits and a mixed chaotic and ordered phase space including periodic elliptic islands [Meiss, 1992].

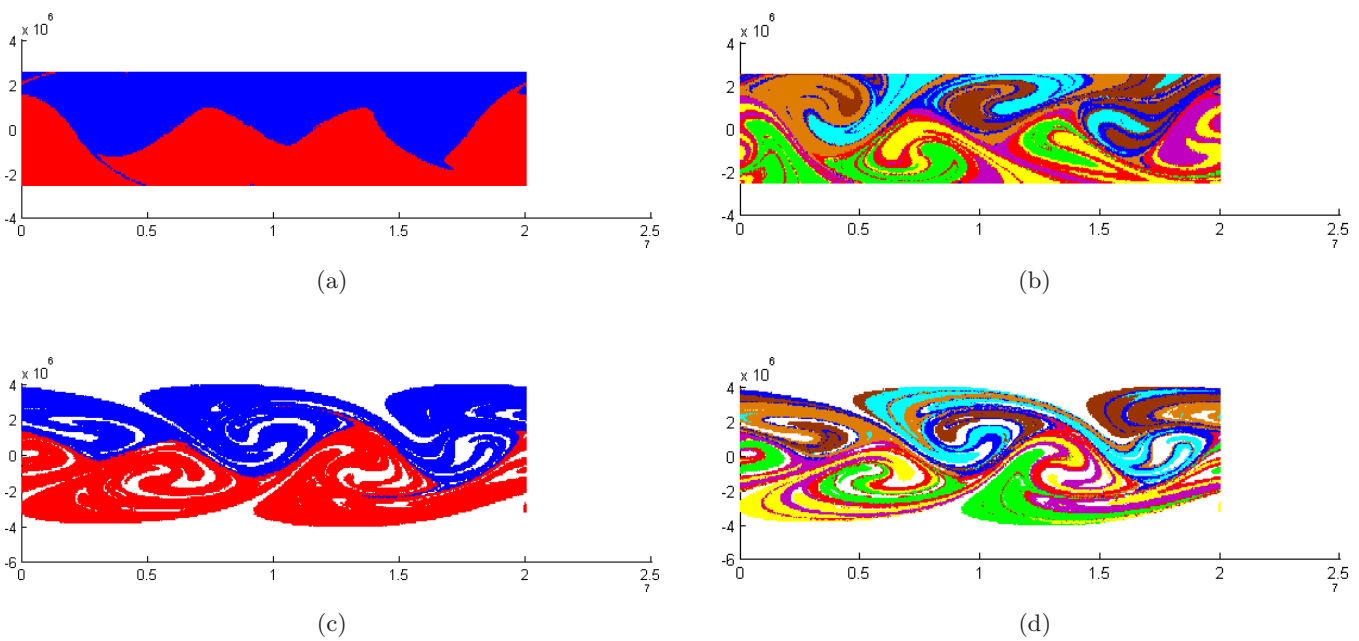
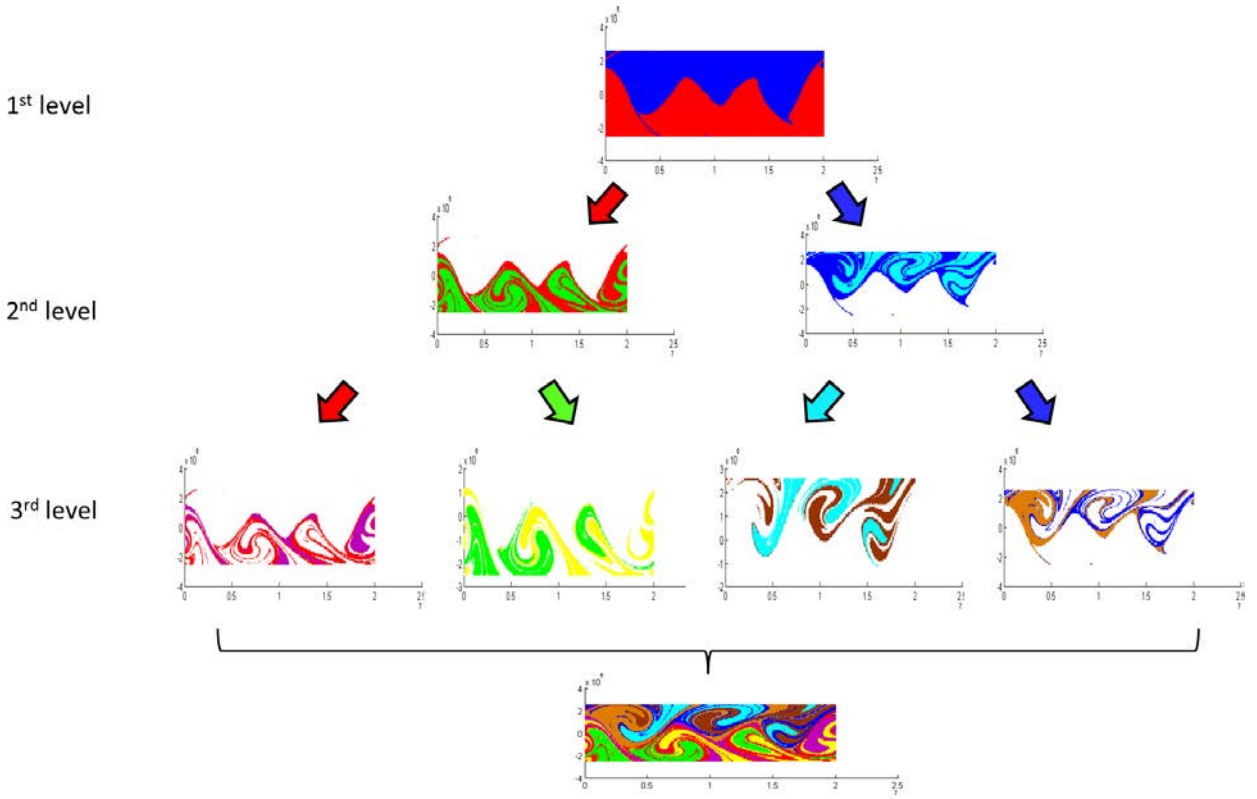
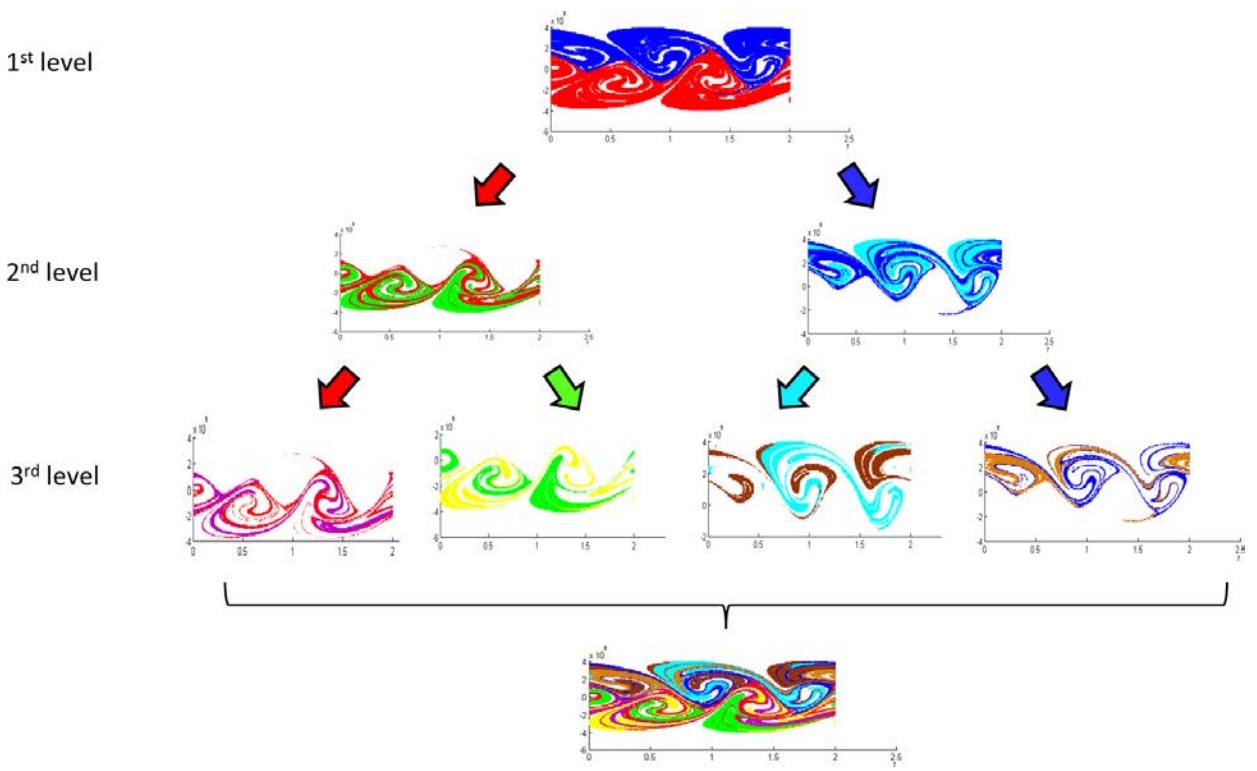


