



Finite-time entropy: A probabilistic approach for measuring nonlinear stretching

Gary Froyland^a, Kathrin Padberg-Gehle^{b,*}

^a School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2031, Australia

^b Institute of Scientific Computing, Technische Universität Dresden, 01062 Dresden, Germany

ARTICLE INFO

Article history:

Received 14 December 2011

Received in revised form

1 June 2012

Accepted 25 June 2012

Available online 29 June 2012

Communicated by J. Dawes

Keywords:

Transfer operator

Entropy

Finite-time Lyapunov exponent

Set-oriented approach

Transport

ABSTRACT

Transport and mixing processes in dynamical systems are often difficult to study analytically and therefore a variety of numerical methods have been developed. Finite-time Lyapunov exponents (FTLEs) or related stretching indicators are frequently used as a means to estimate transport barriers. Alternatively, eigenvectors, singular vectors, or Oseledets vectors of numerical transfer operators find almost-invariant sets, finite-time coherent sets, or time-asymptotic coherent sets, respectively, which are minimally dispersed under the dynamics. While these families of approaches (geometric FTLEs and the probabilistic transfer operator) often give compatible results, a formal link is still missing; here we present a small step towards providing a mathematical link.

We propose a new entropy-based methodology for estimating finite-time expansive behaviour along trajectories in autonomous and nonautonomous dynamical systems. We introduce the finite-time entropy (FTE) field as a simple and flexible way to capture nonlinear stretching directly from the entropy growth experienced by a small localised density evolved by the transfer operator. The FTE construction elucidates in a straightforward way the connection between the evolution of probability densities and the local stretching experienced.

We develop an extremely simple and numerically efficient method of constructing an estimate of the FTE field. The FTE field is instantaneously calculable from a numerical transfer operator – a transition matrix of conditional probabilities that describes a discretised version of the dynamical system; once one has such a transition matrix, the FTE field may be computed “for free”. We also show (i) how to avoid long time integrations in autonomous and time-periodic systems, (ii) how to perform backward time computations by a fast matrix manipulation rather than backward time integration, and (iii) how to easily employ adaptive methods to focus on high-value FTE regions.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The numerical analysis of transport and mixing processes in dynamical systems and aperiodic flows has been the subject of much research over the past 25 years. This is largely due to the wide range of applications including astrodynamics, molecular dynamics, fluid dynamics, climate and ocean dynamics; see e.g. [1–4] for discussions of transport and mixing phenomena.

Assuming that the motion of a particle is determined by an underlying autonomous or nonautonomous dynamical system allows for an application of different concepts from dynamical systems theory to analyse and quantify transport. The developed approaches fall roughly into two classes (i) *geometric* methods which make use of invariant manifolds and Lagrangian coherent

structures and (ii) *probabilistic* techniques which are based on the notions of almost-invariant and finite-time coherent sets.

Geometrical structures such as invariant manifolds are known to play a key role in dynamical transport and mixing. In particular, in the context of lobe dynamics [5,6,3] one makes use of the fact that segments of invariant manifolds of hyperbolic objects form partial transport barriers. In the aperiodic time-dependent settings, finite-time hyperbolic material surfaces [7–9] have been proposed as generalisations of invariant manifolds that form barriers to mixing. These Lagrangian coherent structures are often studied computationally based on finite-time Lyapunov exponents (FTLE), see e.g. [7,8,10,4,9,11]. This quantity measures the growth of an infinitesimal perturbation under the dynamics. Ridges in the FTLE field may indicate transport barriers, see [9] for the mathematical conditions on Lagrangian coherent structures.

Probabilistic approaches study the evolution of densities and aim at detecting regions in phase space that are minimally dispersed under the action of the dynamical system. These regions are known as *almost-invariant* or metastable sets in autonomous

* Corresponding author.

E-mail addresses: g.froyland@unsw.edu.au (G. Froyland), kathrin.padberg@tu-dresden.de (K. Padberg-Gehle).

systems. The concept arose about 15 years ago in the context of dynamical systems [12,13] and time-symmetric Markov processes [14,15]. The main theoretical and computational tool is the Perron–Frobenius (or transfer) operator and almost-invariant sets were estimated heuristically from eigenfunctions of the Perron–Frobenius operator. Further theoretical and computational extensions have since been constructed [16–18]. Very recently a transfer operator-based treatment of coherent sets in nonautonomous systems has been proposed in the time-asymptotic [19,20] and finite-time [21,22] settings.

Both families of approaches (geometric and probabilistic) have a number of advantages and disadvantages, but give very similar answers in case studies; see [23] for a detailed discussion as well as for further references on prior work. However, formal results on the connection between the two approaches are still missing.

