Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Network-based study of Lagrangian transport and mixing

Kathrin Padberg-Gehle[1] and Christiane Schneide[1]

[1]Leuphana Universität Lüneburg, Institute of Mathematics and its Didactics, Applied Mathematics Group, Scharnhorststr. 1, D-21335 Lüneburg, Germany

Correspondence to: padberg@leuphana.de

Abstract. Transport and mixing processes in fluid flows are crucially influenced by coherent structures and the characterization of these Lagrangian objects is a topic of intense current research. While established mathematical approaches such as variational or transfer operator based schemes require full knowledge of the flow field or at least high resolution trajectory data, this information may not be available in applications. Recently, different computational methods have been proposed to identify coherent behavior in flows directly from Lagrangian trajectory data. In this context, spatio-temporal clustering algorithms have been proven to be very effective for the extraction of coherent sets from sparse and possibly incomplete trajectory data. Inspired by these recent approaches, we consider an unweighted, undirected network, where Lagrangian particle trajectories serve as network nodes. A link is established between two nodes if the respective trajectories come close to each other at least once in the course of time. Classical graph concepts are then employed to analyze the resulting network. In particular, local network measures such as the node degree serve as indicators of highly mixing regions, whereas spectral graph partitioning schemes allow us to extract coherent sets. The proposed methodology is very fast to run and we demonstrate its applicability in two geophysical flows - the Bickley jet as well as the antarctic stratospheric polar vortex.

1 Introduction

The notion of coherence in time-dependent dynamical systems is used to describe mobile sets that do not freely mix with the surrounding regions in phase space. In particular, coherent behavior has a crucial impact on transport and mixing processes in fluid flows. The mathematical definition and numerical study of coherent flow structures has received considerable scientific interest for the last two decades. The proposed methods roughly fall into two different classes, geometric and probabilistic approaches, see Allshouse and Peacock (2015) for a discussion and comparison of different methods. Geometric concepts aim at defining the boundaries between coherent sets, i.e. codimension-1 material surfaces in the flow that can be characterized by variational criteria (see Haller (2015) for a recent review). Central to these constructions is the Cauchy-Green strain tensor, which is derived from the derivative of the flow map. Thus, full knowledge of the flow field or at least high resolution trajectory data is required for these methods to work successfully. This applies also to other geometric concepts such as shape coherence (Ma and Bollt (2014)). Probabilistic methods aim at defining sets that are minimally dispersive while moving with the flow. The main theoretical tools are transfer operators, i.e. linear Markov operators that describe the motion of probability densities under the action of the nonlinear, time-dependent flow. The different constructions in this family of approaches are reviewed in Froyland and Padberg-Gehle (2014), also highlighting the crucial role of diffusion in this setting. Recently, a dynamic

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics Discussions
Open Access
EGU

Laplacian framework has been introduced by Froyland (2015), where explicit diffusion is no longer required in the analytical and computational framework. While for this approach fast and accurate algorithms have been developed in Froyland and Junge (2015), the classical transfer operator setting requires the integration of many particle trajectories for the numerical approximation of the infinite dimensional operator. Here again, full knowledge of the underlying dynamical system is needed, which may not be available in applications. Moreover, all discussed approaches assume that the nonautonomous dynamics is represented by a flow map, which, by construction, only considers the starting and end points of each particle trajectory, but neglects the dynamics between the initial and final points in time.

To overcome these problems, different computational methods have been proposed to identify coherent behavior in flows directly from Lagrangian trajectory data, such as obtained from particle tracking algorithms. One of the earliest attempts is the braiding approach proposed by Allshouse and Thiffeault (2012), where trajectories are classified according to their intertwining pattern in space-time. This method is mathematically sound, but computationally demanding and currently restricted to two-dimensional flows. Trajectory-based approaches have also been introduced by Mancho et al. (2013) and Budišić and Mezić (2012). They use time-integrated quantities along trajectories, which again requires knowledge of the underlying dynamical system. Finally, Williams et al. (2015) attempt to reconstruct the transfer operator from limited amount of trajectory data.

Very recently, spatio-temporal clustering algorithms have been proven to be very effective for the extraction of coherent sets from sparse and possibly incomplete trajectory data (see e.g. Froyland and Padberg-Gehle (2015); Hadjighasem et al. (2016); Banisch and Koltai (2016); Schlueter-Kuck and Dabiri (2016)). Here, distance measures between trajectories are used to define groups of trajectories that remain close and/or behave similarly in the time span under investigation. All these methods can deal with sparse and incomplete trajectory data and do respect the dynamics of the entire trajectories, not just the end points. While c-means clustering as used by Froyland and Padberg-Gehle (2015) is computationally inexpensive and works well in example systems (see also Allshouse and Peacock (2015)), spectral clustering approaches as in Hadjighasem et al. (2016); Banisch and Koltai (2016); Schlueter-Kuck and Dabiri (2016), appear to be more robust, but require considerable computational effort.

Inspired by these recent approaches, our aim is to design a reliable but computationally inexpensive method for studying coherent behavior as well as mixing processes directly from Lagrangian trajectory data. For this, we consider an unweighted, undirected network, where Lagrangian particle trajectories serve as network nodes. A link is established between two nodes if the respective trajectories come close to each other at least once in the course of time. This construction is similar in spirit to the concept of recurrence networks (see e.g. Donner et al. (2010)), but here in a spatio-temporal setting. We note that also the discretized transfer operator has been viewed and treated as a network, see e.g. Dellnitz and Preis (2003); Dellnitz et al. (2005); Padberg et al. (2009); Lindner and Donner (2016); Ser-Giacomi et al. (2015).