Fig. 3. Relative coherence in the Rossby system Eq. (18). The first level partition (left-hand side) and third level partition (right-hand side) of both the initial and final status of the zonal flow. Compare to the hierarchical structure as emphasized in Fig. 4. (a) and (b) $t = 0$; (c) and (d) $t = 10$ days.



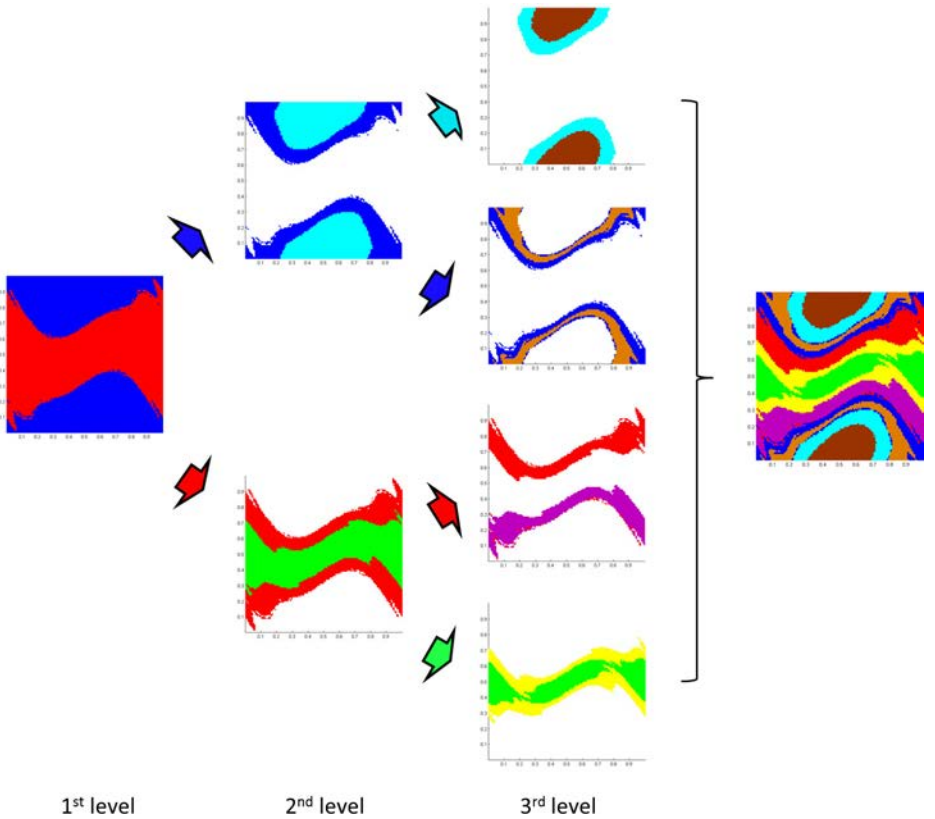
(a)



(b)

Fig. 4. Relative coherence hierarchy in the Rossby system Eq. (18). Coloring and hierarchy tree structure as in (a) and (b). Compare also to Fig. 3. (a) $t = 0$ and (b) $t = 10$ days.

(a)



(b)

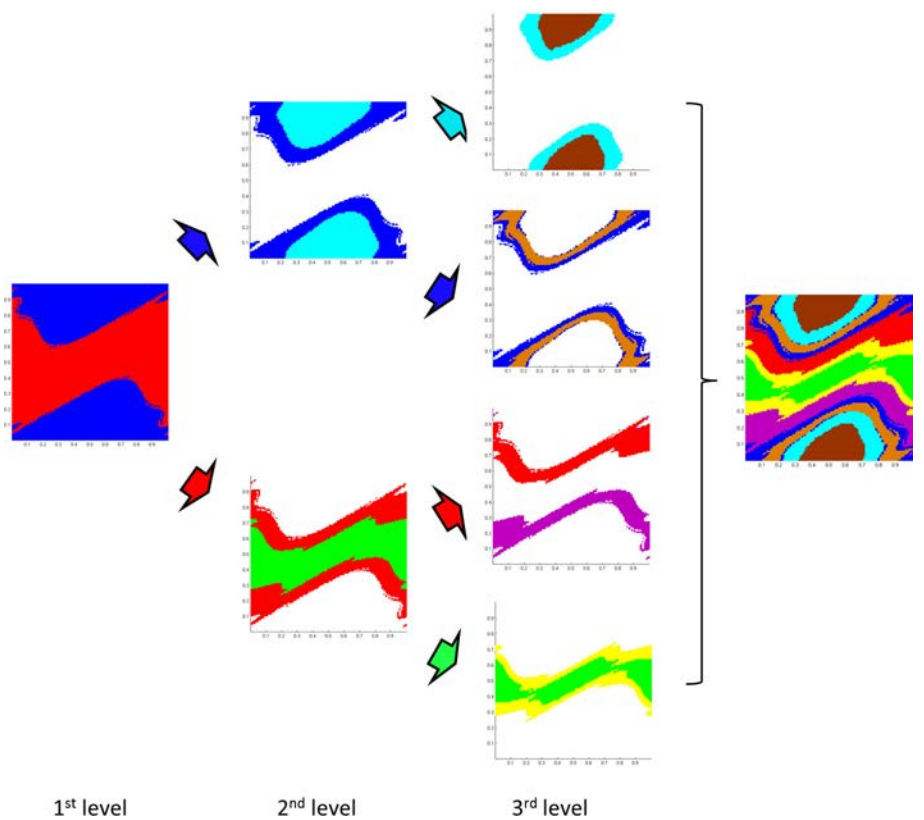


Fig. 5. Standard Map Eq. (19), structured relative coherence hierarchy tree arranged as in (a) and (b). (a) At the beginning and (b) after 10 iterations.