Work related to present paper includes [24,25], where it was observed that information about stretching such as measured by FTLE can also be extracted from a discretised model of the smooth dynamics via an associated directed graph. Certain expansion concepts related to the neighbourhood of a vertex in the graph were studied and it was found that vertices with high expansion correspond to regions in phase space with high FTLE values. In [26] the covariance of an image of an initial density under the dynamics is computed, and from this an estimate of stretching is obtained. In [27] a very similar approach for stochastic differential equations is proposed, whereby the covariance matrix of stochastically integrated sample trajectories is computed to obtain an estimate of “stretching”. The main point of [26] is to attempt to measure nonlinear stretching effects on a small, but not infinitesimal, neighbourhood, although these effects are only partly captured by the use of covariance, which is itself a linear statistic. The methods [26,27] heavily depend on the final geometric shape of the image of the small neighbourhood. For example, in two-dimensions a long thin linear image (stretching only) would receive a high “stretching” value, while the same image curled around itself (stretching + folding) would receive a low “stretching” value; in contrast, the method put forward in the present paper would assign essentially equal stretching values in these two cases.

In this paper we give a mathematically precise definition of nonlinear stretching based on the evolution of probability densities and present a very efficient computational approach that makes direct use of a discretised transfer operator. We measure the growth rate of uncertainty generated by a small initial random ϵ -perturbation of the original nonlinear dynamics. This uncertainty is captured by measuring the entropy increase experienced by a probability density under the action of the dynamical system. We show that under very general conditions (the dynamics need not even be differentiable) a well-defined stretching quantity, called *finite-time entropy* (FTE), exists, which defines a continuous field on the state space. Under mild conditions on the differentiability of the system, we prove the existence of a “deterministic limit” of the FTE fields as $\epsilon \rightarrow 0$, and show that this limit is comparable to FTLE fields. Thus, we may view FTE as a formal nonlinear generalisation of FTLE that operates on ϵ -scales, rather than infinitesimal scales.

We detail a very fast numerical approximation scheme—especially in the cases of autonomous flows/maps and time-periodic flows. Our numerical approach easily extends to aperiodic flows. In addition to being more computationally efficient than the approaches of [24–26] and simpler to implement, our approach has a number of theoretical advantages, including (i) provision of a truly *nonlinear* stretching estimate at every ϵ -perturbation or discretisation level, (ii) robustness to discretisation or ϵ -perturbation level used, with a convergence result in the zero perturbation limit, and (iii) continuity of the FTE field for every ϵ -perturbation level.

The paper is organised as follows: in Section 2 the new probabilistic concept of finite-time entropy is introduced and in Section 3 we show how this approach naturally fits into a set-oriented framework. Section 4 describes the extension to nonautonomous dynamical systems. Computational aspects are addressed in Section 5, before we demonstrate the strength of our approach with several example systems and compare to other concepts in Section 6. We conclude with a discussion about open topics and future work in Section 7.

2. Definition and properties of finite-time entropy

For the development of the theoretical background we consider a discrete dynamical system $T : X \rightarrow X$, where T is assumed to be a diffeomorphism on a compact set $X \subset \mathbb{R}^d$. By T^k , $k \in \mathbb{Z}$ we will denote the k -fold application of T (or the inverse system T^{-1} , in case $k < 0$). Note that T might, for instance, be a time-1 flow map of some autonomous ordinary differential equation. As will be shown later in Section 4 there is a straightforward extension to nonautonomous dynamical systems.

The growth of infinitesimal perturbations in the initial conditions under the dynamics of the map T^k , $k \in \mathbb{Z}$, is typically measured in terms of the maximum finite-time Lyapunov exponent (FTLE)

$$\mathcal{E}_k(x_0) = \frac{1}{2|k|} \log (\lambda_{\max}[DT^k(x_0)^T DT^k(x_0)]), \quad (1)$$

where $DT^k(x_0) = \prod_{i=0}^{k-1} DT(x_i)$, $T(x_i) = x_{i+1}$, $i = 0, \dots, k-1$ denotes the total derivative of T^k at x_0 (analogously for the inverse system) and $\lambda_{\max}[M]$ is the largest eigenvalue of a matrix M . Thus, $\mathcal{E}_k(x_0)$ measures the largest relative length growth rate of any vector under the action of the linearised system $DT^k(x_0)$. In the following we want to describe a probabilistic framework that enables one to measure the growth of uncertainties under the nonlinear dynamics.

