

Lagrangian coherent sets in turbulent Rayleigh-Bénard convection

Christiane Schneide ¹, Martin Stahn ^{2,*}, Ambrish Pandey,³ Oliver Junge ⁴, Péter Koltai ²,
Kathrin Padberg-Gehle ¹ and Jörg Schumacher ^{3,5}

¹*Institut für Mathematik und ihre Didaktik, Leuphana Universität Lüneburg, D-21335 Lüneburg, Germany*

²*Institut für Mathematik, Freie Universität Berlin, D-14195 Berlin, Germany*

³*Institut für Thermo- und Fluidodynamik, Technische Universität Ilmenau, D-98684 Ilmenau, Germany*

⁴*Zentrum Mathematik, Technische Universität München, D-85748 Garching, Germany*

⁵*Tandon School of Engineering, New York University, New York, New York 11201, USA*



(Received 9 July 2019; revised manuscript received 20 September 2019; published 11 November 2019)

Coherent circulation rolls and their relevance for the turbulent heat transfer in a two-dimensional Rayleigh-Bénard convection model are analyzed. The flow is in a closed cell of aspect ratio four at a Rayleigh number $Ra = 10^6$ and at a Prandtl number $Pr = 10$. Three different Lagrangian analysis techniques based on graph Laplacians (distance spectral trajectory clustering, time-averaged diffusion maps, and finite-element based dynamic Laplacian discretization) are used to monitor the turbulent fields along trajectories of massless Lagrangian particles in the evolving turbulent convection flow. The three methods are compared to each other and the obtained coherent sets are related to results from an analysis in the Eulerian frame of reference. We show that the results of these methods agree with each other and that Lagrangian and Eulerian coherent sets form basically a disjoint union of the flow domain. Additionally, a windowed time averaging of variable interval length is performed to study the degree of coherence as a function of this additional coarse graining which removes small-scale fluctuations that cause trajectories to disperse quickly. Finally, the coherent set framework is extended to study heat transport.

DOI: [10.1103/PhysRevE.100.053103](https://doi.org/10.1103/PhysRevE.100.053103)

I. INTRODUCTION

Thermal turbulent convection acts as one essential driving mechanism in many turbulent flows in nature spanning a wide range of examples starting from stellar interiors [1] via planetary cores [2] to atmospheric motion [3] and transport dynamics in lakes and ponds [4]. An idealized model of thermal convection is Rayleigh-Bénard convection (RBC), in which a fluid layer placed between two solid horizontal plates is uniformly heated from below and cooled from above [5]. This particular setting contains already many of the properties which can be observed in natural flows. One is the formation of large-scale coherent patterns when RBC is investigated in horizontally extended domains [6–11]. These coherent sets, which have been detected in the Eulerian frame of reference, are termed turbulent superstructures as the characteristic horizontal scale extends the height of the convection layer. In thermal convection flows, they consist of convection rolls and cells that may, however, be concealed in instantaneous velocity fields by turbulent fluctuations. Large-scale circulations in Rayleigh-Bénard convection exist also in smaller domains or cells and have been analyzed, for example, by proper orthogonal decomposition (POD) [12–14] (see also the Appendix of Verma [15] for POD in RBC). In the Eulerian frame, large-scale patterns show up prominently either after time averaging or as the primary POD modes, for both tem-

perature and velocity fields. This is illustrated in Fig. 1 where two coherent circulation rolls are present with narrow regions of upwelling hot and downwelling cold fluid.

At the core of the data-driven analysis in the Lagrangian frame of reference is the concept of a *coherent set* [16–19], a region in the fluid volume that only weakly mixes with its surrounding and which often stays regularly shaped (non-filamented) under the evolution by the flow. Such regions can be determined in two ways, by set-oriented [20,21] or manifold-based methods (see [18,22] for recent reviews of both concepts). The manifold-based approach comprises Lagrangian coherent structures (LCS), i.e., minimal curves in two dimensions and surfaces in the three-dimensional case that enclose coherent sets [23]. This framework was extended recently to include weak diffusion across the manifold [24].

Coherent sets were originally introduced based on transfer operators [16,17]. These are linear operators that evolve densities under the action of the flow. Coherent sets can be identified from the leading singular functions of this operator. More recently, in Ref. [25] they have been characterized as sets which possess a minimal boundary-to-volume ratio for the entire flow duration. Different approaches have been introduced recently that make use of spatiotemporal clustering algorithms applied to Lagrangian trajectory data [26–31]. These algorithms aim at identifying coherent sets as groups of trajectories that remain close to each other in the time interval under investigation.

In this work, we will focus on the latter of these Lagrangian approaches. The three methods that we are going to apply

*Corresponding author: plonka@mi.fu-berlin.de

characterize coherence via regularized linear operators that are directly approximated on the basis of the Lagrangian trajectory data in the convection flow. These are the (i) transfer operator which is regularized by a diffusion kernel [26], the (ii) graph Laplacian operator that characterizes a network of Lagrangian trajectories in the flow [27], and the (iii) dynamic Laplacian which characterizes sets with minimal averaged boundary-to-volume ratio [25,32] or, equivalently, almost invariant sets of a time-dependent heat flow [19]. In all cases, gaps in the discrete eigenvalue spectrum of the operator under consideration give an indication of the number of coherent sets. As will be seen, all three Lagrangian methods detect the same core regions of the large-scale circulation rolls as the coherent sets in which fluid particles remain together for the longest time. With progressing time these coherent sets get increasingly smaller in their spatial extent as expected for turbulent flows.

A second aspect of this work is therefore to extend these Lagrangian concepts and to perform the analysis on data which are averaged in time over a window of variable length. Similar to the Eulerian studies which were mentioned at the beginning of this introduction, we want to investigate coherence as a function of this additional coarse graining which removes small-scale fluctuations in the flow that typically cause a fast separation of Lagrangian trajectories that are initially close together.

A final aspect of this work is to adapt the presented analysis directly to the transfer of heat. The presented Lagrangian methods can then be used to investigate heat coherence in RBC.

Here, we study RBC in a two-dimensional closed box of aspect ratio four. Note that the large- and small-scale quantities show similar scalings in two- and three-dimensional RBC [34–36] for large Prandtl numbers. Therefore, very-long-time temporal evolution of the convective flow configurations has been studied in two-dimensional settings [37,38]. The objective of this work is to take such a simple two-dimensional turbulent flow at a moderate Rayleigh number and to demonstrate and compare the Lagrangian concepts and ideas.

In Sec. II we introduce the numerical model and the data set. Section III gives a further motivation for coherence. The Lagrangian methods which we will compare are introduced in Sec. IV and applied to the convection data in Sec. V. Heat coherence is discussed in Sec. VI and we conclude in Sec. VII.

II. RAYLEIGH-BÉNARD CONVECTION FLOW

Conservation of mass, momentum, and internal energy lead to equations which govern the dynamics of RBC. In the Boussinesq approximation [5] they are given in a nondimensional form by

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \theta \mathbf{e}_z + \sqrt{\frac{\text{Pr}}{\text{Ra}}} \nabla^2 \mathbf{u}, \quad (1)$$

$$\frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta = u_z + \frac{1}{\sqrt{\text{PrRa}}} \nabla^2 \theta, \quad (2)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (3)$$

where $\mathbf{u} = (u_x, u_z)$, θ , and p are the velocity, temperature deviation, and pressure fluctuation fields, respectively. Note

that the temperature fluctuation from the linear conductive profile is related to the total temperature field T as

$$T(x, z, t) = T_{\text{bottom}} - \frac{\Delta T}{H} z + \theta(x, z, t), \quad (4)$$

where T_{bottom} is the temperature at the bottom plate. Equation (4) is given here in physical units. Equations (1)–(3) were nondimensionalized using the height of the simulation domain H of as the characteristic length scale, the free-fall velocity $u_f = \sqrt{\alpha g \Delta T H}$ as the characteristic velocity, and the temperature difference ΔT between the top and bottom plates as the characteristic temperature. The main governing parameters of RBC are the Rayleigh number Ra and the Prandtl number Pr . The Rayleigh number signifies the strength of thermal driving force compared to dissipative forces, and the Prandtl number is the ratio of the kinematic viscosity and thermal diffusivity of the fluid. They are defined as

$$\text{Ra} = \frac{\alpha g \Delta T H^3}{\nu \kappa}, \quad (5)$$

$$\text{Pr} = \frac{\nu}{\kappa}, \quad (6)$$

where α , ν , κ are the thermal expansion coefficient, the kinematic viscosity, and the thermal diffusivity of the fluid, respectively. The acceleration due to gravity g points downward.

We assume that the fields $u_x, u_z, \theta \in H^1(\Omega \times [0, \tau])$ with $\Omega = [0, L_x] \times [0, H]$ and total integration time τ . Here, H^1 is the Sobolev space of square integrable functions with square integrable derivatives. Equations (1)–(3) are solved using a pseudospectral solver TARANG [39] in a two-dimensional box of aspect ratio $\Gamma = L_x/H = 4$. Stress-free (or free-slip) boundary conditions for the velocity field are employed at all the walls. This implies that the corresponding normal velocity component and the normal derivative of the tangential velocity component vanish to zero, respectively. For the temperature field, isothermal (adiabatic) boundary conditions are applied in the vertical (horizontal) direction. To satisfy these boundary conditions, the temperature fluctuation and velocity components are expanded in sine and cosine basis functions. This results to

$$u_x(x, z, t) = \sum_{k_x, k_z} 4\hat{u}_x(k_x, k_z, t) \sin(k_x x) \cos(k_z z), \quad (7)$$

$$u_z(x, z, t) = \sum_{k_x, k_z} 4\hat{u}_z(k_x, k_z, t) \cos(k_x x) \sin(k_z z), \quad (8)$$

$$\theta(x, z, t) = \sum_{k_x, k_z} 4\hat{\theta}(k_x, k_z, t) \cos(k_x x) \sin(k_z z), \quad (9)$$

where $\mathbf{k} = (k_x, k_z)$ is the wave vector [38] with $k_x = m\pi/L_x$ and $k_z = n\pi/H$; m, n being integers. We perform direct numerical simulation for $\text{Pr} = 10$, $\text{Ra} = 10^6$, and $\Gamma = 4$ using 513×129 uniformly spaced grid points. The time advancement is done using fourth-order Runge-Kutta method (RK4), and the fields are dealiased using the 2/3 rule. We refer to [38–40] for more details on the numerical simulations.

The presented analyses require Lagrangian particle tracks which are evaluated together with the turbulent flow. Each individual tracer particle i is advected in the velocity field corresponding to

$$\frac{dx_i}{dt} = \mathbf{u}(x_i, t). \quad (10)$$

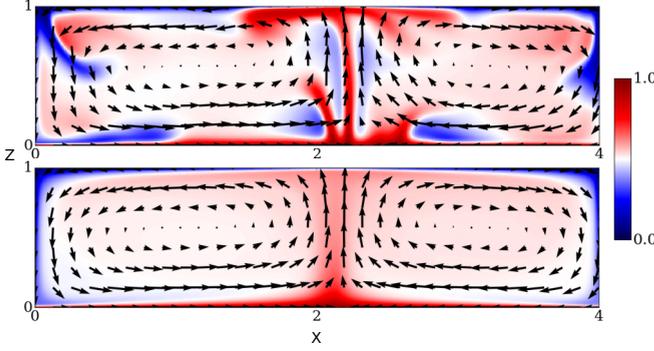


FIG. 1. Two-dimensional RBC benchmark flow with Rayleigh number $Ra = 10^6$ and Prandtl number $Pr = 10$. Top: instantaneous flow configuration exhibiting a pair of counter-rotating circulation rolls. Temperature contours are shown as colors and the velocity field is indicated by arrows. Bottom: time-averaged flow configuration. Averaging is performed for the total duration of the time integration which is 500 free-fall time units.