The first level of Fig. 5 shows two large coherent pairs in the Standard Map, which are filled by blue and red, respectively. By prior knowledge on this benchmark problem, it is clear that the boundary between the primary blue and red coherence estimates the cantorus remnant of the golden area resonance, an expected act came as this is known to remain a primary pseudo-barrier to transport when $K = 1.2$, still not much larger than $K_{\text{cr}} = 0.971635$.

Iteratively repeating the process according to the algorithm for each of X_1, Y_1, X_2 and Y_2 yields four relatively coherent structures which are colored blue, light blue, red and green in the second level. The third level in Fig. 5 tells us there are eight different such structures in total in the first three levels. The outcome partition here shows a familiar depiction of the resonance layers known to be due to cantorus pseudo-barriers which cause the famously slow transport for the standard map [Meiss, 1992]. Notice that the lack of stopping criterion could be due to the great deal of symmetries in this system.

4.4. *The Gulf of Mexico*

In our last example, we consider the Mexico Gulf. The data is the same as was used in [Bollt *et al.*, 2010] and formed by the method in [Bleck, 2002; Halliwell, 2004]. The difference between the Gulf model and the above three examples in the Gulf is an open system, that is, there is water entering and exiting the region. See Fig. 6. This is the reason why in Fig. 7, the shapes of the whole Gulf water of the initial status and final status are slightly different at the bottom and top regions.

By Theorem 1, we need at least 12 807 000 points, the Lipschitz constant M of this model is 2.6911. We choose 20 000 000 points uniformly and randomly in the water region as the initial status, with more details of data. See [Bollt *et al.*, 2010]. The final status is the position of these points after 6 days. We use 32 867 triangles $\{B_i\}_{i=1}^{32\,867}$ as a partition of X and 32 359 triangles $\{C_j\}_{j=1}^{32\,359}$ as a partition of Y. After applying our subdivision method on these triangles, the results are shown in Fig. 7. In this example, we set $\rho_0 = 0.9998$ as the threshold, the stopping criterion. Therefore, the number