For this we consider the evolution of probability densities under the action of the dynamical system. Denote the Lebesgue measure on X by m and let $f \in L^1(X, m)$ be a density; that is, $f \geq 0$ and $\int f dm = 1$. The natural push-forward of f under the map T is given by the Perron–Frobenius operator $\mathcal{P} : L^1(X, m) \rightarrow L^1(X, m)$ defined by

$$\mathcal{P}f(x) = \frac{f(T^{-1}x)}{|\det DT(T^{-1}x)|}. \quad (2)$$

For this definition to make sense, T should be a non-singular transformation with respect to Lebesgue measure (see eg. [28, p. 42]). In [Theorem 2.7](#) we ask that $|\det DT|$ be uniformly bounded away from zero; this guarantees (and is slightly stronger than) non-singularity.

We begin by providing a formal definition of finite-time entropy and later develop a very efficient set-oriented approach for its computation in Sections 3 and 5.

Definition 2.1. Let $f \in L^1(X)$ be a density. The *differential entropy* of f is defined by $h(f) = -\int_S f \log f dm$, where S is the support of f .

The notion of differential entropy goes back to Boltzmann and we refer the reader to Chapter 9 [28] for a discussion in the dynamical systems context. The following well-known lemma elucidates the relationship between entropy and local expansion and contraction.

Lemma 2.2. $h(\mathcal{P}^k f) = h(f) + \int f \log |\det DT^k| dm$.

Proof. We prove the result for $k = 1$ and the result immediately extends to $k \in \mathbb{Z}$ by applying the $k = 1$ result to the map T^k .

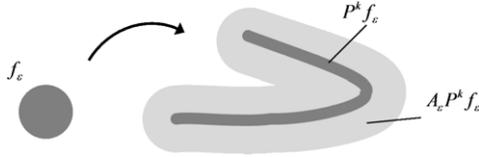


Fig. 1. Action of the smoothing operator \mathcal{A}_ϵ in two dimensions. An initial density f_ϵ (dark disc on left) gets stretched and folded (dark, thin sausage on right) under the evolution of the Perron–Frobenius operator. The additional application of the smoothing operator fattens the support of the image density (light, thick sausage on right); this thickening is more pronounced along contracting directions. This has the (desired) effect that the finite-time entropy (Eq. (4)) measures predominantly the nonlinear stretching.

$$\begin{aligned} h(\mathcal{P}f) &= - \int \mathcal{P}f \log \mathcal{P}f \, dm(x) \\ &= - \int \frac{f(T^{-1}x)}{|\det DT(T^{-1}x)|} \cdot \log \frac{f(T^{-1}x)}{|\det DT(T^{-1}x)|} \, dm(x) \\ &= - \int f(y) \cdot \log(f(y)/|\det DT(y)|) \, dm(y) \\ &\quad (\text{change of variables } y = T^{-1}x), \\ &= h(f) + \int f(y) \cdot \log |\det DT(y)| \, dm(y). \quad \square \end{aligned}$$

Thus the increase in differential entropy from f to $\mathcal{P}^k f$ is the weighted average (according to f) of the field of the sum of all finite-time Lyapunov exponents. In the case of volume-preserving systems, since $DT \equiv 1$, it immediately follows that $h(\mathcal{P}^k f) = h(f)$ for all $k \in \mathbb{Z}$. Further statements relating the entropy and conditional entropy of f to those of $\mathcal{P}^k f$ may be found in [28]; original work for discrete and continuous state spaces is contained in [29,30], respectively. As we are interested in measuring nonlinear stretching we need to find a way to isolate expansion from contraction in the above formula. Therefore we introduce a smoothing operator to neutralise the entropy-reducing effect of the contractive dynamics.

Definition 2.3. Let $B_\epsilon(x_0)$ denote an ϵ -ball centred on x_0 , and $f_{\epsilon,x_0} := (1/m(B_\epsilon(x_0)))\mathbf{1}_{B_\epsilon(x_0)}$ a uniform density supported on $B_\epsilon(x_0)$. The ϵ -smoothing operator \mathcal{A}_ϵ is the averaging operator with radius ϵ defined by $\mathcal{A}_\epsilon f(x) := (1/m(B_\epsilon(x))) \int_{B_\epsilon(x)} f \, dm$.

Under the action of the Perron–Frobenius operator a uniform density supported on a ball of radius ϵ will be stretched, folded and contracted as prescribed by the underlying map T . The application of the ϵ -smoothing operator fattens this image density by increasing the density support. This fattening has a dominant effect along contracting directions; see Fig. 1 for an illustration. Thus the entropy increase experienced in this set-up will predominantly be related to stretching along expanding directions and this is what we wish to measure via finite-time entropy:

Definition 2.4. Given $\epsilon > 0$ we define the finite-time entropy at x_0 by

$$FTE_\epsilon(x_0, k) := \frac{1}{|k|} [h(\mathcal{A}_\epsilon \mathcal{P}^k f_{\epsilon,x_0}) - h(f_{\epsilon,x_0})]. \quad (3)$$

We define the deterministic limit of finite-time entropy at x_0 by

$$FTE(x_0, k) := \lim_{\epsilon \rightarrow 0} FTE_\epsilon(x_0, k). \quad (4)$$

In words, $FTE_\epsilon(x, k)$ is the rate of increase in entropy experienced in an ϵ -neighbourhood of x over k iterations of T , followed

by an ϵ -perturbation. This rate of increase in entropy in a neighbourhood of x may also be loosely thought of as the rate of increase in the uncertainty in the future position of $T^k x$ under the random ϵ -perturbed dynamics.