We use classical graph concepts and algorithms to analyze our trajectory-based flow network. Local network measures such as node degree or clustering coefficient highlight regions of strong mixing and thus capture very similar information as described by the trajectory encounter number very recently introduced by Rypina and Pratt (2016). In particular, spectral graph partitioning schemes for the solution of a balanced cut problem (Shi and Malik (2000)) allow us to efficiently extract coherent sets of the underlying flow.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

The paper is organized as follows. In section 2 we describe our network construction. This is followed by a discussion of network analysis tools in section 3, where we review several, simple local network measures as well as the spectral graph partitioning approach by Shi and Malik (2000). In section 4 we apply the methodology to two different example systems, a Bickley jet as well as the stratospheric polar vortex. We close the paper with a discussion and an outlook on future work.

2 Networks of Lagrangian flow trajectories

In the following, we assume that we have $n \in \mathbb{N}$ Lagrangian particle trajectories from a flow simulation or from a particle tracking experiment in physical space $\mathbb{R}^d$, $d = 2$ or 3. In practice, the particle positions may be given at discrete times $\{0, 1, \dots, T\}$. We denote the trajectories by x_i, $i = 1, \dots, n$, and its position at a certain time instance $t = 0, \dots, T$ by $x_{i,t} \in \mathbb{R}^d$. We now set up a network of these Lagrangian trajectories $x_1, \dots, x_n$ and link two trajectories if they come ϵ-close to each other at least once in the course of time. Such an undirected, unweighted network is uniquely described by a symmetric adjacency matrix $A \in \{0,1\}^{n,n}$. In practice, we construct this from the given Lagrangian flow trajectories by setting

$$A_{ij} = \begin{cases} \max\limits_{0 \le t \le T} \chi_{B_\epsilon(x_{i,t})}(x_{j,t}), & i \neq j \\ 0, & i = j \end{cases}, \tag{1}$$

where χ_B denotes the indicator function of a set $B \subset \mathbb{R}^d$. So $A_{ij} = 1$, that is, a link is established between trajectories x_i and x_j, if and only if at one or more time instances t, $x_{j,t}$ can be found in an ϵ-ball $B_\epsilon(x_{i,t})$ centered at $x_{i,t}$ and thus the trajectories x_i and x_j have come ϵ-close.

By an appropriate choice of ϵ one ensures that the network defined by (1) is connected. For instance, if the trajectories are initialized on a regular grid, then a natural lower bound to ϵ would be the mesh size. In the case that particles are randomly distributed, ϵ has to be chosen accordingly. We will study different choices of ϵ in section 4.

Alternatively, the network might be set up by linking the k-nearest neighboring trajectories at each time instance for some $k \in \mathbb{N}$. While this allows us to get rid of the problem of a suitable choice of ϵ it means that we have to choose a reasonable k. In two-dimensional systems a natural choice would be $k = 4$ mimicking the five point stencil, similarly in three-dimensional systems $k = 6$. If trajectories are initialized on a regular grid this choice again ensures that the resulting network is connected. Our own preliminary studies have indicated that this procedure gives very similar results to the ϵ-based definition in (1) but requires slightly longer computational run times. Therefore, we will not pursue this construction in the present work.

3 Network analysis

Here, we briefly discuss standard analysis concepts for networks (see e.g. Newman (2003)) and relate them to features of the underlying flow. In particular, we will describe how to extract coherent structures by solving a graph partitioning problem, the balanced minimum cut problem as proposed by Shi and Malik (2000) (see also Hadjighasem et al. (2016)).

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

3.1 Degree matrix and graph Laplacian

From the adjacency matrix A one can derive two other important matrices to describe the network. The degree matrix D is a diagonal matrix with $D_{ii} = d_i$ where d_i is the degree of node x_i, i.e. $D_{ii} = \sum_{j=1}^{n} A_{ij}$, that is the number of links attached to node i. In our setting, $d_i \in \mathbb{N}$, $i = 1, \ldots, n$. By construction, for our network the degree of a node is non-zero, so there are no isolated nodes.

The non-normalized Laplacian is formed by $L = D - A$, where D is the degree matrix and A the adjacency matrix. By the construction of A and D, L is symmetric and the entries of L are

$$L_{ij} = \begin{cases} -A_{ij} & i \neq j \\ D_{ii}, & i = j \end{cases}. \tag{2}$$

and thus $L \in \mathbb{Z}^{n,n}$.

The normalized symmetric graph Laplacian $\mathcal{L} \in \mathbb{R}^{n,n}$ is defined as

$$\mathcal{L} = I_n - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}. \tag{3}$$

$\mathcal{L}$ has non-negative real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. $w_1 = D^{\frac{1}{2}}\mathbf{1}$ is eigenvector to eigenvalue $\lambda_1 = 0$. The other eigenvalues and corresponding eigenvectors can be characterized variationally in terms of the Rayleigh quotient of $\mathcal{L}$. We come back to this in section 3.3.

3.2 Local network measures

Node degree

The degree of a node encodes how many other nodes are connected to it and thus it measures the immediate importance of that node. In our setting, it measures how many different trajectories come close to the trajectory represented by the respective node. The node degree d is encoded in the diagonal elements $d_i = D_{ii}$ of the diagonal degree matrix D, with

$$d_i = \sum_j A_{ij}, \quad i = 1, \ldots, n. \tag{4}$$

Average degree of neighboring nodes

Here one considers the average node degree of the neighbors of a node x_i, defined as

$$\langle d \rangle_{nn,i} = \frac{\sum_j A_{ij} d_j}{d_i}, \quad i = 1, \ldots, n. \tag{5}$$

Due to the averaging over all neighboring degrees, $\langle d \rangle_{nn}$ appears to be smoother compared to the simple node degree d. Both the node degree d and the average degree of neighboring nodes $\langle d \rangle_{nn}$ are very related to the trajectory encounter number as recently introduced by Rypina and Pratt (2016). Similar concepts of stretching measures are discussed in Padberg et al. (2009) and Froyland and Padberg-Gehle (2012). Thus, both d and $\langle d \rangle_{nn}$ are expected to be large in mixing regions.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Local clustering coefficient

The local clustering coefficient C indicates how strongly connected certain subgraphs are. It does so by measuring what proportion of the neighbors of x_i are neighbors themselves:

$$C_i i = \frac{\#\text{ triangles connected to } x_i}{\#\text{ triples centered around } x_i} = \frac{(A^3)_{ii}}{d_i(d_i-1)}, \quad i = 1, \ldots, n. \tag{6}$$

In contrast to the quantities that measure the number of other trajectories that a particle encounters in the course of time, the local clustering coefficient C is expected to be large for trajectories in regular regions (i.e. for which d or $\langle d \rangle_{nn}$ is small).