We simulate $i = 1 \dots N$ particle trajectories with $N = 5000$. Time integration is done again by the RK4 method. The interpolation of the velocity field on the particle position applies cubic splines.

We start our simulation with random noise for velocity and temperature fields, \mathbf{u} and θ , as the initial condition and continue until a statistically steady state after 2000 free-fall times t_f is reached. Here, $t_f = H/u_f$. The time-averaged flow structure exhibits a pair of counter-rotating circulation rolls as shown in the bottom panel of Fig. 1. Hot fluid rises in the central region whereas cold fluid falls near the sidewalls. The velocity and temperature fields at all the grid points are written out every $0.1 t_f$.

III. LAGRANGIAN COHERENCE IN A TURBULENT FLOW

In the dynamical systems perspective, we consider the turbulent convection flow as a mapping in the state or phase space $\Omega \subset \mathbb{R}^2$. Let

$$\Phi_{t_0}^t : \mathbf{x}(t_0) \mapsto \mathbf{x}(t) = \Phi_{t_0}^t(\mathbf{x}(t_0), t_0) \quad (11)$$

denote the flow map, which takes fluid particles from their initial location $\mathbf{x}_0 = \mathbf{x}(t_0)$ at time t_0 to their spatial location $\mathbf{x}(t)$ at time t in correspondence with the velocity field \mathbf{u} of the RBC flow. This mapping is given by the differential equations (10) such that $\Phi_{t_0}^t(\mathbf{x}(t_0), t_0)$ solves the corresponding initial value problem. The advecting flow is simultaneously determined by solving the Boussinesq equations for the RBC flow. The phase space Ω is equipped with a reference measure μ and a sequence of nonsingular time-dependent flow maps $\Phi_{t_0}^{t_0+\Delta t}, \dots, \Phi_{t_0+(n-1)\Delta t}^{t_0+n\Delta t}$ for n time steps Δt . We construct a single flow map via a successive application of flow maps over smaller time steps Δt , namely, as $\Phi = \Phi_{t_0}^{t_0+\Delta t} \circ \dots \circ \Phi_{t_0+(n-1)\Delta t}^{t_0+n\Delta t}$.

An overarching goal is to detect and locate slow mixing dynamical structures. These structures should be macroscopic in size and by “slow mixing” we have in mind a geometric mixing rate that is slower than $1/\Lambda$ where Λ is the largest positive Lyapunov exponent that measures the exponential separation

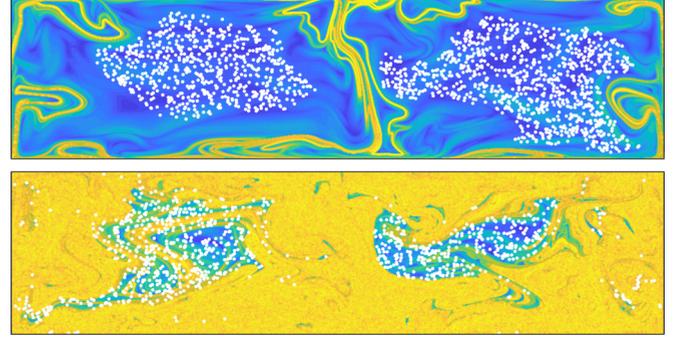


FIG. 2. Finite-time Lyapunov exponent field computed in forward time over $20t_f$ and $200t_f$ in the top and bottom panels, respectively. Large values are indicated by yellow. Superimposed are particles (white dots) at initial positions ($t = 2000t_f$) which belong to two coherent sets as identified by the method described in Sec. IV A (see also Figs. 4 and 7).

of initially infinitesimally distant neighboring particles. Thus, such slow mixing cannot be explained by local stretching, but is instead due to the way in which the dynamics acts globally. Following [25], we wish to partition the state space $\Omega = \mathbb{A} \cup \mathbb{A}^C$ into a disjoint union of \mathbb{A} and its complement \mathbb{A}^C at initial time and $\Omega = \mathbb{B} \cup \mathbb{B}^C$ at final time such that we optimize the coherence ratio which can be thought of as

$$\rho(\mathbb{A}, \mathbb{B}) = \frac{\mu(\mathbb{A} \cap \Phi^{-1}(\mathbb{B}))}{\mu(\mathbb{A})} + \frac{\mu(\mathbb{A}^C \cap \Phi^{-1}(\mathbb{B}^C))}{\mu(\mathbb{A}^C)}, \quad (12)$$

and that quantifies how much of \mathbb{A} ends up in \mathbb{B} after applying the flow map Φ (augmented by a small random perturbation, to be precise) with respect to the reference volume of \mathbb{A} and similarly for \mathbb{A}^C and \mathbb{B}^C . In a nutshell, Eq. (12) quantifies in our example the area content of the subset that stays connected from a Lagrangian point of view.

We seek for an optimal 2-partition into slow mixing (or coherent) sets and the remainder which becomes rapidly well mixed under the action of convective turbulence. Coherent means now that these sets are almost invariant under the combined forward-backward dynamics of the flow map Φ (including a small amount of diffusion for regularization). The exact version of (12) and the connection to the transfer operator \mathcal{P} and its singular values or functions can be found in [17].

Central to manifold-based concepts of coherence are LCS [23]. These are material surfaces that extremize a certain stretching or shearing quantity, such as measured by finite-time Lyapunov exponents (FTLEs) [41]. Ridges in the FTLE field from a forward time computation highlight repelling LCS. In Fig. 2 we show the forward time FTLE fields for the two different time windows considered in Sec. VC. While for the short time interval (top) isolated ridges can be observed, this is no longer the case for the long time interval (bottom). This makes it difficult to unambiguously define coherent sets from the FTLE ridges (see also [28] for related observations) and, therefore, in this work, we focus on set-based approaches. In Fig. 2 we also show the results of a set-based computation (in this case using the minimum distance spectral trajectory clustering method introduced in

Sec. IV A), which identifies coherent sets as the core regions of the large-scale circulation rolls that appear to be characterized by low FTLE values. We come back to this point in Sec. V C.

IV. GRAPH LAPLACIAN-BASED COHERENT STRUCTURE DETECTION

All algorithms to be introduced work with the following set of Lagrangian data. Let the positions of N particles \mathbf{x}_i at $\mathcal{T} + 1$ time instances t_k be given, i.e., the trajectory data consist of

$$\{\mathbf{x}_i(t_k) \in \mathbb{R}^d \mid i \in \{1, \dots, N\}, k \in \{0, \dots, \mathcal{T}\}\} \subset \mathbb{R}^d, \quad (13)$$

where $\mathbf{x}_i(t_{k+1}) = \Phi_{t_k}^{t_{k+1}} \mathbf{x}_i(t_k)$, for $i \in \{1, \dots, N\}$ and $k = 0, \dots, \mathcal{T} - 1$. Trajectories in a coherent set will stay close to each other over a long time in contrast to trajectories not belonging to this specific coherent set. Diverging trajectories indicate filamentation of a set which has a large diffusive outward transport with respect to the dynamics. Based on this notion, the algorithms evaluate the ε neighborhood of the trajectories using distinct models of the diffusion process. If possible, this is followed by the construction of a rate matrix Q , which will be introduced in the following sections. The solution of an eigenvalue problem yields eigenvalues λ_ℓ which satisfy $0 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. The eigenvectors corresponding to the dominant eigenvalues (i.e., close to zero) are used to cluster the trajectory data into coherent sets.

A. Distance spectral trajectory clustering

The idea of the network-based analysis, published in [27], is to interpret each Lagrangian trajectory $\{\mathbf{x}_i(t_k)\}_k$ as node of a network consisting of N nodes. A link between two nodes $\{\mathbf{x}_i(t_k)\}_k$ and $\{\mathbf{x}_j(t_k)\}_k$ is created if and only if the minimum distance of the two trajectories for at least one time instance t_k is smaller than a prespecified cutoff radius ε . Thereby, network measures as, e.g., the node degree or local clustering coefficient, can be used to distinguish coherent sets from incoherent flow [42–44]. In order to partition the network into independent sets, we use a version of the normalized cut approach [45], which is based on spectral graph theory. This approach aims at maximizing the intracluster connectivity and simultaneously minimizing the intercluster connectivity. An approximate solution of the normalized cut problem can be achieved solving the generalized eigenvalue problem [45]

$$Ly = \lambda Dy, \quad (14)$$

where $L = D - A$ is the non-normalized graph Laplacian, D is the degree matrix, and A is the adjacency matrix. Here, we vary the approach in the way that we use $L' = A - D$ and, for the sake of direct comparability with the second method, we

Algorithm 1 Minimum distance spectral trajectory clustering.

1: Define adjacency matrix

$$A_{ij} = \begin{cases} 1, & \|\mathbf{x}_i(t_k) - \mathbf{x}_j(t_k)\| \leq \varepsilon \\ & \text{for at least one } k \in \{0, \dots, \mathcal{T}\} \text{ and } i \neq j \\ 0, & \text{otherwise.} \end{cases}$$

2: Define degree matrix

$$D_{ij} = \begin{cases} \sum_{j'} A_{ij'}, & i = j \\ 0, & \text{otherwise.} \end{cases}$$

3: Define normalized Laplacian $Q^{\text{nw}} = D^{-1}(A - D)$.

4: Partition trajectories based on dominant eigenvectors of Q^{nw} .

state (14) in the equivalent form¹

$$Q^{\text{nw}}y = \lambda y, \quad \text{where } Q^{\text{nw}} = D^{-1}L'. \quad (15)$$

Note that the eigenvalues of (14) and (15) differ only in their signs. The multiplicity m of the first eigenvalue $\lambda_1 = 0$ equals the number of connected components in the network. We construct the network such that the zero eigenvalue is simple. Then, the number of eigenvalues $\lambda_2, \dots, \lambda_m$ close to 0 determines the number of weakly linked subgraphs. The eigenvectors corresponding to these eigenvalues are used to extract m clusters. In this paper, this postprocessing is done by the k -means algorithm [46]. The corresponding pseudocode is given in Algorithm 1.

We note that the adjacency matrix in Algorithm 1 could be constructed as a weighted matrix instead of the current binary one. One possibility is to measure the average distance between two trajectories

$$\bar{d}(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\mathcal{T} + 1} \sum_{k=0}^{\mathcal{T}} \|\mathbf{x}_i(t_k) - \mathbf{x}_j(t_k)\|, \quad (16)$$

and set $A_{ij} = \bar{d}(\mathbf{x}_i, \mathbf{x}_j)^{-1}$, $i \neq j$, as was suggested in [31]. In general, relaxing the binary structure of A by introducing weights is a refinement of the dynamical information contained in A , and is used in the method discussed next.