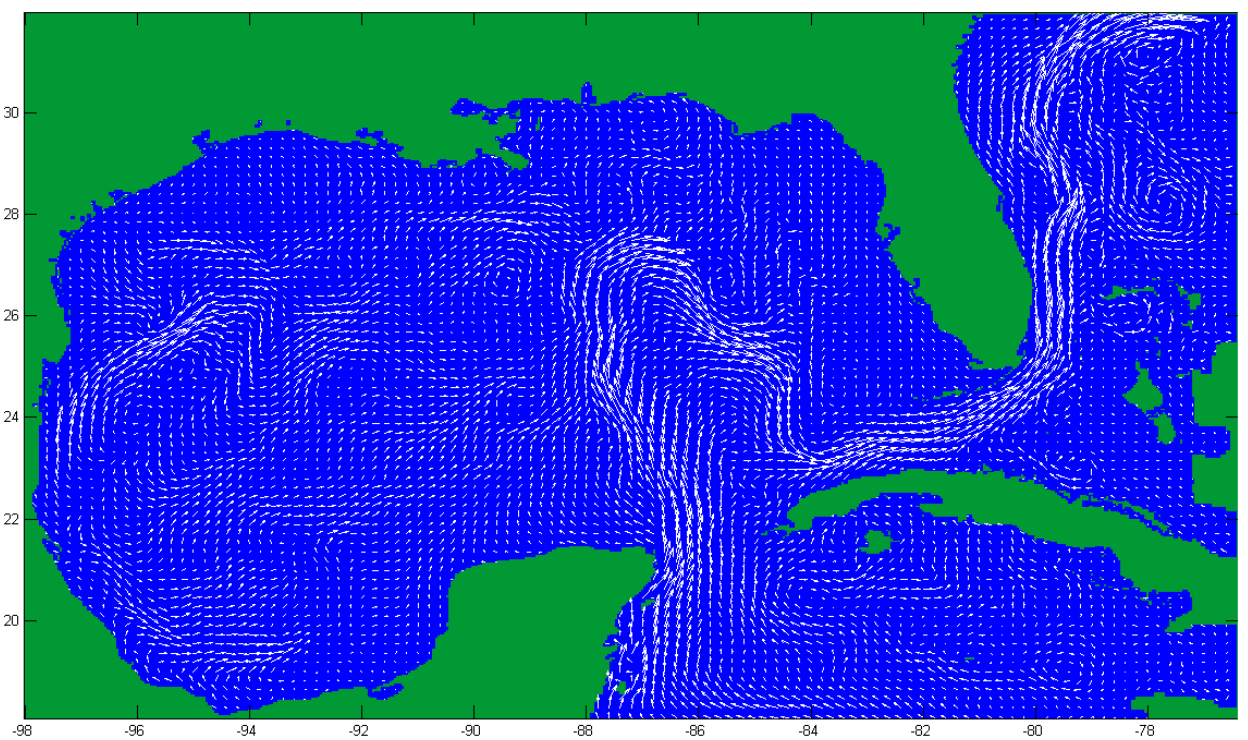
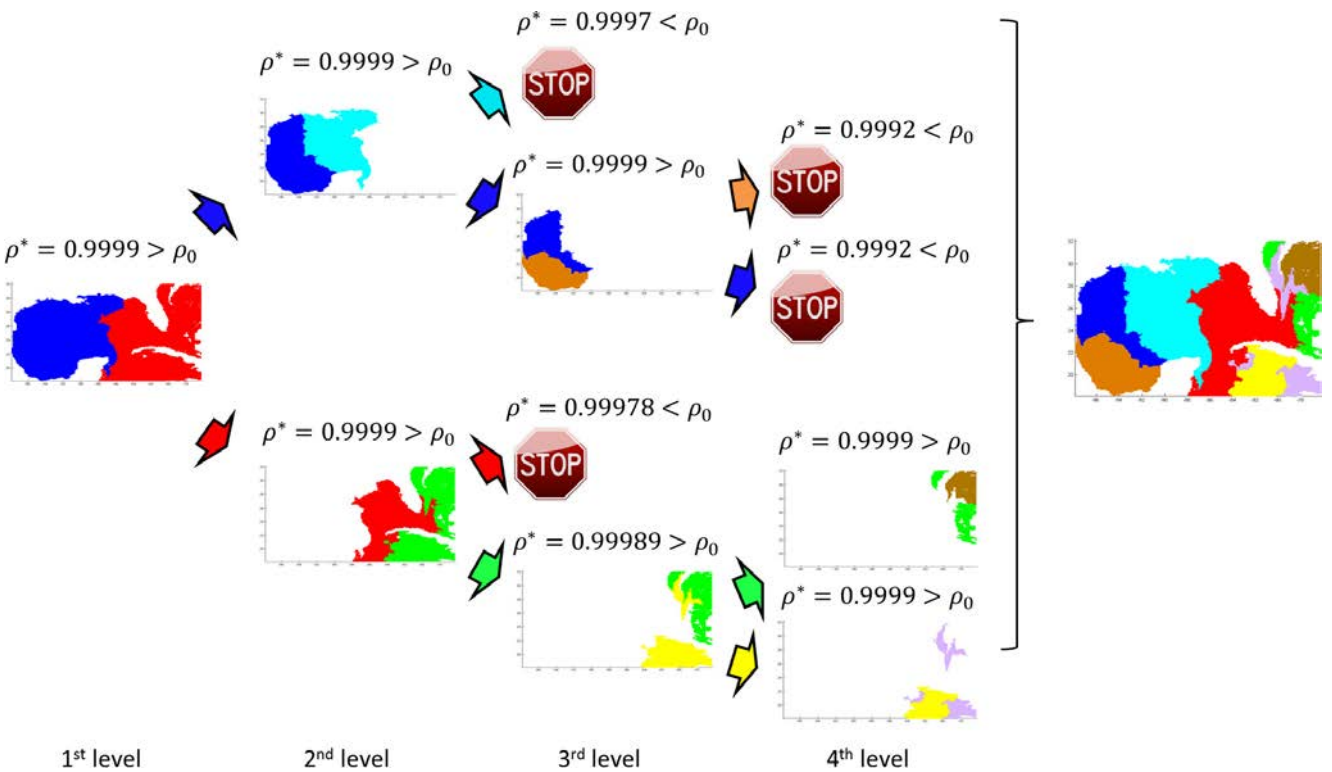
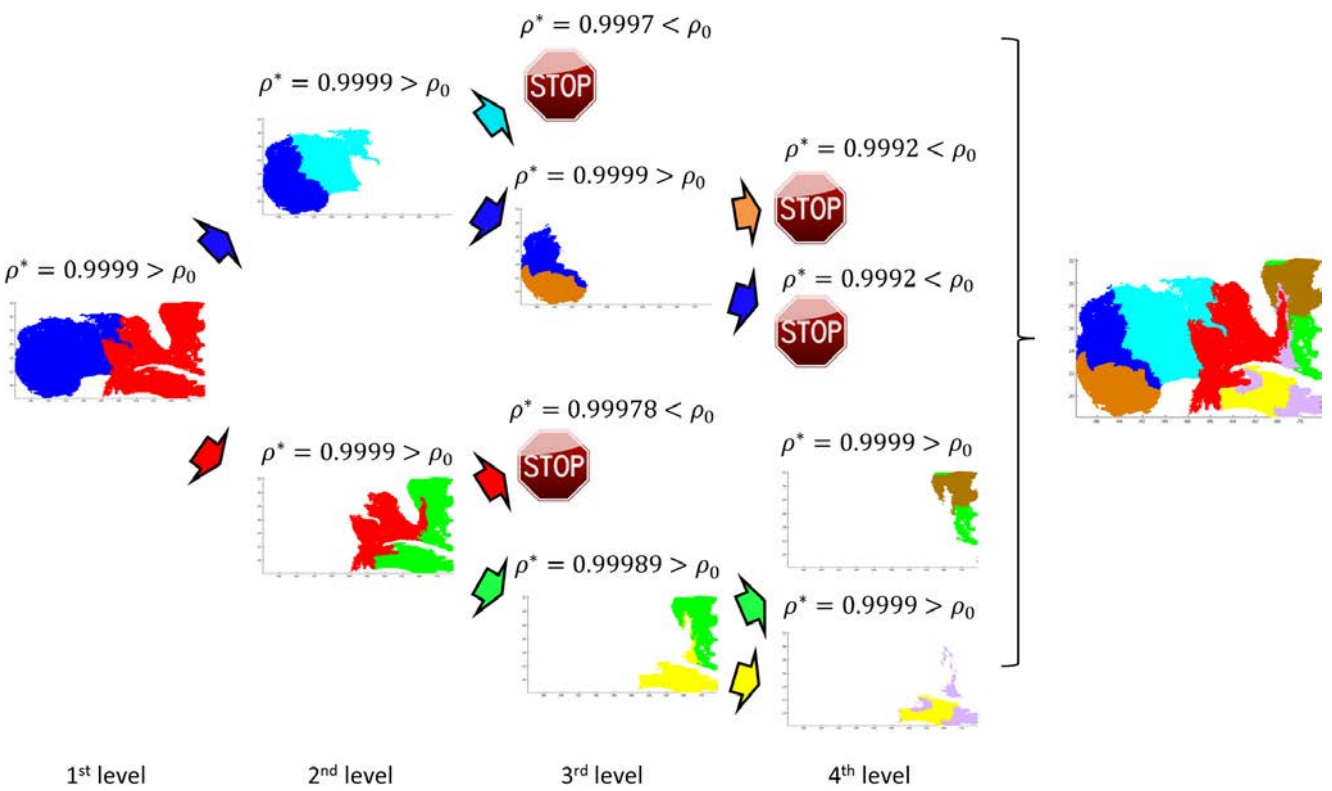


Fig. 6. Vector field describing surface flow in the Gulf of Mexico on May 24, 2010, computed using the HYCOM model [HYCOM, 2010]. Note the coherence of the Gulf Stream at this time. Oil spilling from south of Louisiana could flow directly into the Gulf Stream and out towards the Atlantic. This is an open system relative to this window shown. Horizontal and vertical units are degrees longitude (negative indicates west longitude) and degrees latitude (positive indicates north latitude), respectively.



(a)



(b)

Fig. 7. Hierarchical relative coherence in the Gulf of Mexico following the flow according to vector field data as illustrated in Fig. 6. Tree structure and relatively coherent pairs coloring as in (a) and (b). (a) $t = 0$ and (b) $t = 6$ days.

of relatively coherent pairs is not equal 2^q, where q is defined in Eq. (11).

5. Conclusions

We have defined a concept of relative coherence based on relative measure, as a generalization of coherent pairs. We have also introduced a recursive method of detecting relatively coherent structures under flows in a finite time, based on relative measures, with respect to the restricted Frobenius–Perron operator. Relative measures are used to build a hierarchy of relatively coherent pairs at successive levels, which can be illustrated in a natural tree structure of relative coherence.

We have demonstrated the method with the double gyres, the standard map, a Rossby wave system and data from the Gulf of Mexico. These examples have included hierarchical structure, open and closed systems, a system known only through data, and the use of the stopping criterion.