3.3 Spectral graph partitioning

As outlined above, the normalized symmetric graph Laplacian $\mathcal{L}$ has non-negative real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. The second smallest eigenvalue $\lambda_2 \geq 0$ is called algebraic connectivity or Fiedler eigenvalue of a graph (Fiedler (1973)). This eigenvalue is non-zero if and only if the network is connected. More generally, the number of connected components of the network appears as the multiplicity of the eigenvalue zero of the Laplacian matrix. If $\lambda_2 > 0$ but very close to zero, then the network is nearly decoupled and the sign structure of the corresponding eigenvector determines the two communities in the network (Fiedler cut) If λ_i,. $i = 2, \ldots, k$ for some $k < n$ are close to zero and there is a spectral gap between λ_k and λ_{k+1}, then the network is nearly decoupled into k communities. The corresponding eigenvectors $w_2, \ldots, w_k$ carry information about the location of these communities. This can be verified by considering the Rayleigh quotient of the normalized graph Laplacian. Shi and Malik (2000) used this concept to solve a balanced cut problem for defining communities in the network that are characterized by minimum communication between different communities and maximum communication within communities. Such nearly decoupled subgraphs correspond to bundles of trajectories that are internally well connected but only loosely connected to other trajectories. This is indicative of coherent behavior (see also Hadjighasem et al. (2016)). Instead of considering the eigenvalue problem $\mathcal{L}w = \lambda w$, Shi and Malik (2000) propose to solve the equivalent generalized eigenvalue problem

$$Lv = \lambda Dv. \tag{7}$$

As both L and D are symmetric and have integer entries, eigenvalue problem (7) turns out to be numerically more convenient than the original one. It has the same eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ and the eigenvectors are related by $w_i = D^{\frac{1}{2}} v_i$, $i = 1, \ldots, n$. In particular, $v_1 = \mathbf{1}$. The number of leading eigenvalues (i.e. eigenvalues close to zero) indicates the number of nearly decoupled subgraphs. An application of a standard k-means clustering algorithm can then be employed to extract the sets of interest from the corresponding eigenvectors.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

4 Examples

4.1 Bickley jet

As our first example we consider the Bickley jet proposed by Rypina et al. (2007). It is defined by the streamfunction

$$\Psi(x,y,t) = -U_0 L \tanh(y/L) + \sum_{i=1}^{3} A_i U_0 L \,\mathrm{sech}^2(y/L)\cos(k_i x - \sigma_i t). \tag{8}$$

We use same the parameter values as in Hadjighasem et al. (2016); Banisch and Koltai (2016), i.e. $U_0 = 5.414$, $A_1 = 0.0075$, $A_2 = 0.15$, $A_3 = 0.3$, $L = 1.770$, $c_1/U_0 = 0.1446$, $c_2/U_0 = 0.205$, $c_3/U_0 = 0.461$, $k_1 = 2/r_e$, $k_2 = 4/r_e$, $k_3 = 6/r_e$ where $r_e = 6.371$ as well as $\sigma_i = c_i k_i$, $i = 1,2,3$. Here, we have dropped the physical units for brevity. The physical assumptions underlying the model equations and the parameters are described in detail in Rypina et al. (2007).

Initial conditions are chosen in the domain $M = [0,20[\times[-3,3]$ and are numerically integrated on the time interval $[10,30]$ using a 4th order adaptive Runge-Kutta scheme and periodic boundary conditions in x-direction. We output the particle positions at integer time steps. We consider two sets of initial conditions, which we will refer to as cases (i) and (ii) in the following:

(i) 12,200 points from a regular grid on M with grid size 0.1

(ii) 1,000 random points uniformly distributed on M.

For the first high-resolution setting (i) we choose $\epsilon = 0.1$, corresponding to the distance between neighboring grid points, as well as $\epsilon = 0.15$ and $\epsilon = 0.2$. For the sparse setting (ii), we consider $\epsilon = 0.5$. Significantly smaller values of ϵ did not produce a connected network in this case. For comparison purposes, we also test $\epsilon = 1$ and $\epsilon = 2$ in this case.

In Figure 1 the local network measures for case (i) are plotted with respect to the initial conditions. The left column contains the results for $\epsilon = 0.1$ and the right column for $\epsilon = 0.2$. The top row displays the node degree d. As expected, d is high in mixing regions, i.e. where trajectories meet many other trajectories and low in the regular regions, i.e. the six vortices and the jet core. Whereas the result for $\epsilon = 0.1$ appears a bit spurious, that for $\epsilon = 0.2$ is much sharper. The average node degree of neighboring nodes $\langle d \rangle_{nn}$ (middle row) gives a very pronounced indication of regular and mixing flow behavior, even for $\epsilon = 0.1$. The bottom row shows the clustering coefficient C. It has the expected high values for regular regions only for $\epsilon = 0.2$ (and $\epsilon = 0.15$, not shown), but appears to have some singularities in the vortex cores for $\epsilon = 0.1$. Overall, the average node degree $\langle d \rangle_{nn}$ appears to give the clearest and most robust indication of regular and mixing regions in the flow.