B. Time-averaged diffusion maps

The theory of diffusion maps introduced by Coifman and Lafon in [47] has been successfully applied to a variety of nonlinear dimensionality reduction problems. In Ref. [26] the framework of diffusion maps is used to analyze transport in dynamical systems and to find coherent sets solely based on possibly sparse or incomplete Lagrangian trajectory data. The idea is to introduce a diffusion process on the data points and

¹Assuming D has no zeros on the diagonal, i.e., every trajectory “meets” another at least once. If not, that is there is some i with $D_{ii} = 0$, we can either delete the trajectory from the set completely, or if we want to keep it as an independent entity not connected to any other, we can set the i th row of Q^{nw} equal to the i th row of the identity matrix. The latter step produces a cluster containing only the i th trajectory.

detect points that can be reached more easily from one another by the diffusion process. This is done via the eigenvectors of this diffusion process (or operator) which provide intrinsic coordinates of the data set.

The time-averaged diffusion map algorithm from [26] (called “space-time diffusion map” therein) proceeds as follows. Given the trajectory data as in (13), the algorithm looks for tight bundles of trajectories. To achieve this, a diffusion map matrix P_k is constructed at every time instance t_k by row normalizing a similarity matrix K_{t_k} . This matrix is constructed using a rotation invariant kernel which is given by

$$k_\delta(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{\delta}\right) \mathbb{1}_{[0,r]}, \quad (17)$$

with the characteristic function $\mathbb{1}_{[0,r]}$ introducing a cutoff at some radius r . Hence, the similarity matrix $(K_{t_k})_{ij} = k_\delta(\mathbf{x}_i(t_k), \mathbf{x}_j(t_k))$ is only dependent on the Euclidean distances $\|\mathbf{x}_i(t_k) - \mathbf{x}_j(t_k)\|$ in \mathbb{R}^2 between the point pairs $\mathbf{x}_i(t_k), \mathbf{x}_j(t_k)$. The parameter δ can be seen as the strength (or duration) of the diffusion, and characterizes what will be considered close; hence, its square root is a length scale of spatial resolution. In practice, δ together with the kernel function will determine a cutoff radius r , beyond which the similarity is negligibly small (here $r = \sqrt{2\delta}$). This is similar to ε in Sec. IV A, and allows for efficient numerical computation. Now, P_k corresponds to a random walk (Markov chain) constructed on all existing data points (or particle positions) at t_k with two points getting a higher jump probability the closer they are. Using time averaging,

$$P^{\text{dm}} = \frac{1}{\mathcal{T} + 1} \sum_{k=0}^{\mathcal{T}} P_{t_k}, \quad (18)$$

we obtain the space-time diffusion matrix P^{dm} which describes a Markov chain on all trajectories. We note that averaging the matrices P_{t_k} yields different results to using diffusion maps with the averaged distances $\bar{d}(\mathbf{x}_i, \mathbf{x}_j)$ from above, and has a different interpretation (cf. [26]).

Again, for the reason of comparability, we will work with the rate matrix

$$Q^{\text{dm}} = \frac{1}{\delta}(P^{\text{dm}} - I), \quad (19)$$

where I denotes the identity matrix. By construction, the eigenvalues λ_ℓ of Q^{dm} satisfy $0 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Just as in Sec. IV A, the dominant eigenvectors of Q^{dm} are now used to cluster the trajectory data into coherent sets.

We remark that the algorithm, as introduced here in Algorithm 2, corresponds to a particular choice of scaling in diffusion maps, the so-called $\alpha = 0$ case [26,47]. Originally in Ref. [26] a different scaling was used. We refer to that work for the details, and note only that for uniformly distributed data there is no practical difference. With the correct scaling, it is shown in Ref. [26, Theorem 3 and Eq. (21)] that if the number of trajectories goes to infinity, Q^{dm} converges to the so-called *dynamic Laplacian* [19,25,48] whose eigenvectors characterize Lagrangian coherent sets. This makes the time-averaged diffusion-map method a consistent data-based approach for finding coherent sets.

Algorithm 2 Diffusion-map based analysis of Lagrangian data.

- 1: Define kernel $k_\delta(\mathbf{x}, \mathbf{y}) = \exp(-\|\mathbf{x} - \mathbf{y}\|^2/\delta) \mathbb{1}_{[0,r]}(\|\mathbf{x} - \mathbf{y}\|)$ with cutoff radius r .
 - 2: Define similarity matrices $K_{t_k,ij} = k_\delta(\mathbf{x}_i(t_k), \mathbf{x}_j(t_k))$.
 - 3: Define Markov matrices $P_k = D_{t_k}^{-1} K_{t_k}$, with diagonal degree matrices $D_{t_k,ii} = \sum_j K_{t_k,ij}$.
 - 4: Define time-averaged diffusion-map matrix $P^{\text{dm}} = \frac{1}{\mathcal{T}+1} \sum_{k=0}^{\mathcal{T}} P_{t_k}$, and corresponding time-averaged diffusion-map Laplacian $Q^{\text{dm}} = \frac{1}{\delta}(P^{\text{dm}} - I)$.
 - 5: Partition trajectories based on dominant eigenvectors of P^{dm} (which are the same as those of Q^{dm}).
-

C. Discrete dynamic Laplacian

A third approach to the detection of Lagrangian coherent sets from sparse and possibly incomplete Lagrangian trajectory data has been developed in [32]. It is based on the geometric idea [25] that Lagrangian coherent sets can be characterized by a small boundary-to-volume ratio: Whenever the length of the boundary of some advecting set $\Phi_{t_0}^t(\mathbb{A})$, $\mathbb{A} \subset \Omega$, is small in relation to its area consistently for all times t , diffusive transport of some passive scalar over its boundary (induced by small random perturbations to $\Phi_{t_0}^t$) will be small. Consequently, the coherence ratio (12) of the pair $\mathbb{A}, \Phi_{t_0}^t(\mathbb{A})$ will be large even in the presence of small perturbations to $\Phi_{t_0}^t$.

Equivalently (cf. [19]), we can characterize a Lagrangian coherent set as a material set which is almost invariant under the flow of the Lagrangian diffusion equation [49] for some diffusive scalar quantity S . In Lagrangian coordinates, this equation is given by [19]

$$\frac{\partial S}{\partial t} = \tilde{\kappa}_L \Delta^t S, \quad (20)$$

with $\Delta^t S = \nabla \cdot (\mathcal{D}(\mathbf{x}, t) \nabla S)$ and the dimensionless $\tilde{\kappa}_L \ll 1$. Here, the advection by the flow map that deforms a material set has been encoded in a Lagrangian eddy diffusivity (i.e., the inverse Cauchy-Green strain tensor) which is given by

$$\mathcal{D}(\mathbf{x}, t) = D\Phi_{t_0}^t(\mathbf{x})^{-1} D\Phi_{t_0}^t(\mathbf{x})^{-\top}, \quad (21)$$

where $D\Phi_{t_0}^t$ is the Jacobian of the flow map.

In order to compute coherent sets via this approach, we first need to remove the time dependence from the diffusion operator Δ^t . This can be achieved by time averaging, i.e., by considering the operator

$$\bar{\Delta} = \frac{1}{\mathcal{T} + 1} \sum_{k=0}^{\mathcal{T}} \Delta^{t_k}, \quad (22)$$

called the *dynamic Laplacian* [25].

In a second step, one considers the eigenproblem $\bar{\Delta} v = \lambda v$ with appropriate boundary conditions (here, we have used Neumann boundary conditions). Lagrangian coherent sets are then given by sublevel sets of the eigenvectors at the leading

Algorithm 3 Dynamic Laplacian based analysis of Lagrangian data.

-
-
- 1: Choose a finite element (e.g., the piecewise linear triangular element).
 - 2: **for** $k = 0, \dots, \mathcal{T}$ **do**
 - 3: Construct a mesh of $\{\mathbf{x}_i(t_k), i = 1, \dots, N\}$.
 - 4: Compute K^k **end do**.
 - 5: Compute $\bar{K} = \frac{1}{\mathcal{T}+1} \sum_{k=0}^{\mathcal{T}} K^k$ and M .
 - 6: Partition trajectories based on the dominant eigenvectors of (\bar{K}, M) .
-
-

eigenvalues of $\bar{\Delta}$, i.e., those closest to 0 (cf. [20,25,50] for more details).

The dynamic Laplacian $\bar{\Delta}$ in this eigenproblem can be discretized by standard finite-element methods (FEM), leading to the generalized eigenproblem $\bar{K}v = \lambda Mv$ with eigenvalues $0 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. The stiffness matrix \bar{K} is the average $\bar{K} = \frac{1}{\mathcal{T}+1} \sum_{k=0}^{\mathcal{T}} K^k$ of the stiffness matrices at each time t_k :

$$K_{ij}^k = - \int_{\Omega} \langle \nabla \phi_i^k, \nabla \phi_j^k \rangle dx. \quad (23)$$

Here, the functions ϕ_i^k are the finite-element basis functions on a triangulation of the data points $\mathbf{x}_i(t_k)$. For the mass matrix M , it suffices to compute

$$M_{ij} = \int_{\Omega} \phi_i^0 \phi_j^0 dx \quad (24)$$

only at the initial time t_0 .

A typical finite element in step 1 of Algorithm 3 is the linear triangular Lagrange element, i.e., the mesh in step 3 consists of triangles and the basis functions are the piecewise linear nodal basis functions. This computational mesh can be constructed as the Delaunay triangulation of the data points $\{\mathbf{x}_i(t_k), i = 1, \dots, N\}$ at each time step. Whenever the data set has a very irregular hull, an α shape will be more appropriate. Typically, α is then chosen minimal such that the triangulation is connected (cf. `alphashape` in `MATLAB`). We refer to Ref. [32] and to the documentation of the packages `FEMDL` (`MATLAB`) and `CoherentStructures.jl` (`JULIA`), which are available from `GITHUB`, for more details and examples.

The triangulation induces a network of trajectories for each time step. There is an interpretation of this discretization of the dynamic Laplacian from a graph Laplacian perspective, detailed in the Appendix, Sec. A 1.

Finally, we want to state that there is no direct link to some rate matrix Q as in the other two approaches since the inverse of M will be in general not sparse.

V. COMPARISON OF DIFFERENT LAGRANGIAN METHODS

In order to compare the methods introduced above, we choose different perspectives. We compare the approaches first on a theoretical level (Sec. V A) and regarding their structure (Sec. V B). Finally, we apply all these three methods to the same trajectory data set introduced in Sec. II and compare their results (Sec. V C).

A. Theoretical level

We start by noting that Q^{nw} and Q^{dm} are so-called *rate matrices*. A rate matrix Q has the properties $\sum_j Q_{ij} = 0$ for all i and $Q_{ij} \geq 0$ for $i \neq j$. It defines a time-continuous Markov chain in the following sense. The process being in state i at current time, (i) will stay in i for a random amount of time τ , where τ is an exponentially distributed random variable with rate $-Q_{ii}$,² and (ii) then will jump to the next state j randomly (and independently of τ) with probability $Q_{ij}/|Q_{ii}|$. Thus, the larger the absolute value of Q_{ii} , the faster the jump occurs on average. The quantity $1/|Q_{ii}|$ is called the (*average/expected*) *holding time* of state i .