References

Bleck, R. [2002] "An oceanic circulation model framed in hybrid isopycnic-Cartesian coordinates," *Ocean Model.* **4**, 55–88.

Bollt, E. M., Luttman, A., Kramer, S. & Basnayake, R. [2010] "Measurable dynamics analysis of transport in the Gulf of Mexico during the oil spill," *Int. J. Bifurcation and Chaos* **22**, to appear.

Bollt, E. M. & Santitissadeekorn, N. [2012] *Applied and Computational Measurable Dynamics* (SIAM, to appear, 2013), pp. 73–101; 105–153.

Dellnitz, M. & Junge, O. [1998] "An adaptive subdivision technique for the approximation of attractors and invariant measures," *Comput. Vis. Sci.* **1**, 63–68.

Dellnitz, M. & Junge, O. [2000] "Set oriented numerical methods for dynamical systems," *Handbook of Dynamical Systems* (Elsevier B.V.), Chapter 5, pp. 221–264.

Ding, J., Li, T. Y. & Zhou, A. [2002] "Finite approximations of Markov operators," *J. Comput. Appl. Math.* **147**, 137–152.

Froyland, G. & Padberg, K. [2009] "Almost-invariant sets and invariant manifolds — Connecting probabilistic and geometric descriptions of coherent structures in flows," *Physica D* **238**, 1507–1523.

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

Guder, R., Dellnitz, M. & Kreuzer, E. [1997] "An adaptive method for the approximation of the generalized cell mapping," *Chaos Solit. Fract.* **8**, 525–534.

Guder, R. & Kreuzer, E. [1999] "Control of an adaptive refinement technique of generalized cell mapping by system dynamics," *J. Nonlin. Dyn.* **20**, 21–32.

Halliwell, G. R. [2004] "Evaluation of vertical coordinate and vertical mixing algorithms in the HYbrid-Coordinate Ocean Model (HYCOM)," *Ocean Model.* **7**, 285–322.

HYCOM [2010] "Hybrid Coordinate Ocean Model (HYCOM)," http://www.hycom.org/.

Junge, O. [2001] "An adaptive subdivision technique for the approximation of attractors and invariant measures: Proof of convergence," *Dyn. Syst.* **16**, 213–222.

Lasota, A. & Mackey, M. C. [1994] *Chaos, Fractals, and Noise Stochastic Aspects of Dynamics*, 2nd edition (Springer).

Meiss, J. D. [1992] "Symplectic maps, variational principles, and transport," *APS, Rev. Mod. Phys.* **64**, 795–848.

Perko, L. [2006] *Differential Equations and Dynamical Systems*, 3rd edition (Springer).

Rypina, I. I., Brown, M. G., Beron-Vera, F. J., Kocak, H., Olascoaga, M. J. & Udovydchenkov, I. A. [2007] "On the Lagrangian dynamics of atmospheric zonal jets and the permeability of the stratospheric polar vortex," *J. Atmosph. Sci.* **64**, 3595–3610.

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

Appendix A

On Computational Complexity

In this section, we analyze the computational complexity to properly develop an Ulam–Galerkin matrix. The following discussion is based on the Lipschitz constant and Gronwall's inequality. See [Perko, 2006]. There are some other works related to this topic. See [Dellnitz & Junge, 1998; Guder *et al.*, 1997; Guder & Kreuzer, 1999; Junge, 2001].

The arguments here are premised on a very simple idea that all we need to consider is the boundary before and the boundary after a flow. Let $f : X \times [t_0, t] \to X$ be a flow and X be compact. We use $X(t_0)$ and $X(t)$ to denote the status of X at time t_0 and t under the flow f, respectively. For an arbitrary point $x \in X$, we set $x(t_0)$ and $x(t)$ as the positions of x at time t_0 and t. Consider a square box with length q, we use such size q box to make partitions for $X(t_0)$ and $X(t)$, which are $B = \{B_i\}_{i=1}^m$ and $C = \{C_j\}_{j=1}^n$, m and n are positive integers.

Theorem A.1 [Theorem 1]. *For a Lipschitz flow f in the plane $X \in \mathbb{R}^2$ and partitions $B = \{B_i\}_{i=1}^m$ and $C = \{C_j\}_{j=1}^n$ consist of identical boxes with length $q > 0$ for $X(t_0)$ and $X(t)$, where $t_0 < t$. Let $P_{i,j}$ be from Eq. (6) and $\hat{P}_{i,j}$ be from Eq. (5). For each $P_{i,j}$ built by N uniformly sampled points from a B_i-q-box, $\exists$ is a confidence coefficient $r \in (0,1)$ which depends on the Lipschitz constant M of f, and the density δ of sampling points N. According to Eqs. (A.7) and (A.13), it follows that $r \to 0$ as $N \to \infty$.*

Note that the confidence coefficient r may vary for different B_i. If we use local Lipschitz constants for each box B_i and this discussion could lead to adaptive grids, but here we simply use a uniform Lipschitz constant from a global grid.

As a direct conclusion, we have the following corollary.

Corollary A.1 [Corollary 4.1]. *Alternatively, choosing $r \in (0,1)$ for a B_i box, the number of uniform sampling points $N(r, B_i)$ of B_i box follows from Eqs. (A.7) and (A.13), such that we can control the error $|P_{i,j} - \hat{P}_{i,j}|$ term by term as desired.*

Notice that the number $N(r, B_i)$ may vary for different i and the total sample points number is the sum of all $N(r, B_i)$, $i = 1, 2, \ldots, m$, see Eq. (A.14). The corollary allows us to develop adaptive sampling for different boxes designed according to local Lipschitz constant values rather than the global Lipschitz M we use in the theorem.

Theorem 1 may be used to estimate the term by term errors between the matrices P and $\hat{P}$. And Corollary 4.1 is always applied on deciding the total number of uniform sampled points. Therefore they both are important for the numerical approximation of Ulam's matrix.

Before we prove the theorem, we consider the error types in the matrix building process. Suppose $\hat{P}_{i,j} = \alpha$, which means if we uniformly and ideally choose N sample points in B_i box, there will be $\text{Round}(N\alpha)$ points or $\text{Round}(N\alpha)+1$ points in the C_j box. We will explain the word "ideally" at the end of subsection B.

We now discuss another type of error. For convenience of illustration, we assume B_i is a box of partition of $X(t_0)$, there will be a box C_{j-1} of the partition of $X(t)$ sharing all of the boundary of B_i. That is, they totally overlap. The triangulation we use is handled similarly, but rectangles will simplify this discussion even if triangles allow for the powerful Delaunay triangulation algorithms in practice. For convenience, we discuss here computational cost of boxes instead of triangles, but one box can be easily changed to two triangles by cutting through the diagonal.