In Figure 2 we repeat the study for the low-resolution case (ii), using $\epsilon = 0.5$ (left column) and $\epsilon = 1$ (right column). The results are very much comparable to the high-resolution case (i), with the average node degree $\langle d \rangle_{nn}$ (middle row) giving again a good indication of the different flow regimes, especially for small ϵ, where the node degree d only produces spurious results. For both choices of ϵ the local clustering coefficient C picks up the cores of the six vortices. We note that when increasing ϵ further (not shown), both the node degree and clustering coefficient will continue to highlight mixing and vortical regions respectively, whereas the average node degree will fail at some point as - due to the enlarged neighborhood - it will average over different flow regimes and thus blur the local information.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

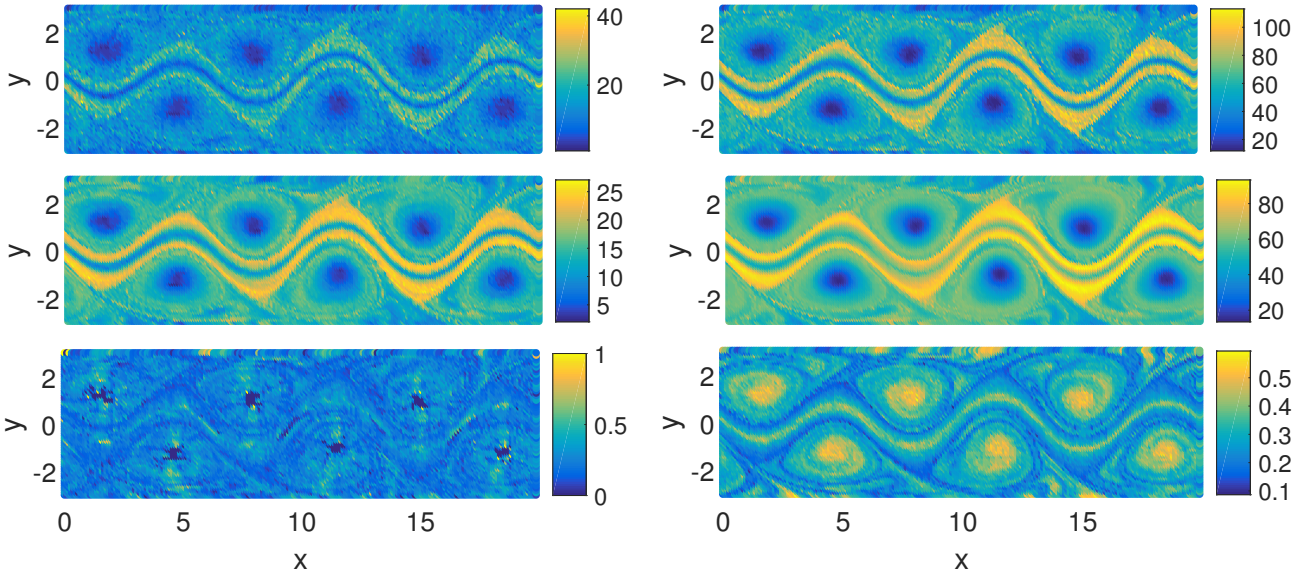

Figure 1. Network measures for high resolution initial conditions (case (i)) in the Bickley jet for $\epsilon = 0.1$ (left) and $\epsilon = 0.2$ (right) . From top to bottom: node degrees d and $\langle d \rangle_{nn}$ and clustering coefficient C.

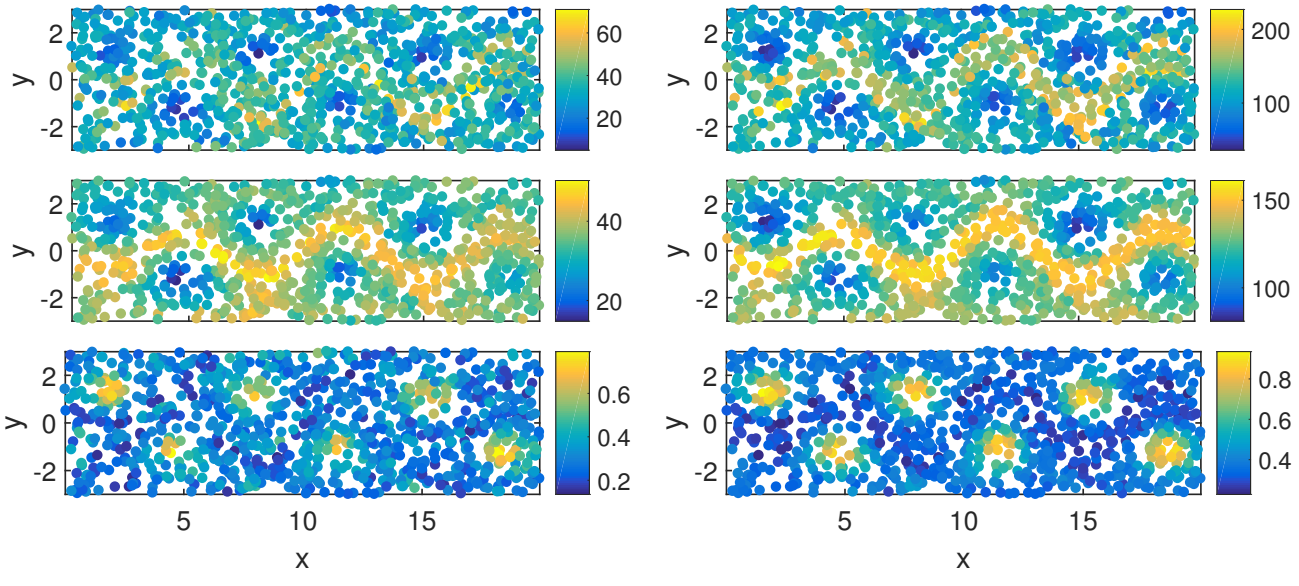

Figure 2. Network measures for 1,000 random initial conditions (case (ii)) in the Bickley jet for $\epsilon = 0.5$ (left) and $\epsilon = 1$ (right). From top to bottom: node degrees d and $\langle d \rangle_{nn}$ and clustering coefficient C.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