We refer to the Appendix, Sec. A 2, for further technical details, and only note here briefly that the jump processes Q^{nw} and Q^{dm} introduce jumps of mean length $\mathcal{O}(\varepsilon)$ and $\mathcal{O}(\sqrt{\delta})$, respectively, and this distance governs the finest scales they can resolve. Coherent sets below these scales can not be detected. The FEM-based discretization of the dynamic Laplacian has no explicit scale parameter, however, the average diameter of the triangulation can implicitly be viewed as such.

B. Structure

The descriptions in Sec. IV suggest that the algorithms construct similar objects and have partially even similar steps. All three algorithms are by their very nature frame independent (as they only use mutual distances of trajectories), fast, and construct sparse matrices that encode the space-time behavior. They require the following three parameters as user input:

(1) A subset $\Theta \subset \{0, \dots, \mathcal{T}\}$ of all time instances at disposal. Not every coherent set might be present for the entire observation range of a fully nonautonomous flow, thus a natural choice is to restrict the time interval of consideration to $\Theta = \{k_0, \dots, k_1\}$, where $0 \leq k_0 < k_1 \leq \mathcal{T}$. Also, if the sampling time step $\max_k |t_k - t_{k-1}|$ is small compared to typical dynamic timescales, the data can be subsampled with respect to time without essentially altering the results.

(2) Proximity parameters δ and ε (not for the FEM approach). These govern the minimal spatial scales on which the methods can detect coherence. In the FEM method this scale will implicitly be defined by the size of the elements resulting from the triangulation. For uniformly distributed Lagrangian particles, the diameter of the resulting triangulation can be interpreted as the implicit closeness parameter.

(3) The number of clusters. Gaps in the spectrum of Q^{dm} , Q^{nw} , or (\bar{K}, M) can indicate natural choices for the number of clusters. Some arguments for this approach are collected in [51].

All methods discussed here are highly suitable for sparse and incomplete data (see [26], in particular Sec. V therein, [27,32]).

²A random variable is exponentially distributed with rate ρ , if the probability that $\tau > t$ for any $t > 0$ is $e^{-\rho t}$. Equivalently, the probability density of τ is given by $f_{\tau}(t) = \rho^{-1} e^{-\rho t}$.

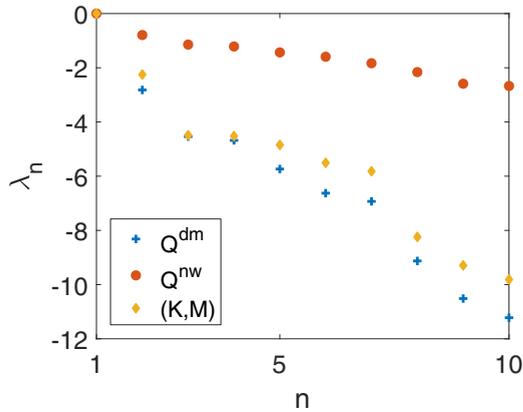


FIG. 3. Eigenvalues of $2\epsilon^{-1}Q^{nw}$ (circles), Q^{dm} (crosses), and (\bar{K}, M) (diamonds). The leading 10 eigenvalues of the matrices of each approach for the short time interval setting are shown.

C. Numerical method

In the following, we discuss how the three methods compare to each other when applied to the turbulent convection flow data introduced in Sec. II for times $t > 2000t_f$. In view of the Lagrangian coherence, we will look at a short time interval of $\mathcal{T} = 200$ steps and a long time interval of $\mathcal{T} = 2000$ steps with $0.1t_f$ per step. We thus start at $t = 2000.0t_f$ and end at $\tau = 2020.0t_f$ and $\tau = 2200.0t_f$, respectively. These time intervals are chosen since the average circulation (or eddy turnover) time of one roll was found to be $t \approx 20t_f$, calculated using the maximum circumference of the roll covering half of the box and the root mean square velocity (see [11] for more details). Thus, the shorter (longer) time interval approximately corresponds to 1 circulation (10 circulations). We initialize $N = 5000$ uniformly distributed particles in order to analyze mass transport. The particles are advected using the snapshot files of the velocity field and for comparability all three algorithms are applied to the same trajectory data set.

At this point, we would like to seize the opportunity to explain some options for the choice of the similarity parameters δ and ϵ . Some general criteria for the choice of δ and ϵ are (i) sparsity of Q^{dm} and Q^{nw} to achieve, for example, 5% nonzero entries, (ii) a stable spectrum which implies that for a series of δ , ϵ the dominant spectrum of Q^{dm} or Q^{nw} does not vary qualitatively, thus conserving relative distances and gaps of eigenvalues, (iii) edge density for the graph- or network-based approaches [52].

Aside from those technical constraints, there are physical reasons that account for the turbulent convection flow under consideration. For Rayleigh-Bénard convection that is considered here, we use the Nusselt number Nu , a dimensionless number that quantifies the global turbulent heat transfer across the plane, and the expected nondimensional radius of a convection roll r_{roll} to obtain bounds for both cutoff scales:

$$\delta_T = \frac{1}{2Nu} < \epsilon < \frac{r_{roll}}{2}. \quad (25)$$

We recall here that $\epsilon = \sqrt{2\delta}$. The lower bound in (25) is determined by the mean thickness of the thermal boundary layer $\delta_T = 0.028$ for the present Nusselt number of

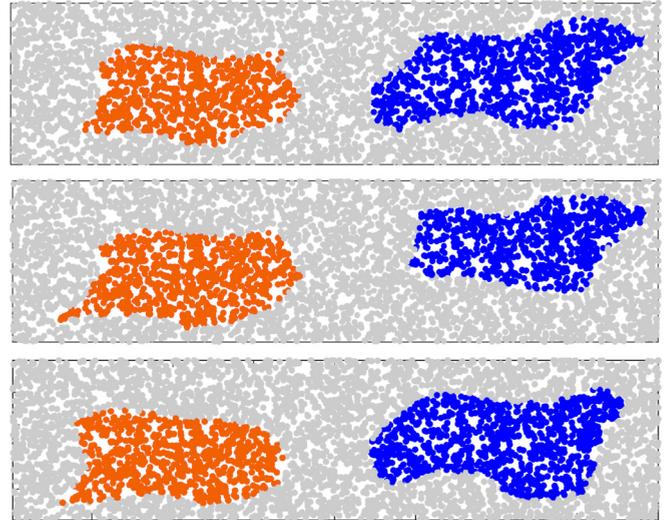


FIG. 4. Clustering with respect to the second and fifth eigenvectors of Q^{nw} (top) and of Q^{dm} (center) and the second and third eigenvectors of (\bar{K}, M) (bottom). Particles in different coherent sets at time $t = 2010t_f$ are colored according to k -means clustering.

$Nu = 17.7$. Recall that all length scales are expressed in units of height H . The viscous boundary layer thickness can be determined for stress-free boundary conditions by a method that has been suggested by Petschel *et al.* [53]. They determined a so-called dissipation layer thickness from the intersection points (close to the bottom and top walls) of the line- and time-averaged profile of the kinetic energy dissipation rate with its plane mean. The analysis for the present flow gives a viscous boundary layer thickness of $\delta_v \approx 0.039$ which is slightly larger than the thermal boundary layer thickness.

The dynamics further restricts the choice of both closeness parameters. In the short time interval, the trajectories cover only one circulation and very few Lagrangian particles will switch from the left roll to the right roll or vice versa (see Fig. 1). Thus, we have to take a small cutoff parameter in order to capture relevant dynamical structures. For the long time interval, the spectra of eigenvalues are typically not stable in case of a small closeness parameter. To satisfy these restrictions for both methods, we get bounds on ϵ for the short and long time intervals of

$$\epsilon \in \begin{cases} [0.03, 0.07] : \tau = 2020t_f, \\ [0.085, 0.25] : \tau = 2200t_f. \end{cases} \quad (26)$$

As we will see further below, our choices satisfy $\delta_T < \delta_v < \epsilon$.

In the following, we want to visualize the differences and similarities of the results of the three algorithms. In general, the coherent sets are expected to be larger for the network-based analysis compared to the time-averaged diffusion-map method. The reason for this behavior is that the network-based approach does not take into account how close trajectories pass by (as long as it is closer than the threshold ϵ) and how long they reside in the vicinity of each other, thus having a less pronounced and hence larger “dynamic neighborhood.” As the strength of the diffusion decreases exponentially with increasing distance between particles diffusion, the transition from coherent structure to incoherent flow is more easily detected

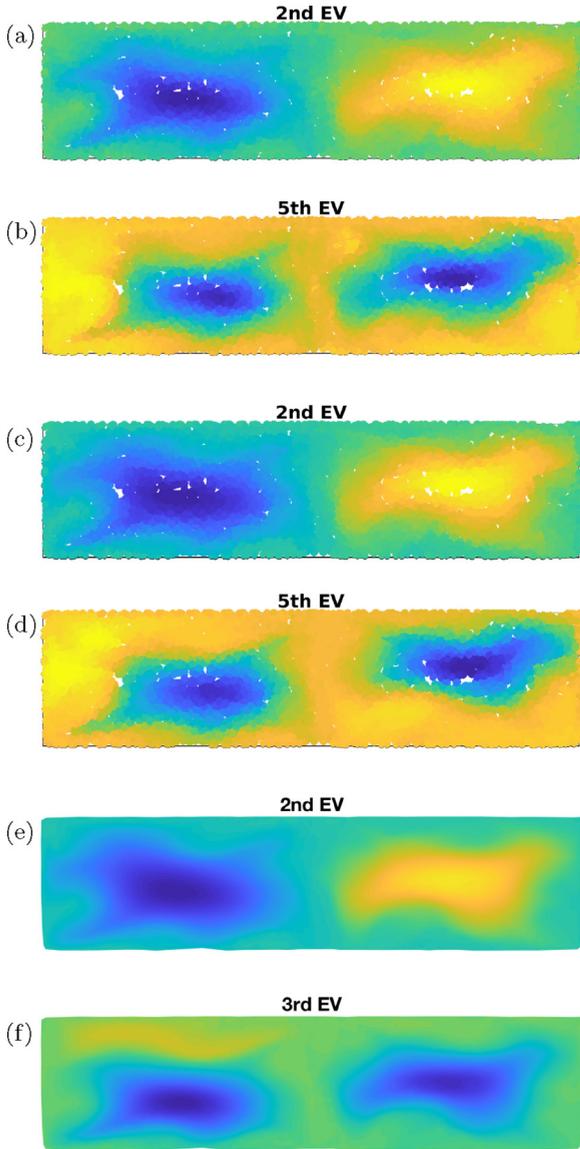


FIG. 5. Particles colored according to eigenvectors of Q^{nw} (a), (b), Q^{dm} (c), (d), and (\bar{K}, M) (e), (f) in the normalized range $[-1, 1]$. Negative (positive) values are indicated by dark (bright) contours. The second and fifth (respectively third) eigenvectors at $t = 2010t_f$.

for diffusion maps. Consequently, the definition of coherence is more strict in the diffusion-map method compared to the network-based method and this will lead to smaller coherent sets.