First, recall that, the Gronwall's inequality is,

$$|x_1(t) - x_2(t)| \le |x_1(t_0) - x_2(t_0)| e^{M|t-t_0|}, \quad \text{(A.1)}$$

where M is the Lipschitz constant. Assuming f is uniformly continuously differentiable,

$$M = \max_{(x,t)\in X\times[t_0,t]} \left|\frac{\partial f}{\partial x}(x,t)\right|. \quad \text{(A.2)}$$

A.1. *An open cover*

We assume the time interval $[t_0, t]$ is relatively short. Without loss of generality, we choose a q box B_i as a subdomain, there will be a C_{j-1} at the same place as B_i. Figure 8 is an illustration of the subdomain under a flow f in the short time interval. That is, the different time status of the box overlap in the most major part. However, we wish to catch the behaviors of the flow as much as possible, so we need a proper number of initial points in the domain so that they can be transported everywhere including the small nonoverlapping region, the shadow region in Fig. 8, under the flow.

Intuitively, we may choose a finite open cover of the shadow region and study the preimages of the elements of the cover. See Fig. 9. Since an open cover for a compact space X under a flow is still an open cover, the preimages of open sets still cover X at time t_0. Then we choose a smallest radius of all the preimages as the largest distance among random points. Generally, we catch most of the domain

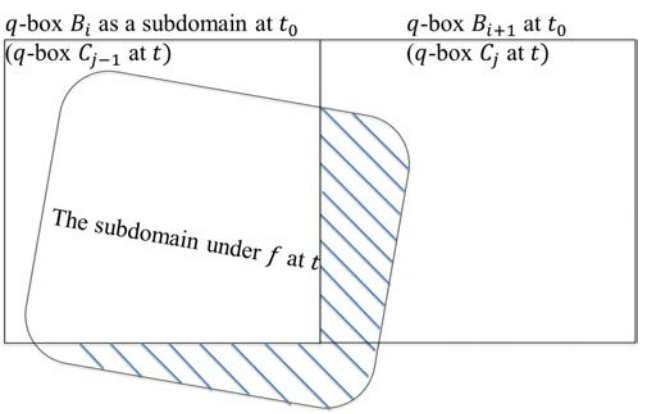


Fig. 8. A small square shape subdomain under flow f through time $t - t_0$.

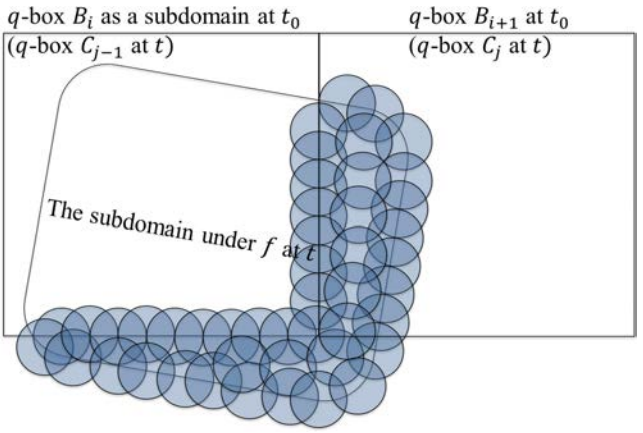


Fig. 9. A finite open cover of the shadow region.

in this way, however, there will still be a problem on building the Ulam–Galerkin matrix. Figure 10 shows a small slice of the shadow region, we now discuss why a usual open cover may not work for our case.

From now on, without of loss of generality, we focus on building the entry $P_{i,j}$. We consider three scenarios of open neighborhoods in the cover. See the colored circles on the right-hand side of Fig. 11. The three open neighborhoods can be denoted by

$$|x_u(t) - x(t)| < l \tag{A.3}$$

where $x_u, u = 1, 2, 3$ are centers of the circles and $l > 0$. By Gronwall's inequality, there is a

$$\delta = \frac{l}{e^{M|t-t_0|}} \tag{A.4}$$

such that if $x(t_0)$ in one of the open neighborhoods defined as follows, the three open neighborhoods on the left-hand side of Fig. 11,

$$|x_u(t_0) - x(t_0)| < \delta \tag{A.5}$$

where $u = 1, 2, 3$ and $x_u(t_0)$ are the centers of the neighborhoods, $x(t)$ will go to one of the three open neighborhoods on the right-hand side of Fig. 11, respectively.

To understand this well, we place all circles in partitions of $X(t_0)$ and $X(t)$. In Fig. 12, B_i and B_{i+1} are boxes from the partition of $X(t_0)$. By our assumption, we set C_{j-1} as the one totally overlapping with B_i, as a direct conclusion, so is C_j with respect to B_{i+1}. See Fig. 12.

We assume that we have built the open sets small enough so that we can only afford to choose only one point from one set due to heavy computation. Now we have the following cases where we are trying to build the Ulam–Galerkin matrix P.

(1) An arbitrary point from the green one in B_i must map across the boundary of B_i and into C_j. Therefore, the Ulam–Galerkin matrix P will have an adjusted value of $P_{i,j}$.
(2) An arbitrary point from the yellow one across B_i and B_{i+1} will adjust to $P_{i,j}$ or $P_{i+1,j}$.
(3) An arbitrary point from the red one in B_i will effect $P_{i,j}$ or $P_{i,j-1}$.

When we build the matrix P, we expect that all the points we choose randomly in a neighborhood only change one certain item of the matrix P, such as $P_{i,j}$. The reason is, for example, for the red one in $X(t)$ of Fig. 12, we can see part of it is still in the C_{j-1} box after flow f, if we choose the initial point $x(t_0)$ in the red one in $X(t_0)$ such that $x(t)$ in $C_j - 1$ in $X(t)$, we lose the information of part of the red one in C_j when we build the matrix.

If we go back to check Eqs. (5) and (6), we easily find that the case that all points chosen from such circles covering $T^{-1}(C_j)$ must be in $T^{-1}(C_j)$, since they focus on $\hat{P}_{i,j}$ and $P_{i,j}$. In other words, each entry of the matrix need more points to describe the mass change, however, Cases 2 and 3 indicate if one can be more accurate, the other one will be effected. A way to eliminate this kind of interaction is to try to avoid choosing points from yellow and

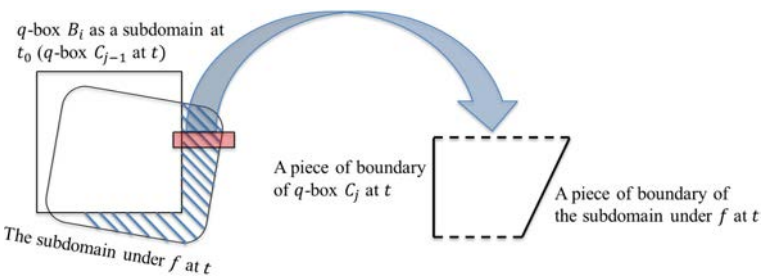


Fig. 10. A small slice of the shadow region with part of boundary.