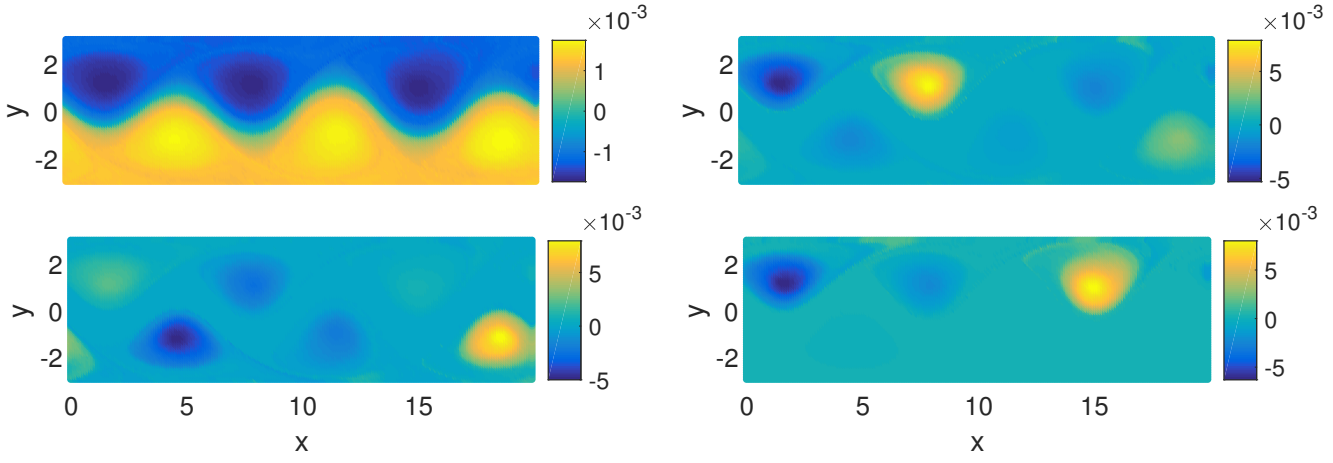

Figure 3. Leading eigenvectors $v_2 - v_5$ (from top left to lower right) of the generalized graph Laplacian eigenvalue problem (7) for the network constructed from high resolution initial data in the Bickley jet (case (i)) with $\epsilon = 0.2$.

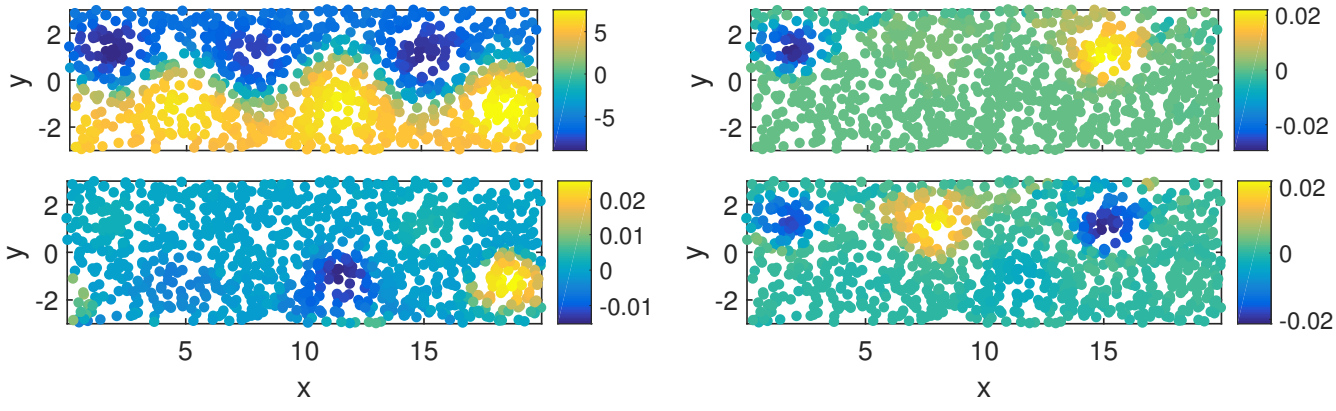

Figure 4. Leading eigenvectors $v_2 - v_5$ (from top left to lower right) of the generalized graph Laplacian eigenvalue problem (7) for the network constructed from 1,000 random initial conditions in the Bickley jet (case (ii)) and $\epsilon = 0.5$.

In Figure 3 the four (non-trivial) leading eigenvectors $v_2, \ldots, v_5$ of the generalized eigenvalue problem (7) are shown for the high-resolution initial conditions (case (i)) with $\epsilon = 0.2$. The eigenvectors highlight the two regions delineated by the jet as well as the different vortices, comparable to the results in Banisch and Koltai (2016). We note that the corresponding figures for $\epsilon = 0.1$ and $\epsilon = 0.15$ would look the same. In the low-resolution case (ii), the leading eigenvectors match those of the high-resolution data case, but in a slightly different order (see Figure 4). This comes from the fact that the four eigenvalues $\lambda_3, \ldots \lambda_5$ all have approximately the same magnitude and are therefore sensitive to perturbations. The ten leading eigenvalues for case (i) and $\epsilon = 0.2$ are displayed in Figure 5 (left), the low resolution case (ii) with $\epsilon = 0.5$ in the middle. These spectra exhibit clear spectral gaps between the second and the third and between the eighth and the ninth eigenvalues.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

The first spectral gap is related to the coherent behavior of the upper and lower part of the cylinder, delineated by the jet core. The second (and larger) spectral gap indicates the existence of altogether eight coherent sets. These can be extracted via a standard k-means clustering (with $k = 8$) of the first eight eigenvectors. The resulting partitions are shown in Figure 6. As expected, the six vortices and the two distinct stream regions are picked up, both in the high resolution (i) and the sparse data case (ii). For the low resolution case (ii) and a choice of $\epsilon = 1$ (or larger) the spectrum is no longer correctly recovered (see Figure 5 (right)).