1. Short time interval up to $\tau = 2020t_f$

For the particular parameter choices $\delta = 0.002$ and $\varepsilon = 0.0632$, we now show the results for the short time span in Figs. 3–5. For visualization purposes we plot the eigenvalues of $2\varepsilon^{-1}Q^{nw}$.

The second eigenvector gives a separation of the left and right circulation rolls. The fifth eigenvector (respectively third for the FEM approach) gives a separation of the gyre cores and the background. We omit the plots for the third and fourth eigenvectors of Q^{nw} and Q^{dm} since they correspond to

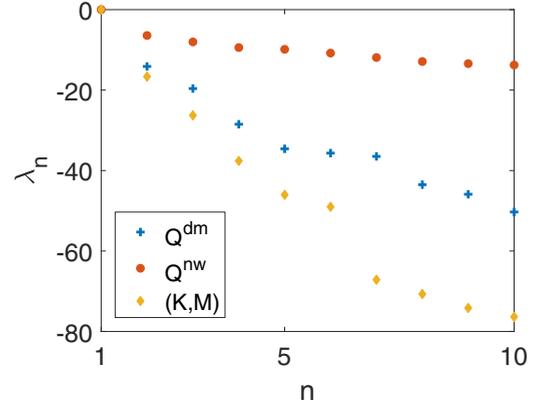


FIG. 6. Eigenvalues of $2\varepsilon^{-1}Q^{nw}$ (circles), Q^{dm} (crosses), and (\bar{K}, M) (diamonds). Leading 10 eigenvalues of the relevant matrices of each approach for the long time interval setting.

subpartitions of the left and the right rolls only. They basically split each roll into halves, i.e., bottom and top or left and right for the third or fourth eigenvector, respectively. This occurrence can be explained in different ways. These (third and fourth) eigenvalues can be considered as higher multiplicities of the second eigenvalue interacting with the numerics. Furthermore, these subpartitions are still valid coherent sets which can be explained as follows. Within the selected time span, the main motion is the rotation. However, even though most particles complete one rotation and the rotational speed varies with radius, this difference is not large enough to effectively separate particles from the inner (core) regions and the outer (background) regions. Therefore, the subpartitions of the circulation rolls are more coherent than the gyre cores. As we will see later for the longer time span, these subpartitions do not occur, implying that they are then less coherent than the gyre cores.

2. Long time interval up to $\tau = 2200t_f$

In the following, we visualize the results for the particular parameter choices of $\delta = 0.005$ and $\varepsilon = 0.1$ in Figures 6–8. Again, for visualization purposes we plot the eigenvalues of $2\varepsilon^{-1}Q^{nw}$. We still get the separation of the left and the right sides from the second eigenvector and the separation of the gyre cores and the background from the third eigenvector. We also note that the spectral gap is more prominent in the eigenvalue spectrum of Q^{dm} compared to Q^{nw} . This could be advantageous in case of an unknown number of coherent sets.

3. Comparison of the three methods

Figure 9 compares the results by highlighting the trajectories that have been assigned differently by the network-based and diffusion-map method. This is done with a symmetrical difference of coherent sets A^{nw} and A^{dm} , given by $(A^{nw} \setminus A^{dm}) \cup (A^{dm} \setminus A^{nw})$. For simplicity, we only plot the symmetric differences for the Q^{nw} and Q^{dm} results. Furthermore, we observe a correlation between the distances between consecutive eigenvalues (see Fig. 6).

In the short time setting the coherent sets resulting from the FEM method are comparable in size and shape to the

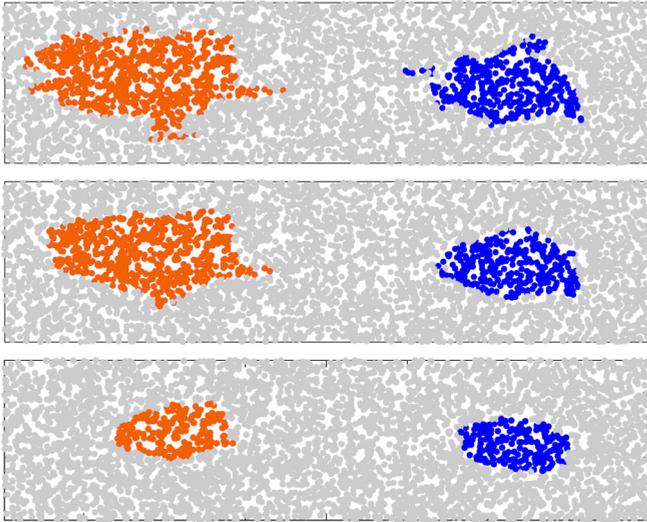


FIG. 7. Clustering with respect to the second eigenvector of Q^{nw} (top), Q^{dm} (center), and (\bar{K}, M) (bottom). Particles in different coherent sets at time $t = 2100t_f$ are colored according to k -means clustering.

ones detected by the other two methods. For the long time interval, however, they are significantly smaller. This might be explained by the original construction of the dynamic Laplacian via dynamic isoperimetry in Ref. [25]: the method detects sets which keep a small boundary-to-volume ratio over the entire time interval of the flow evolution. It is thus a stricter criterion than the ones in the other two approaches.

Figure 10 displays the magnitude of the time-averaged velocity field. Small values can be identified in the elliptic centers of the counter-rotating rolls, the hyperbolic regions in the center of the top and bottom plates where Lagrangian particle pairs separate, and the four corners. The elliptic centers are the regions where the particles stay closely together for the longest time.

4. Lagrangian particle advection in time-averaged flow

In the Eulerian frame of reference, temporal averaging has to be performed in order to reveal the coherent large-scale patterns in the flow clearly [11]. In the following, we adapt these ideas to the present Lagrangian analysis. We therefore carry out a time averaging in the following sense:

$$\hat{\mathbf{x}}_i(t_{\hat{k}}) = \frac{1}{\mathcal{T}_W} \sum_{j=\hat{k}}^{\hat{k}+\mathcal{T}_W-1} \mathbf{x}_i(t_j), \quad (27)$$

where the number of time steps \mathcal{T}_W depends on the chosen window size Δt in free-fall time units. The time-averaged trajectories $\{\hat{\mathbf{x}}_i(t_{\hat{k}})\}_{\hat{k}}$ are assembled such that they represent the same time interval of length $200t_f$, i.e., \hat{k} depends on Δt . Here, we apply the minimum distance spectral trajectory clustering method (Sec. IV A) to the time-averaged trajectories. Equivalent results can be achieved by applying the time-averaged diffusion-map method. The degree of coherence of the independent sets, interpreted as the size of the spectral gap, is improved by the time averaging. Figure 11 shows the eigenvalues for the original setting, i.e., with no time

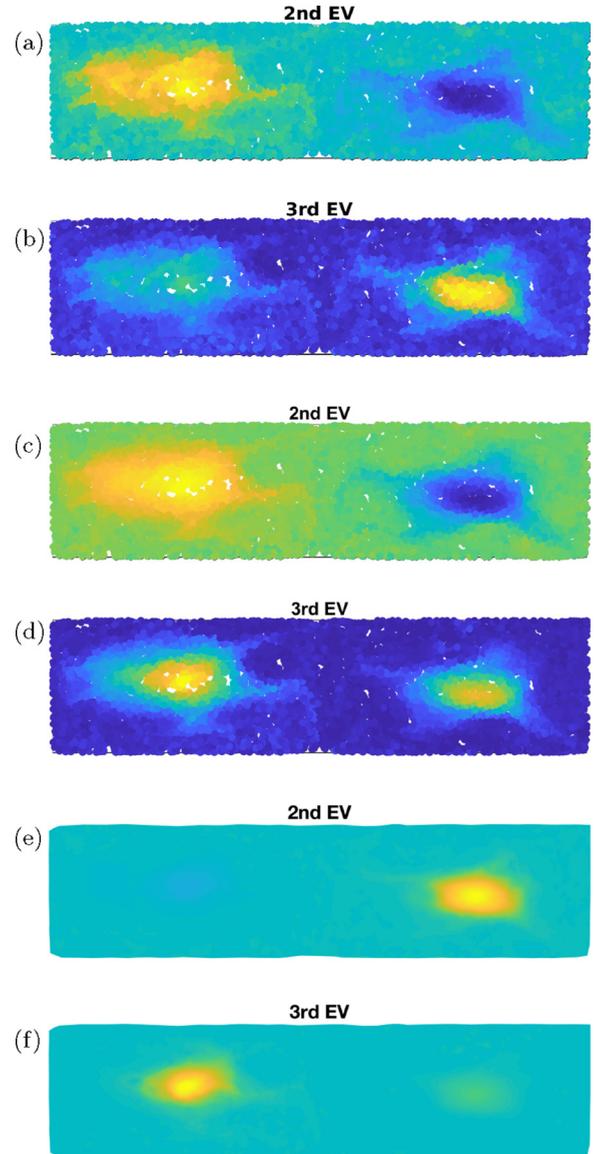


FIG. 8. Particles colored according to eigenvectors of Q^{nw} (a), (b), Q^{dm} (c), (d), and (\bar{K}, M) (e), (f) in the normalized range $[-1, 1]$. Negative (positive) values are indicated by dark (bright) contours. The specific eigenvector is indicated at the top of each panel. The first nontrivial eigenvectors at $t = 2100t_f$.

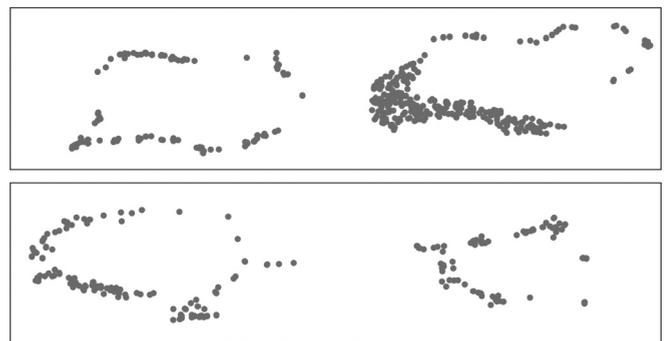


FIG. 9. Symmetrical difference of the clusters obtained from Q^{nw} and Q^{dm} at $t = 2010t_f$ (top) and $t = 2100t_f$ (bottom).

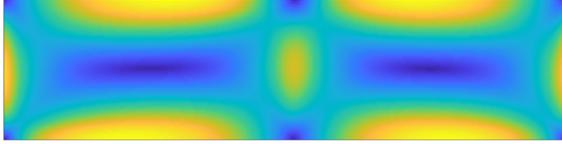


FIG. 10. Magnitude of the time-averaged advecting velocity field in the plane as a filled contour plot. The averaging time is for $2000 t_f$. The largest magnitudes are bright contours, the smallest ones are dark contours.

averaging, and for time-averaged trajectories with window sizes $\Delta t = 5t_f$ and $14t_f$ (which correspond with $\mathcal{T}_W = 50$ and 140 , respectively). For $\Delta t = 5t_f$, a prominent gap is visible between the third and fourth eigenvalues. This corresponds to three almost decoupled sets, i.e., dynamically independent flow regions, which are the two cores of the convection rolls and the background. With increasing window size transitions from one side of the domain to the other side, which occur rarely for individual trajectories in the original setting, are removed and the Lagrangian time-averaged trajectories $\{\hat{x}_i(t_k)\}_k$ mostly remain in the initial flow region. Thereby, a segmentation into inner and outer cores is formed on both sides of the domain, corresponding to four almost decoupled sets. The spectral gap is shifted to appear between the fourth and fifth eigenvalues. This segmentation of the time-averaged trajectories is visualized in Fig. 12. The transition from three to four almost decoupled sets occurs around a window size of $\Delta t = 10t_f$. We can thus conclude that an additional averaging enhances the long-term coherence of the sets significantly even for the present two-dimensional case for which the small-scale dispersion by turbulence remains moderate.