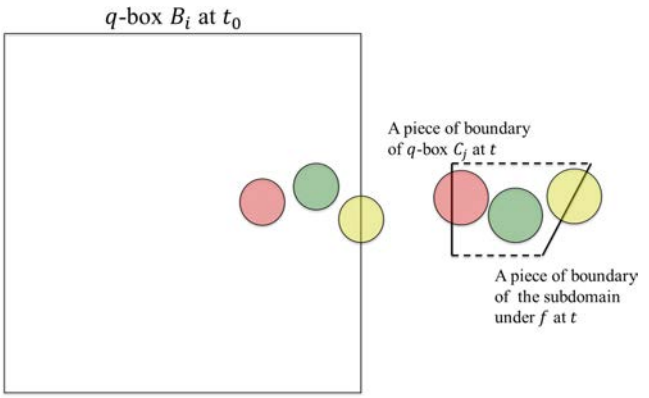


Fig. 11. The circles on the right-hand side are open neighborhoods we choose, and the corresponding colored circles on the left-hand side are from Gronwall's inequality.

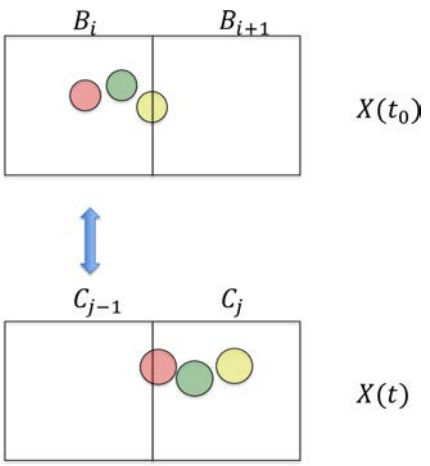


Fig. 12. The open neighborhoods in partitions of $X(t_0)$ and $X(t)$.

red ones, such that the entries will be independent of each other. However, we do not know the specific positions of these red and yellow ones in $X(t_0)$. And also if so, we lose information of red and yellow ones.

A.2. *An adjusted open cover*

Consider shrinking the discs on the trapezoid region's boundary in Fig. 11, so that we can reduce the probability of sampling points chosen from yellow and red discs. Figure 13 shows that we can

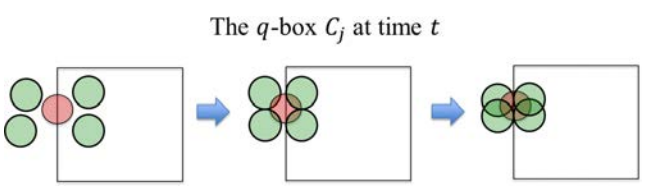


Fig. 13. The process to reduce an open neighborhood.

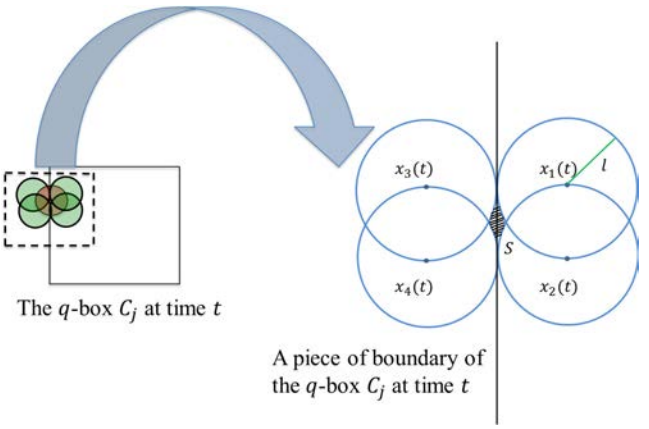


Fig. 14. The structure built by four green open neighborhoods in $X(t)$.

reduce a red open neighborhood by four green open neighborhoods. We assume all points from a green neighborhood only change one element in the matrix, according to Case 1. Then we build the structure as shown in Fig. 14 by drawing four circles with radius l. The region S surrounded by the four circles is the region that leads to Case 2 or Case 3. Notice this structure is in $X(t)$; we now consider the status of the structure in time t_0 which is in $X(t_0)$. It can be shown that relative position of the four circles will hold at time t_0 by topology properties of the flow. See, Fig. 15. Moreover, by Gronwall's inequality, the radius δ of these circles at time t_0 is smaller than l; and then the surrounding region S' at time t_0 will be smaller than S.

We can decrease the area of S' by shrinking the area of S. On the other hand, the smaller l is, the smaller S will be. Therefore, we can reduce the probability of a random point chosen from S' by making a smaller S', which indicates we need a

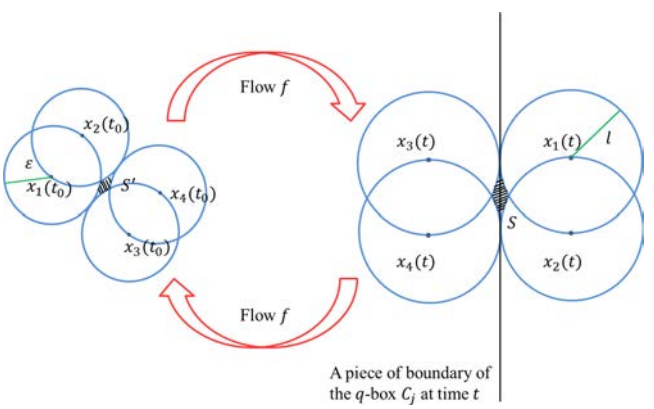


Fig. 15. The relative position of the circles hold for different times.

smaller l. However, as l becomes smaller, the radius δ will become smaller by Gronwall's inequality. If we set the largest distance among the uniform initial points to be smaller than δ, we need more sample points when δ become smaller. By the geometry of the S' structure, the relationships between δ and the area of S' is,

$$\mu(S') = \delta^2 \left(2 - \frac{\sqrt{3}}{2} - \frac{\pi}{3}\right), \tag{A.6}$$

where μ is the measure defined on $X(t_0)$. Thus, the size of S' cannot be zero, which will lead to infinitely many initial points. We have to use a confidence coefficient as a measure for the probability we discussed above.

At the end of this subsection, we talk about choosing points *ideally*. We say, if all the initial points are chosen from $X(t_0)$ except all the S' area, these points are sampling *ideally*.

A.3. *A confidence coefficient*

Note that there are finite number S' regions in B_i box due to a finite open cover. We now consider a ratio number r such that,

$$r = \frac{\sum_{S' \in B_i \cap (\bigcup_j f^{-1}(C_j))} \mu(S')}{\mu(B_i)}. \tag{A.7}$$

Notice in the theorem, we use uniformly sampling points which automatically eliminates the effect of $P_{i+1,j}$ from Case 2, but not the whole Case 2. Even for uniform sampling points, we still do not know if some points lie in such S' regions. Thus, r can be considered as a confidence coefficient that indicates the probability Case 2 or Case 3 occurs. Note that r is for all C_j related to a fixed B_i. If we consider an r for a B_i, it turns out r is the error from all missed transitions about B_i.

Then the only question left is how many S' regions are in B_i box. It is not easy to find the number of S' directly. However, since the number of S' regions in B_i is the number of S regions on the part of the shadow region's boundary, see Fig. 16, which is a direct conclusion from our discussion in the above section, we can calculate the number of S instead of S'.