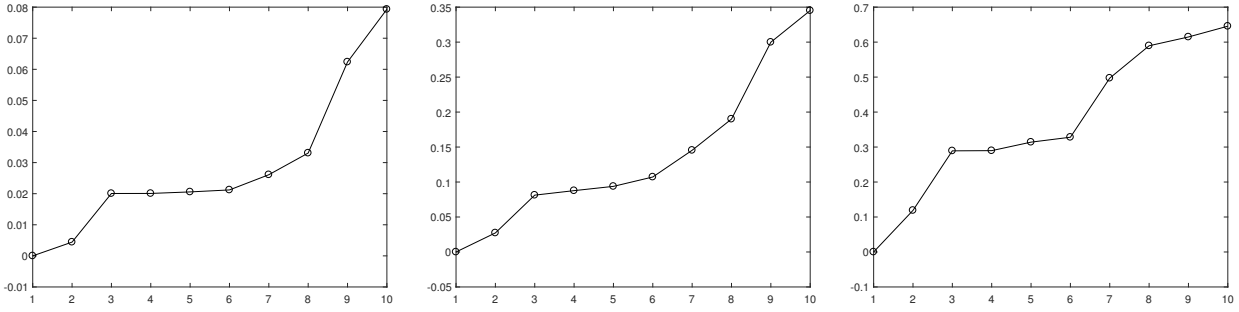

Figure 5. Leading eigenvalues of the generalized graph Laplacian eigenvalue problem (7) for the Bickley jet. Left: high-resolution data (case (i), $\epsilon = 0.2$); middle: sparse data case (ii), $\epsilon = 0.5$; right: case (ii), $\epsilon = 1$.

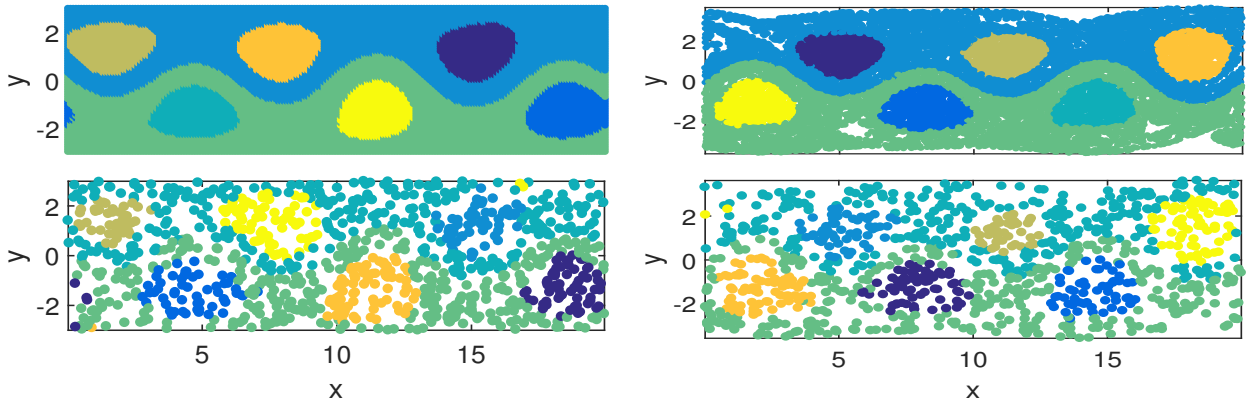

Figure 6. Extraction of eight coherent sets based on a k-means clustering of the eight leading eigenvectors of the generalized eigenvalue problem for the Bickley jet. Coherent sets at initial time ($t = 10$, left) and at final time ($t = 30$, right). Top: high-resolution case (i), $\epsilon = 0.2$; bottom: 1,000 random initial conditions (case (ii)) with $\epsilon = 0.5$.

Finally, we note that the proposed approach is computationally inexpensive, with total run times of $< 2s$ for the sparse data case (ii) and $< 40s$ for the high resolution case (i) using MATLAB (R2016a) on a single processor, see Table 1 for details.

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Table 1. Computation times (in seconds)

	trajectory integration	computation of A	eigenvalue problem
(i) 12,200 points ($\epsilon = 0.1$)	13.4s	23.0s	1.1s
(i) 1,000 points ($\epsilon = 0.5$)	1.6s	0.15s	0.14s

4.2 Stratospheric polar vortex

As a second example we study the transport and mixing dynamics in the stratospheric polar vortex over Antarctica. The coherent behavior of the polar vortex has already been numerically studied using transfer operator methods (Froyland et al. (2010)). For the computation of particle trajectories we use two-dimensional velocity data from the ECMWF Interim data set[1]. The global ECMWF data is given at a temporal resolution of 6 hours and a spatial resolution of a 121×240 grid in longitude and latitude directions respectively. As in Froyland et al. (2010) we focus on the stratosphere over the southern hemisphere. We consider the flow from September 1, 2002 to October 31, 2002 on a 600 K isentropic surface. For the integration of particle trajectories, we seed initial data on a 64×64 grid centered at the South Pole (square of side lengths $12,000$km), with a mesh size of 187.5km. A 4th order Runge Kutta scheme with a constant step size of 45min and linear interpolation in space and time are used and we output the particle positions every six hours. For the construction of the trajectory network we choose $\epsilon = 187.5$km, i.e. the grid spacing.