VI. HEAT COHERENCE IN CONVECTION FLOW

A. Temperature from a passive scalar perspective

The dynamical structure of heat transport can be analyzed from different perspectives. Although temperature is not a passive scalar, but a prognostic variable, once the evolution of the full system is known (i.e., the velocity \mathbf{u} and total

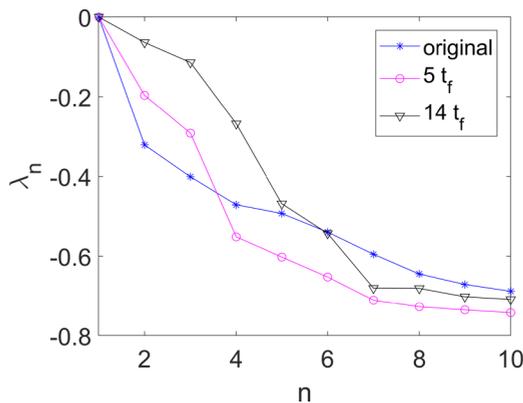


FIG. 11. Effect of time averaging in the Lagrangian analysis. Leading 10 eigenvalues of Q^{bw} of the original trajectories and time-averaged trajectories with window sizes $\Delta t = 5t_f$ and $14t_f$ for a time interval of $200.0t_f$.

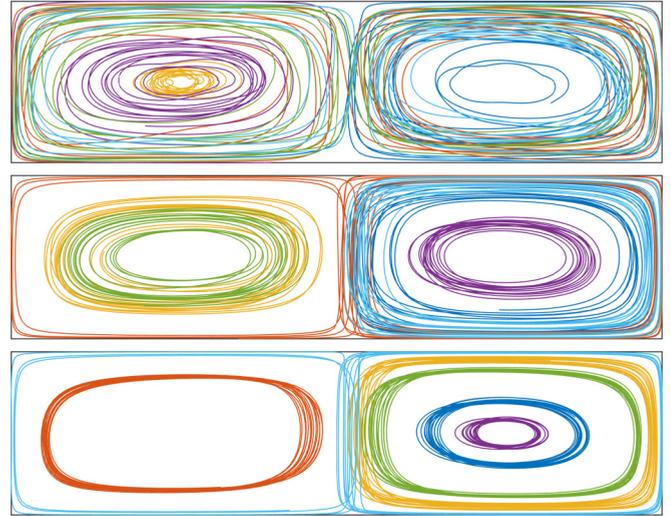


FIG. 12. Comparison of trajectories for the original and time-averaged Lagrangian analysis. Top: trajectories of six random particles for a time interval of $200.0t_f$. Middle (bottom): six time-averaged trajectories for $\Delta t = 5t_f$ ($\Delta t = 14t_f$).

temperature T fields are computed), we can view heat as a passive scalar with evolution governed by the advection-diffusion equation [see also Eq. (2)]

$$\frac{\partial T}{\partial t} = \frac{1}{\sqrt{\text{PrRa}}} \nabla^2 T - \nabla \cdot (\mathbf{u}T) =: -\nabla \cdot \Psi, \quad (28)$$

with boundary conditions that we have defined in Sec. II. Here, Ψ denotes the total dimensionless heat flux vector and the average $\langle \Psi_z \rangle_{x,t} = \text{Nu}/\sqrt{\text{RaPr}}$. The profile $T_0 = 1 - z$ is the dimensionless linear equilibrium temperature profile [see also Eq. (4)]. Furthermore, Γ_D and Γ_N are parts of the domain boundary, where prescribed constant temperature (i.e., Dirichlet condition on Γ_D) and insulating wall (i.e., Neumann condition on Γ_N) boundary conditions are applied. To recall, an insulated side wall implies that the normal derivative vanishes, $\frac{\partial T}{\partial n} = 0$.

This suggests that coherence with respect to heat could be analyzed by the approach from Sec. III. The situation is more delicate, however, as the identification of “heat packages” with particles has to be done properly. One issue is, for instance, that heat is not a conserved quantity and can enter and exit at the top and bottom boundaries Γ_D . Thus, the boundary conditions in (28) have to be incorporated in the analysis. We start by briefly discussing a popular line of approaches that is however inappropriate for heat-coherence investigations.

In the so-called “heatline” approaches [33,54–57], the temperature field is considered as a concentration field driven by some velocity field \mathbf{u}^H . In order to get the correct flux, $\Psi = \mathbf{u}^H T$ has to hold, giving

$$\mathbf{u}^H(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t) - \frac{1}{\sqrt{\text{PrRa}}} \frac{\nabla T(\mathbf{x}, t)}{T(\mathbf{x}, t)}. \quad (29)$$

Note that the system (28) is translation invariant in the following sense: For any constant $T_c \in \mathbb{R}$ and given solution T of (28), the translated field $T' = T + T_c$ solves (28) if

the initial and Dirichlet boundary conditions are translated accordingly, i.e., $T'_\Gamma = T_\Gamma + T_c$ and $T'_0 = T_0 + T_c$. This means, intuitively, that the evolution does not change even if we express temperature with respect to an arbitrary reference value.

In contrast, the “heat velocity” field \mathbf{u}^H changes under translation of the temperature field, giving completely different “heat trajectories.” Nevertheless, by construction, the *global temperature field* T is advected correctly by \mathbf{u}^H . This means that the usefulness of the “heat velocity field” is restricted to considerations involving the global temperature distribution, but *internal fluctuations* of heat and the *internal transport* of heat are biased by the choice of reference temperature.

It should be remarked at this point that in [33,58] this problem is alleviated by splitting *convective* and *conductive* contributions of heat transport, where the translational invariance is taken up completely by the conductive part. To the remaining (convective) flux a velocity field $\mathbf{u}^{\text{conv}} = \Phi^{\text{conv}}/T^{\text{conv}}$ can be assigned, which then describes the convective heat transfer in a Lagrangian manner. However, as we would like to describe structures governing the entire heat transport (and not just its convective part), we take a different route in the following.

B. Randomly evolving heat packages and induced transport

The microscopic evolution of heat can be described by advection and stochastic fluctuations. Let us consider the stochastic differential equation

$$\dot{\mathbf{x}}(t) = \mathbf{u}(\mathbf{x}(t), t) + \sqrt{\frac{4}{\text{PrRa}}} \boldsymbol{\eta}(t), \quad (30)$$

where $\mathbf{x}(t)$ denotes the random position of a particle driven by the drift \mathbf{u} and by white noise $\boldsymbol{\eta}$ (i.e., $\text{Cov}[\boldsymbol{\eta}(t), \boldsymbol{\eta}(s)] = \delta(t - s)$, where δ denotes the Dirac delta distribution). The probability distribution $p(\mathbf{x}, t)$ of $\mathbf{x}(t)$ satisfies the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = \frac{1}{\sqrt{\text{PrRa}}} \nabla^2 p - \nabla \cdot (\mathbf{u}p). \quad (31)$$

If we replace p by T , this is identical with the evolution equation in (28). Thus, scaling T to have integral one (setting the total heat to be unity), we can express the evolution of heat as the evolution of the probability distribution of an ensemble of particles, evolving mutually independently with their motions governed by (30). This ensemble has initial distribution T_0 . Single particles can thus be viewed as “heat packages,” each carrying one unit of heat.

We will illustrate some adaptations necessary according to the boundary conditions in the Rayleigh-Bénard convection problem. Neumann zero boundary conditions naturally translate into reflecting particles at the side walls Γ_N . Furthermore, we have $T = 0$ at the top and $T = 1$ for the temperature in dimensionless units. These conditions simply translate into absorbing every trajectory hitting the top lid, and reinjecting a new one at a random position at the bottom lid. Trajectories are reflected once they try to exit at the bottom lid. Note that the dynamical equation (30) is independent of the choice of a reference temperature, which is well aligned with the intuition that the dynamics of single heat packages should not depend

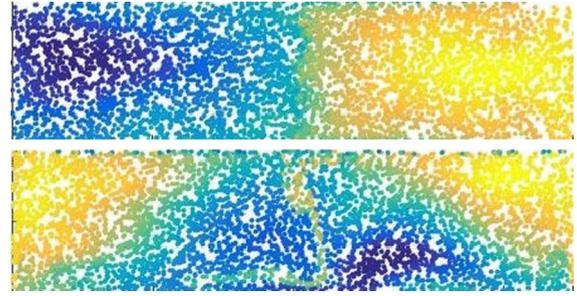


FIG. 13. Short-term heat-coherence analysis in two-dimensional Rayleigh-Bénard flow. Second eigenvector (top) and third eigenvector (bottom) are shown. Negative (positive) values are indicated by dark (bright) colors. Eigenvectors of Q^{dm} at the initial time.

on the reference temperature. The total heat transport depends on the reference temperature through the initial distribution T_0 . The absorption and reinjection of trajectories at the boundaries result in some trajectories with a shorter life span than the considered total integration time since they leave before the end or are seeded later.

Combining [26, Theorem 3] with [59, Sec. 4.1] shows that eigenvectors of the newly obtained matrix P^{dm} are relaxed solutions of the problem [cf. (12)]

$$\max_{\mathbb{A}} \frac{T(\mathbb{A} \rightarrow \mathbb{A})}{T(\mathbb{A})} + \frac{T(\mathbb{A}^C \rightarrow \mathbb{A}^C)}{T(\mathbb{A}^C)}, \quad (32)$$

where $\mathbb{A} \subset \{1, \dots, N\}$ is a subset of all trajectories (including trajectories seeded later), \mathbb{A}^C is its complement, $T(\mathbb{A})$ is the total heat content of the set \mathbb{A} (as every particle is a heat package with one unit of heat, this is the cardinality of \mathbb{A}), and $T(\mathbb{A} \rightarrow \mathbb{B})$ is the heat moving from set \mathbb{A} to set \mathbb{B} under the heat dynamics. Thus, $T(\mathbb{A} \rightarrow \mathbb{A})$ describes the heat remaining in \mathbb{A} . Solutions of (32) attempt to partition the domain into two subdomains between which there is as little heat exchange as possible, while the normalization by $T(\mathbb{A})$ and $T(\mathbb{A}^C)$ avoids highly unbalanced partitions where one of the sets contains almost all of the heat.