We next give a method to count the number of S. For consistency, we use the same shape shadow region, however, the method can deal with a more general case. We now only consider the shadow

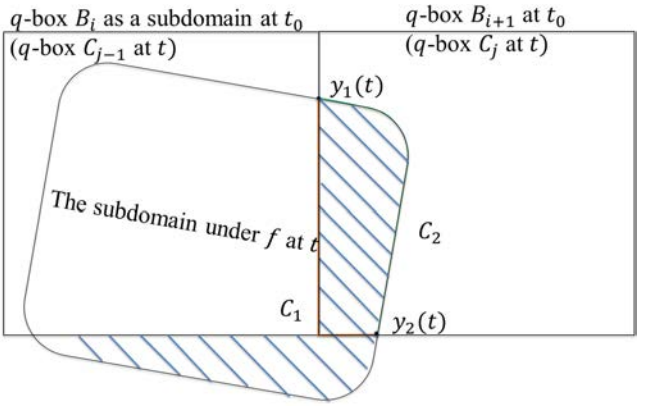


Fig. 16. The shadow of region's boundary is split into an orange curve C_1 and a green curve C_2.

region's boundary in C_j. In Fig. 16, we separate part of the shadow region by two curves with different colors, C_1 and C_2. We use $\mathrm{Num}(S, C_1)$ and $\mathrm{Num}(S, C_1)$ to denote the number of S on C_1 and C_2, respectively; and $|C|$ denotes the length of a curve C with respect to the corresponding measure. And then we have the following relationships,

$$\begin{aligned}&\mathrm{Num}(S', B_i, C_j)\\ &\quad = \mathrm{Num}(S, C_1) + \mathrm{Num}(S, C_2),\end{aligned} \tag{A.8}$$

where $\mathrm{Num}(S', B_i, C_j)$ denotes the number of S' in B_i related to C_j.

For curve C_1, the orange one, it is easy to get $\mathrm{Num}(S, C_1)$, for it consists of line segments,

$$\mathrm{Num}(S, C_1) \approx \left\lceil \frac{|C_1|}{l} \right\rceil. \tag{A.9}$$

For curve C_2, however, $|C_2|$ is hard to be determined directly. Remember that the S regions on C_2 correspond to some S' regions in B_i, so we go back to study the status of C_2 at time t_0.

In Fig. 17, we can see there are two intercepts $y_1(t)$ and $y_2(t)$. We consider the positions of y_1 and y_2 at time t_0, i.e. $y_1(t_0)$ and $y_2(t_0)$. $y_1(t_0)$ and $y_2(t_0)$ must lie on the part of q-box B_i's boundary, by assuming an orientation preserving property. We conclude that C_2 at time t_0 must be the red line segments in Fig. 17, denoted by $C_2(t_0)$, so the number of S' on $C_2(t_0)$, $\mathrm{Num}(S', C_2(t_0))$, is

$$\mathrm{Num}(S', C_2(t_0)) \approx \frac{1}{2} \left\lceil \frac{|C_2(t_0)|}{\delta} \right\rceil, \tag{A.10}$$

where δ is from Eq. (A.4) for the same l in Eq. (A.9). There is a $1/2$, because for any S' on $C_2(t_0)$ we only have half the area inside B_i by the structure we have

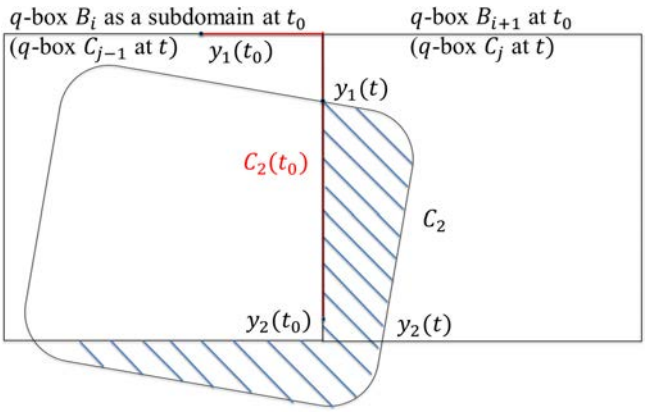


Fig. 17. The red line segment is the status of C_2 at time t_0, which is denoted by $C_2(t_0)$.

built. On the other hand, the following relationships hold, as we discussed,

$$\mathrm{Num}(S, C_2) = \mathrm{Num}(S', C_2(t_0)). \quad (A.11)$$

Thus, by Eqs. (A.8)–(A.11), we can rewrite $\mathrm{Num}(S', B_i)$ as,

$$\mathrm{Num}(S', B_i, C_j) = \left\lceil \frac{|C_1|}{l} \right\rceil + \frac{1}{2} \left\lceil \frac{|C_2(t_0)|}{\delta} \right\rceil. \quad (A.12)$$

Moreover, by Eqs. (A.12) and (A.6), Eq. (A.7) becomes

$$r = \frac{\sum_j \mathrm{Num}(S', B_i, C_j) m(S')}{m(B_i)}. \quad (A.13)$$

Actually, Eqs. (A.4) and (A.13) can be considered as a function between r and l for a given flow and a partition. Therefore, we can use a given r to solve l and then we can get δ_i for each B_i. Normally, we set $r \leq 0.5\%$. Now, we have the equation of the necessary number of initial points for a partition $B = \{B_i\}_{i=1}^k$,

$$\text{Initial Points' Number} = \sum_{B_i} \left\lceil \frac{q^2}{\delta_i^2} \right\rceil. \quad (A.14)$$

A.4. *Proof of Theorem 1*

By following the above subsections, we give a brief proof for Theorem 1.

Given a $P_{i,j}$ built by N uniformly sampling points from B_i q-box with density δ, according to Eqs. (A.13) and (A.7), by adding δ, we can get a r. Clearly, the equations tell us when $r \to 0, \delta \to 0$ and then $N \to \infty$.

Moreover, suppose $\hat{P}_{i,j} = \alpha, \alpha \in (0, 1)$ and the points are ideally chosen. Then we have $\mathrm{Round}(N\alpha)+1$ or $\mathrm{Round}(N\alpha)$ points in C_j for the ideal case. On the other hand, for another error, the worst case is when Nr points all are effected in $P_{i,j}$. Thus, for the general $P_{i,j}$ not built by ideally chosen points, we have

$$|P_{i,j} - \hat{P}_{i,j}| \leq \frac{Nr}{N} + \frac{\alpha}{N} < \frac{r+1}{N}. \quad (A.15)$$

The cases $\alpha = 0, 1$ mean we have all points in C_j or no point in C_j, and then $|P_{i,j} - \hat{P}_{i,j}| = 0$. As we proved above, when $r \to 0, 1/N \to 0$, so P converges to $\hat{P}$ term by term, for each i, j, where $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Theorem 1 tells us, if we uniformly choose enough points, no matter where the positions of these points, we can get an appropriate approximation of Ulam's matrix related to r. The theorem is a strong support for numerically computing the Ulam's matrix without considering the sampling problem. In practice, each r and δ are nearly the same for different B_i because of relatively short time and Lipschitz flow f.