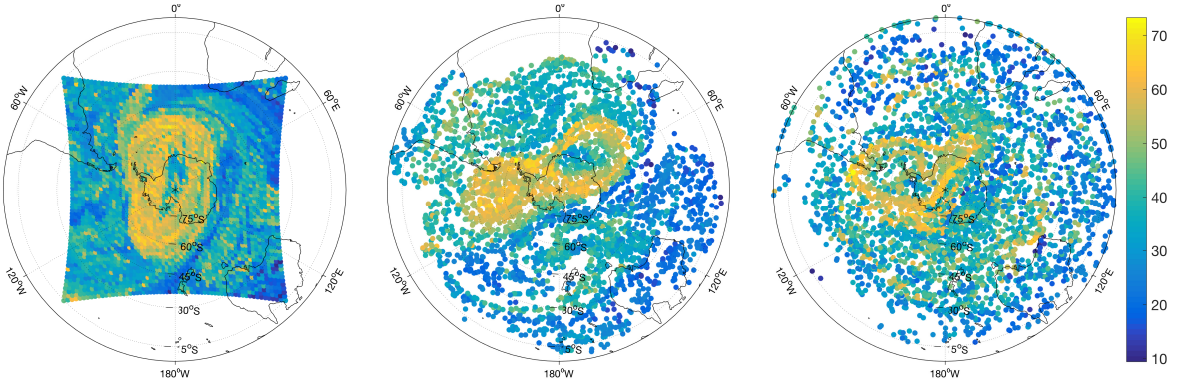

Figure 7. Average node degree of neighboring nodes $\langle d \rangle_{nn}$ of a network constructed from trajectories for the polar vortex flow between September 1 and October 31, 2002. Left: particles at September 1, 2002; middle: September 26, 2002; right: October 31, 2002.

In Figure 7 the average node degree of neighboring nodes $\langle d \rangle_{nn}$ is shown. This network measure highlights a region of strong mixing that appears to delineate the polar vortex. Similar observations have been made using other stretching measures (see e.g. Joseph and Legras (2002); Froyland and Padberg-Gehle (2012)). In Figure 8, the second eigenvector of the generalized

[1] http://data.ecmwf.int/data/index.html

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

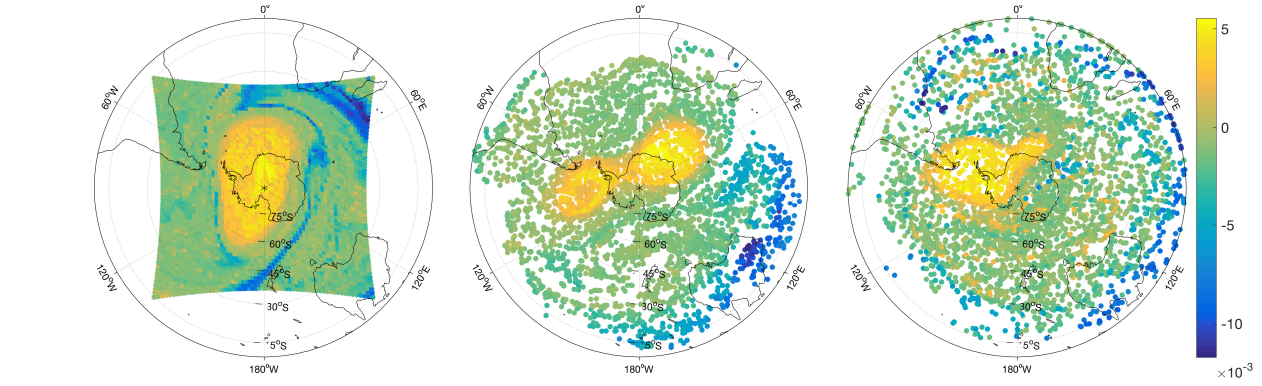

Figure 8. Eigenvector v_2 of generalized graph Laplacian eigenvalue problem for the polar vortex example. Left: particles at September 1, 2002; middle: September 26, 2002; right: October 31, 2002.

graph Laplacian eigenvalue problem (7) is shown. It clearly highlights the polar vortex as a coherent set, confirming the transfer operator based results obtained by Froyland et al. (2010) for a different data set (September 1-14, 2008).

Towards the end of September 2002, the polar vortex splits into two vortices. This is also captured by our network analysis - both in the average node degree of neighboring nodes (Figure 7, middle) and the eigenvector study (Figure 8, middle). The right panels in both figures indicate that one of the two vortices has become unstable and disperses whereas the other vortex increases again by the end of the computation (October 31, 2002).

5 Discussion and conclusion

We have proposed a very simple and inexpensive approach for analyzing coherent behavior and thus transport and mixing phenomena in flows. It is based on a network in which Lagrangian particle trajectories form the nodes. A link is established between two nodes if the respective trajectories come close to each other at least once in the course of time. The resulting network is unweighted and undirected and can be represented by a binary adjacency matrix. Classical local network measures such as node degree and clustering coefficient highlight regions of strong mixing and regular motion, respectively. In particular, a generalized graph Laplacian eigenvalue problem can be used to efficiently and robustly extract coherent sets, even for the case of sparse data. While in this manuscript we have only demonstrated our approach in two-dimensional systems, the extension to three-dimensional flows is straight-forward. In addition, although not illustrated here, our method can easily deal with incomplete trajectory data as only one-time encounters of trajectories are required for setting up the network.

There are some apparent relations to other recently proposed methodologies such as the dynamic isoperimetry framework introduced by Froyland (2015), where a dynamic Laplacian and its spectrum play a central role. The graph Laplacian matrix studied in the present paper appears to be a very coarse but inexpensive and robust approximation of this operator and in a similar way it approximates the diffusion maps used in Banisch and Koltai (2016). There is also an obvious link of the node degree d of our network construction to the trajectory encounter number recently proposed by Rypina and Pratt (2016). A

mathematical analysis of the commonalities and differences between these approaches and our novel network approach is subject to future research.