While the problem (32) is clearly not invariant under translation with respect to a reference temperature, in certain cases it can be argued that it would give highly consistent solutions for any reference value. Let us assume that both $T(\mathbb{A}) \approx T(\mathbb{A}^C)$, and that the trajectories in \mathbb{A} and \mathbb{A}^C cover approximately the same area in physical space. Then, problem (32) and the problem $\min_{\mathbb{A}} \{T(\mathbb{A} \rightarrow \mathbb{A}^C) + T(\mathbb{A}^C \rightarrow \mathbb{A})\}$

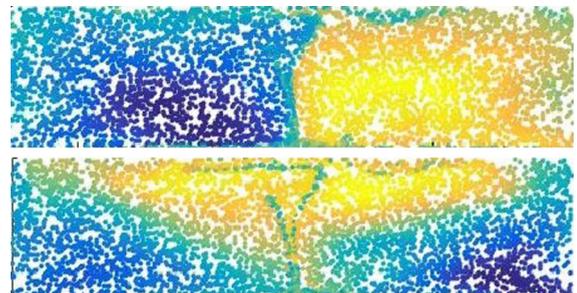


FIG. 14. Same as Fig. 13, but for the final time of the short time interval.

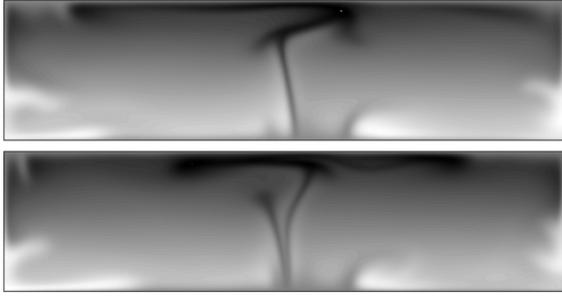


FIG. 15. Temperature contour grayscale plot in relation to the heat-coherence evaluation. Initial slice (top) at $t = 2000t_f$ and final slice (bottom) at $t = 2010t_f$. Dimensionless temperature varies between 0 (bright) and 1 (dark).

have approximately the same solutions, while the latter describes total heat exchange between the two subdomains. Now, the physically relevant quantity is the *net heat flow*, which is the difference of $T(\mathbb{A} \rightarrow \mathbb{A}^C)$ and $T(\mathbb{A}^C \rightarrow \mathbb{A})$, and this is almost invariant under changing the reference measure, if the spatial domains occupied by \mathbb{A} and \mathbb{A}^C are almost of the same size, because the contributions due to translation by a reference value cancel out.

C. Heat coherence in two-dimensional Rayleigh-Bénard convection

In the following, this methodology is applied to the present two-dimensional Rayleigh-Bénard system. We conduct the analysis for the diffusion-map method (Sec. IV B) only since the other two methods gave similar results as we saw in Sec. V. We first look again at a small time interval which means here $\mathcal{T} = 100$ steps which corresponds to a time interval of $10t_f$. If we look at the initial and the final time slices (see Figs. 13 and 14, respectively) we can see a separation of the right and left sides of the domain. Furthermore, the color of the particles that leave the domain fast and particles that enter the domain late is green, which implies that they have a zero value in the eigenvector. These Lagrangian particles correspond to thermal plumes as seen in the temperature field snapshot in Fig. 15. The algorithm identifies them as less relevant for the overall dynamics due to their short life span, which in turn implies that they are highly relevant for the effective heat transport from the bottom to the top. We now look at a longer sequence of $\mathcal{T} = 2000$ steps which corresponds to a time span of $200t_f$. The second eigenvector (see Fig. 16, top panel) reveals now that the cores of the convection rolls are the most coherent features for the temperature evolution. This implies that there is almost no effective heat transport through the cores aside from diffusion which will become increasingly subdominant as the Rayleigh number grows. It is important to note that even though the most heat-coherent sets do not contribute to vertical heat transport in this example, other cases may arise as well. If there would be big bubbles of hot fluid moving (slowly, compared to the considered time span) from the bottom to the top, the algorithm would identify them as heat coherent.

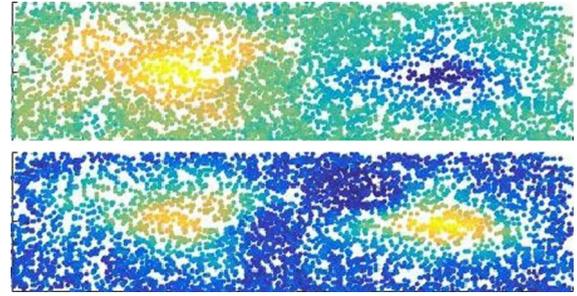


FIG. 16. Long-term heat-coherence analysis in two-dimensional Rayleigh-Bénard flow. Second eigenvector (top) and third eigenvector (bottom) are shown. Negative (positive) values are indicated by dark (bright) colors.

VII. CONCLUSION

The main objectives of this work were to discuss coherence in a simple two-dimensional turbulent convection flow from a Lagrangian point of view and to relate it to the more frequently used perspective of the Eulerian frame of reference. We therefore compared three different Lagrangian approaches: (i) minimum distance spectral-clustering based, (ii) diffusion-map based, and (iii) dynamic Laplacian based analysis of Lagrangian data. We find that all three methods identify similar coherent sets, the core regions of the two convection rolls in our example flow. These are the areas in which neighboring Lagrangian particles would remain close to each other for the longest time before dispersion by small-scale turbulence would tear them apart from each other. Our analysis shows that these regions are the complementary to those which would be highlighted in a time-averaging procedure in the Eulerian frame of reference, namely, the ridges of hot upwelling and cold downwelling fluid between these two circulation rolls. However, the notion of coherence is less strict in the minimum distance spectral-clustering based analysis compared to the diffusion-map based analysis and it varies with interval length for the dynamic Laplacian based analysis.

Furthermore, we introduced the concept of time averaging in the Lagrangian frame and demonstrated that, similar to the Eulerian case, coherence of structures is improved.

Finally, we discussed the concept of heat coherence in the present setting. We therefore suggested an approach to analyze the transport of nonpassive scalar quantities including boundary sources and sinks utilizing the theory of coherent sets. We find that effective heat transport only occurs outside the cores.

How do the present Lagrangian and the (standard) Eulerian description compare to each other? In RBC flows, the prominent structures are flow circulations between the top and bottom plates in connection with rising hot or falling cold thermal plumes. The Eulerian picture in Fig. 1 reveals a space-filling coherent pattern with a characteristic length λ (which is here simply the horizontal extension of a pair of counter-rotating rolls) and a characteristic timescale τ which is proportional to the typical turnover time inside a roll [11]. The Eulerian analysis highlights the regions of the flow that contribute most to the convective heat transfer, namely,

the hot upwellings and the cold downwellings between the counter-rotating rolls. Lagrangian methods are connected to the material transport by the evolving flow. As shown in this work, all Lagrangian methods detect the core regions of the large-scale circulation rolls as the coherent sets in which fluid particles remain together for the longest time. Thus, they reveal regions that form the spatial complement to the Eulerian ones, those that contribute least to the turbulent heat transfer. One major advantage of the discussed Lagrangian analysis methods is their objectivity, i.e., coordinate-frame independence (see, e.g., [23]) that helps to identify barriers to the turbulent transport in the flow. Lagrangian methods provide complementary powerful tools to reduce the dynamics to that of a few relevant degrees of freedom. They may not only be applicable in numerical simulations, but also for the growing number of experimental techniques that provide Lagrangian particle tracks, e.g., [60–62].

The present example was a two-dimensional flow at a moderate Rayleigh number in a working fluid with a large Prandtl number. Such a setup is an appropriate starting point for Lagrangian studies as the temperature field obeys a small diffusivity and the magnitude of turbulent velocity fluctuations remains small. The Reynolds number which quantifies the turbulent momentum transfer is $\text{Re} \approx 96$ in our case. As a part of the future work, we will extend the mathematical foundations of the present Lagrangian framework to temporally averaged turbulence fields and apply these techniques in three-dimensional settings for extended flows. This is necessary to compensate for the enhanced turbulent dispersion which is always probed by Lagrangian methods and which increases as Rayleigh numbers are increased or Prandtl numbers are decreased. These efforts are partly under way and will be reported elsewhere.

ACKNOWLEDGMENTS

This work of C.S., M.S., and A.P. is supported by the Priority Programme DFG-SPP 1881 “Turbulent Superstructures” of the Deutsche Forschungsgemeinschaft (DFG). We acknowledge computing resources by the Large Scale Project pr62se of the Gauss Centre for Supercomputing. The authors thank C. Cierpka, D. Karrasch, and N. Schilling for helpful comments.

APPENDIX: ADDITIONAL MATERIAL

1. Interpretation in terms of graph Laplacians

The stiffness matrices K^k are symmetric and have zero row and column sums [32]. Its diagonal entries are positive and the off-diagonal entries nonpositive in certain important cases (and also in all numerical experiments). Because of this, for the graph G^k with nodes X^k and edges defined by the edges of the triangulation, we can write $K^k = \Pi^k - W^k$, where

$$W_{ij}^k = \begin{cases} -K_{ij}^k, & i \neq j \\ 0, & i = j \end{cases}$$

are the *weights* assigned to edges (i, j) , and $\Pi_{ii}^k = K_{ii}^k$ is a diagonal matrix with the i th diagonal element equal to the

degree of node i in G^k , namely, $\Pi_{ii}^k = -\sum_j K_{ij}^k$. Thus, we can view K^k as an (unnormalized) graph Laplacian.

Note that the entry K_{ij}^k will decrease in magnitude as the distance between the associated data points $\mathbf{x}_i(t_k)$, $\mathbf{x}_j(t_k)$ increases until the Delaunay edge (i, j) no longer exists in G^k .

When we solve the eigenproblem $\bar{K}v = \lambda Mv$, we normalize by the symmetric, non-negative mass matrix M , i.e., based on local area or volume elements that neighboring data points enclose. This normalization is different to the standard graph Laplacian normalization (which is based on node degree only): M is not diagonal and the small number of off-diagonal entries of M coincide with the off-diagonal entries of K^0 and correspond to arcs (i, j) that are in the graph G^0 . In fact, normalizing by the mass matrix *automatically handles nonuniformly distributed data* because if initial points $\mathbf{x}_i^0, \mathbf{x}_j^0$ are far apart, the value of M_{ij} will be commensurately larger.

Note further that since a triangulation is used here, there is no free parameter (like the cutoff radius) to choose and that the method can always yield a decomposition of the entire domain Ω (more precisely, the convex hull of the data points) into coherent sets.

2. Comparison of methods based on the rate-matrix interpretation

From the point of view of rate matrices in Sec. VA, on the one hand, Q^{nw} defines a process where every state has holding time 1 (because $Q_{ii}^{\text{nw}} = 1$), and the process has equal probabilities $1/D_{ii}$ to jump to any of its “neighbors,” where i and j are neighbors if $A_{ij} = 1$. We observe that the proximity parameter ε can also be interpreted as diffusion strength (diffusion coefficient), as the holding times of the Markov process are all one, and the “jumps” cover an ε neighborhood in space. This can be seen from noting that a Brownian diffusion $c B_t$ of strength c has standard deviation $\sqrt{\text{Var}[c B_t]} = c\sqrt{t}$, i.e., if it has strength $c = \varepsilon$, then it produces ε mean deviation in unit time. Thus, in a first order approximation, we could interpret Q^{nw} as a diffusion on trajectories with strength ε .