Acknowledgements. This work is supported by the Priority Programme SPP 1881 Turbulent Superstructures of the Deutsche Forschungsgemeinschaft (PA 1972/3-1). KPG also acknowledges funding from EU Marie-Skłodowska-Curie ITN Critical Transitions in Complex Systems (H2020-MSCA-2014-ITN 643073 CRITICS). We thank Naratip Santitissadeekorn for sharing code for the Antartic polar vortex computations. Publication is supported by Office of Naval Research under Grant No. N00014-16-1-2492 .

Nonlin. Processes Geophys. Discuss., doi:10.5194/npg-2017-4, 2017
Manuscript under review for journal Nonlin. Processes Geophys.
Discussion started: 17 February 2017

References

Allshouse, M. R. and Peacock, T.: Lagrangian based methods for coherent structure detection, Chaos, 25, 097617, 2015.

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

Banisch, R. and Koltai, P.: Understanding the geometry of transport: diffusion maps for Lagrangian trajectory data unravel coherent sets, https://arxiv.org/abs/1603.04709, 2016.

Budišić, M. and Mezić, I.: Geometry of the ergodic quotient reveals coherent structures in flows, Physica D: Nonlinear Phenomena, 241, 1255 – 1269, 2012.

Dellnitz, M. and Preis, R.: Congestion and Almost Invariant Sets in Dynamical Systems, in: Symbolic and Numerical Scientific Computation (Proceedings of SNSC'01), edited by Winkler, F., LNCS 2630, pp. 183–209, Springer, 2003.

Dellnitz, M., Junge, O., Koon, W., Lekien, F., Lo, M., Marsden, J., Padberg, K., Preis, R., Ross, S., and Thiere, B.: Transport in Dynamical Astronomy and Multibody Problems, International Journal of Bifurcation and Chaos, 15, 699–727, 2005.

Donner, R. V., Zou, Y., Donges, J. F., Marwan, N., and Kurths, J.: Recurrence networks—a novel paradigm for nonlinear time series analysis, New Journal of Physics, 12, 033 025, 2010.

Fiedler, M.: Algebraic connectivity of graphs, Czechoslovak Mathematics, 23, 298–305, 1973.

Froyland, G.: Dynamic isoperimetry and the geometry of Lagrangian coherent structures., Nonlinearity, 28, 3587–3622, 2015.

Froyland, G. and Junge, O.: On fast computation of finite-time coherent sets using radial basis functions, Chaos: An Interdisciplinary Journal of Nonlinear Science, 25, 087 409, 2015.

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

Froyland, G. and Padberg-Gehle, K.: Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion, in: Ergodic Theory, Open Dynamics, and Coherent Structures, edited by Bahsoun, W., Bose, C., and Froyland, G., vol. 70 of *Proceedings in Mathematics and Statistics*, chap. 9, pp. 171–216, Springer, 2014.

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

Froyland, G., Santitissadeekorn, N., and Monahan, A.: Transport in time-dependent dynamical systems: Finite-time coherent sets, Chaos, 20, 043 116, 2010.

Hadjighasem, A., Karrasch, D., Teramoto, H., and Haller, G.: Spectral-clustering approach to Lagrangian vortex detection, Phys. Rev. E, 93, 063 107, 2016.

Haller, G.: Lagrangian Coherent Structures, Annual Review of Fluid Mechanics, 47, 137–162, 2015.

Joseph, B. and Legras, B.: Relation between Kinematic Boundaries, Stirring, and Barriers for the Antarctic Polar Vortex, Journal of the Atmospheric Sciences, 59, 1198–1212, 2002.

Lindner, M. and Donner, R.: Spatio-temporal organization of dynamics in a two-dimensional periodically driven vortex flow - a Lagrangian flow network perspective, https://arxiv.org/abs/1603.09704, 2016.

Ma, T. and Bollt, E. M.: Differential Geometry Perspective of Shape Coherence and Curvature Evolution by Finite-Time Nonhyperbolic Splitting, SIAM Journal on Applied Dynamical Systems, 13, 1106–1136, 2014.

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

Newman, M. E. J.: The Structure and Function of Complex Networks, SIAM Review, 45, 167–256, 2003.

Padberg, K., Thiere, B., Preis, R., and Dellnitz, M.: Local expansion concepts for detecting transport barriers in dynamical systems, Communications in Nonlinear Science and Numerical Simulation, 14, 4176 – 4190, 2009.

Rypina, I., Brown, M., Beron-Vera, F., Koçak, H., Olascoaga, M., and Udovydchenkov, I.: On the Lagrangian dynamics of atmospheric zonal jets and the permeability of the stratospheric polar vortex, J. Atmos. Sci., 64, 3595, 2007.

Rypina, I. I. and Pratt, L. J.: Trajectory encounter number as a diagnostic of mixing potential in fluid flows, Nonlinear Processes in Geophysics Discussions, 2016, 1–23, 2016.

Schlueter-Kuck, K. and Dabiri, J.: Coherent structure coloring: identification of coherent structures from sparse data using graph theory, https://arxiv.org/abs/1610.00197, 2016.

Ser-Giacomi, E., Rossi, V., López, C., and Hernández-García, E.: Flow networks: A characterization of geophysical fluid transport, Chaos, 25, 036404, 2015.

Shi, J. and Malik, J.: Normalized Cuts and Image Segmentation, IEEE Trans. Pattern Anal. Mach. Intell., 22, 888–905, 2000.

Williams, M. O., Rypina, I. I., and Rowley, C. W.: Identifying finite-time coherent sets from limited quantities of Lagrangian data, Chaos, 25, 087408, 2015.