On the other hand, Q^{dm} defines a random walk where the i th state has holding time

$$\frac{\delta}{1 - \frac{1}{T} \sum_k \frac{1}{\sum_j K_{k,ij}}}, \quad (\text{A1})$$

which follows from the construction noting that $K_{k,ii} = 1$. We observe immediately that the scale parameter δ features directly as a timescale. As the “range” of the kernel³ and thus the mean jump distance is $\mathcal{O}(\sqrt{\delta})$, a similar consideration as above shows that Q^{dm} can be interpreted as a diffusion of strength 1. Returning to (A1), we see that the holding time grows very large if the trajectory i has only few and distant neighbors (i.e., it is unlikely to jump over to trajectories that are not alike), and approaches δ from above if the trajectory has many close neighbors. The jump probabilities are readily encoded in the entries of P^{dm} , and due to the construction

³The cutoff radius r is defined by $k_\delta(\mathbf{x}, \mathbf{y}) = h(\rho^2/\delta) < \vartheta$ for all $\|\mathbf{x} - \mathbf{y}\| = \rho > r$, with some small threshold parameter ϑ . This directly yields $r = \text{const} \times \sqrt{\delta}$.

involving the similarity matrix, it is more likely that the process jumps to a neighbor which is closer on average (in time).

To summarize this theoretical comparison of the methods from Secs. IV A and IV B, both use spectral clustering with matrices that are interpretable as rate matrices of certain Markov jump processes on the trajectories. By this, those trajectories belong to the same coherent set which are likely to be reached from one another, opposed to unlikely transitions from other trajectories. This behavior is often referred to as *metastability* or *almost invariance*, and its connection to the dominant eigenmodes of the process' jump matrix is well analyzed [63–68]. The difference in the methods considered here relies in characteristics of the associated Markov processes:

The network-based process jumps to its neighbors with equal probabilities (thus utilizing only a binary information about distances, whether they are smaller or larger than ε), while the diffusion-map based process prefers to jump to closer neighbors (thus utilizing a more refined distance information). While the proximity parameter ε can be viewed as strength of the Markov process (diffusion) generated on the trajectories by Q^{nw} , the strength of the analogous diffusion generated by Q^{dm} is always one. As these diffusion processes are discrete in space (because of jumping between a finite set of trajectories), their mean jump distance governs the finest scales they can resolve: $\mathcal{O}(\varepsilon)$ and $\mathcal{O}(\sqrt{\delta})$. Coherent sets below these scales can not be detected.

-
- [1] M. S. Miesch, *Living Rev. Solar Phys.* **2**, 1 (2005).
- [2] U. R. Christensen and J. Aubert, *Geophys. J. Int.* **166**, 97 (2006).
- [3] B. Stevens, *Annu. Rev. Earth Planet. Sci.* **33**, 605 (2005).
- [4] D. Bouffard and A. Wuest, *Annu. Rev. Fluid Mech.* **51**, 189 (2019).
- [5] F. Chillà and J. Schumacher, *Eur. Phys. J. E* **35**, 58 (2012).
- [6] T. Hartlep, A. Tilgner, and F. H. Busse, *Phys. Rev. Lett.* **91**, 064501 (2003).
- [7] J. von Hardenberg, A. Parodi, G. Passoni, A. Provenzale, and E. A. Spiegel, *Phys. Lett. A* **372**, 2223 (2008).
- [8] J. Bailon-Cuba, M. S. Emran, and J. Schumacher, *J. Fluid Mech.* **655**, 152 (2010).
- [9] M. S. Emran and J. Schumacher, *J. Fluid Mech.* **776**, 96 (2015).
- [10] R. J. A. M. Stevens, A. Blass, X. Zhu, R. Verzicco, and D. Lohse, *Phys. Rev. Fluids* **3**, 041501(R) (2018).
- [11] A. Pandey, J. D. Scheel, and J. Schumacher, *Nat. Commun.* **9**, 2118 (2018).
- [12] J. Bailon-Cuba and J. Schumacher, *Phys. Fluids* **23**, 077101 (2011).
- [13] B. Podvin and A. Sergent, *Phys. Fluids* **24**, 105106 (2012).
- [14] S. Paul and M. K. Verma, Proper Orthogonal Decomposition vs Fourier analysis for extraction of large-scale structures of thermal convection, in *Advances in Computation, Modeling and Control of Transitional and Turbulent Flows* (World Scientific, Singapore, 2017), pp. 433–441.
- [15] M. K. Verma, *Physics of Buoyant Flows* (World Scientific, Singapore, 2018).
- [16] G. Froyland, S. Lloyd, and N. Santitissadeekorn, *Physica D (Amsterdam)* **239**, 1527 (2010).
- [17] G. Froyland, *Physica D (Amsterdam)* **250**, 1 (2013).
- [18] M. R. Allshouse and T. Peacock, *Chaos* **25**, 097617 (2015).
- [19] D. Karrasch and J. Keller, [arXiv:1608.05598](https://arxiv.org/abs/1608.05598).
- [20] M. Dellnitz, G. Froyland, and O. Junge, *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems* (Springer, Berlin, 2001), pp. 145–174.
- [21] M. Dellnitz and O. Junge, *Handbook of Dynamical Systems*, Vol. 2 (North-Holland, Amsterdam, 2002), pp. 221–264.
- [22] A. Hadjighasem, M. Farazmand, D. Blazevski, G. Froyland, and G. Haller, *Chaos* **27**, 053104 (2017).
- [23] G. Haller, *Annu. Rev. Fluid Mech.* **47**, 137 (2015).
- [24] G. Haller, D. Karrasch, and F. Kogelbauer, *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9074 (2018).
- [25] G. Froyland, *Nonlinearity* **28**, 3587 (2015).
- [26] R. Banisch and P. Koltai, *Chaos* **27**, 035804 (2017).
- [27] K. Padberg-Gehle and C. Schneide, *Nonlin. Process. Geophys.* **24**, 661 (2017).
- [28] C. Schneide, A. Pandey, K. Padberg-Gehle, and J. Schumacher, *Phys. Rev. Fluids* **3**, 113501 (2018).
- [29] K. L. Schlueter-Kuck and J. O. Dabiri, *J. Fluid Mech.* **811**, 468 (2017).
- [30] G. Froyland and K. Padberg-Gehle, *Chaos* **25**, 087406 (2015).
- [31] A. Hadjighasem, D. Karrasch, H. Teramoto, and G. Haller, *Phys. Rev. E* **93**, 063107 (2016).
- [32] G. Froyland and O. Junge, *SIAM J. Appl. Dyn. Sys.* **17**, 1891 (2018).
- [33] M. F. M. Speetjens, *J. Phys.: Conf. Ser.* **395**, 012033 (2012).
- [34] J. Schmalzl, M. Breuer, and U. Hansen, *Europhys. Lett.* **67**, 390 (2004).
- [35] E. P. van der Poel, R. J. A. M. Stevens, and D. Lohse, *J. Fluid Mech.* **736**, 177 (2013).
- [36] A. Pandey, M. K. Verma, A. G. Chatterjee, and B. Dutta, *Pramana: J. Phys.* **87**, 13 (2016).
- [37] K. Petschel, M. Wilczek, M. Breuer, R. Friedrich, and U. Hansen, *Phys. Rev. E* **84**, 026309 (2011).
- [38] A. Pandey, M. K. Verma, and M. Barma, *Phys. Rev. E* **98**, 023109 (2018).
- [39] M. K. Verma, A. G. Chatterjee, K. S. Reddy, R. K. Yadav, S. Paul, M. Chandra, and R. Samtaney, *Pramana: J. Phys.* **81**, 617 (2013).
- [40] M. K. Verma, S. C. Ambhire, and A. Pandey, *Phys. Fluids* **27**, 047102 (2015).
- [41] G. Haller, *Physica D (Amsterdam)* **149**, 248 (2001).
- [42] K. Padberg, B. Thiere, R. Preis, and M. Dellnitz, *Commun. Nonlinear Sci. Numer. Simul.* **14**, 4176 (2009).
- [43] R. V. Donner, S. Barbosa, J. F. Kurths, and N. Marwan, *Eur. Phys. J. Spec. Top.* **174**, 1 (2009).
- [44] R. Banisch, P. Koltai, and K. Padberg-Gehle, *Chaos* **29**, 063125 (2019).
- [45] J. Shi and J. Malik, *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888 (2000).
- [46] S. P. Lloyd, *IEEE Inf. Theory* **28**, 2 (1982).
- [47] R. R. Coifman and S. Lafon, *Appl. Comput. Harmonic Anal.* **21**, 5 (2006).
- [48] G. Froyland and E. Kwok, *J. Nonlinear Sci.*, 1 (2017), doi: 10.1007/s00332-017-9397-y.

- [49] J.-L. Thiffeault, C. R. Doering, and J. D. Gibbon, *J. Fluid Mech.* **521**, 105 (2004).
- [50] M. Dellnitz and O. Junge, *SIAM J. Numer. Anal.* **36**, 491 (1999).
- [51] A. Denner, Ph.D. thesis, TU Munich, 2017.
- [52] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, *Phys. Rev. E* **81**, 015101(R) (2010).
- [53] K. Petschel, S. Stellmach, M. Wilczek, J. Lülf, and U. Hansen, *Phys. Rev. Lett.* **110**, 114502 (2013).
- [54] S. Kimura and A. Bejan, *J. Heat Transfer* **105**, 916 (1983).
- [55] V. A. Costa, *Appl. Mech. Rev.* **59**, 126 (2006).
- [56] L. Pratt, R. Barkan, and I. Rypina, *Fluids* **1**, 27 (2016).
- [57] S. Balasuriya, N. T. Ouellette, and I. Rypina, *Physica D (Amsterdam)* **372**, 31 (2018).
- [58] M. F. M. Speetjens, *Int. J. Therm. Sci.* **61**, 79 (2012).
- [59] G. Froyland and K. Padberg-Gehle, in *Ergodic Theory, Open Dynamics, and Coherent Structures. PROMS, Vol. 70*, edited by W. Bahsoun *et al.* (Springer, New York, 2014).
- [60] M. Novara, D. Schanz, R. Geisler, S. Gesemann, C. Voss, and A. Schröder, *Exp. Fluids* **60**, 44 (2019).
- [61] C. Cierpka, B. Lütke, and C. J. Kähler, *Exp. Fluids* **54**, 1533 (2013).
- [62] D. Schanz, S. Gesemann, and A. Schröder, *Exp. Fluids* **57**, 70 (2016).
- [63] E. B. Davies, *Proc. London Math. Soc.* **s3-45**, 133 (1982).
- [64] E. B. Davies, *J. London Math. Soc.* **s2-26**, 541 (1982).
- [65] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein, *Commun. Math. Phys.* **228**, 219 (2002).
- [66] G. Froyland, *Physica D (Amsterdam)* **200**, 205 (2005).
- [67] W. Huisinga and B. Schmidt, in *New Algorithms for Macromolecular Simulation*, edited by B. Leimkuhler *et al.*, Lecture Notes in Computational Science and Engineering Vol. 49 (Springer, Berlin, 2016).
- [68] C. Schütte and M. Sarich, *Metastability and Markov State Models in Molecular Dynamics*, Courant Lecture Notes in Mathematics Vol. 24 (AMS, Providence, RI, 2013).