Theorem 2.5. Let $T : X \rightarrow X$ be non-singular. For each fixed $\epsilon > 0$ and k , $FTE_\epsilon(x, k)$ is a continuous function of $x \in X$.

Proof. Fix ϵ, k and x . Let $(x_n)_{n \in \mathbb{N}}$ be a sequence where $x_n \rightarrow x$ as $n \rightarrow \infty$. We show that $FTE_\epsilon(x_n, k) \rightarrow FTE_\epsilon(x, k)$ as $n \rightarrow \infty$.

To do this, we show $|h(\mathcal{A}_\epsilon \mathcal{P}^k f_{\epsilon,x_n}) - h(\mathcal{A}_\epsilon \mathcal{P}^k f_{\epsilon,x})| \rightarrow 0$ and $|h(f_{\epsilon,x_n}) - h(f_{\epsilon,x})| \rightarrow 0$. For the latter, we note that $\|(1/m(B_\epsilon(x_n)))\mathbf{1}_{B_\epsilon(x_n)} - (1/m(B_\epsilon(x)))\mathbf{1}_{B_\epsilon(x)}\|_1 \rightarrow 0$ and apply Lemma A.1. For the former, we note that (i) $\|\mathcal{A}_\epsilon \mathcal{P}^k\|_1 \leq 1$ and (ii) $\mathcal{A}_\epsilon \mathcal{P}^k(1/m(B_\epsilon(x_n)))\mathbf{1}_{B_\epsilon(x_n)}$ and $\mathcal{A}_\epsilon \mathcal{P}^k(1/m(B_\epsilon(x)))\mathbf{1}_{B_\epsilon(x)}$ are bounded densities (bounded because $\mathcal{P}^k(1/m(B_\epsilon(x)))\mathbf{1}_{B_\epsilon(x)}$ is a density and $\mathcal{A}_\epsilon f \leq (1/m(B_\epsilon)) \int f dm$ for any $f \geq 0$). We argue as above and apply Lemma A.1 again. The result follows. \square

For the special case of affine maps we make the following observation:

Lemma 2.6. If T is affine then $FTE_\epsilon(x, k)$ is independent of ϵ .

Proof. We again prove the result only for $k = 1$, as it immediately extends to $k \in \mathbb{Z}$ by applying the $k = 1$ result to T^k . Let $f_{\epsilon,x} := (1/m(B_\epsilon(x)))\mathbf{1}_{B_\epsilon(x)}$ and WLOG assume that the density $f_{\epsilon,x}$ is centred at the origin; we now drop the x subscript. Denote by $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_d > 0$ the d singular values of the linear part of T . Since T is affine, $\mathcal{P}f_\epsilon$ is a uniform density supported on an ellipsoid with radii given by the d values $\epsilon\sigma_1, \epsilon\sigma_2, \dots, \epsilon\sigma_d$. WLOG we assume below that this ellipsoid, denoted $E_\epsilon(0)$ is centred at the origin. Now,

$$\begin{aligned} F_\epsilon(x) &:= (\mathcal{A}_\epsilon \mathcal{P}f_\epsilon)(x) = \frac{1}{m(B_\epsilon(x))} \int_{B_\epsilon(x)} \frac{1}{m(E_\epsilon(0))} \mathbf{1}_{E_\epsilon(0)}(y) \, dy \\ &= \frac{1}{m(B_\epsilon)^2 \prod_{i=1}^d \sigma_i} \int_{B_\epsilon(x)} \mathbf{1}_{E_\epsilon(0)}(y) \, dy \\ &= \frac{\epsilon^d}{m(B_\epsilon)^2 \prod_{i=1}^d \sigma_i} \int_{B_1(x/\epsilon)} \mathbf{1}_{E_1(0)}(z) \, dz, \quad \text{using } z = y/\epsilon, \\ &= \frac{1}{\epsilon^d m(B_1)^2 \prod_{i=1}^d \sigma_i} \int_{B_1(x/\epsilon)} \mathbf{1}_{E_1(0)}(z) \, dz \\ &= (1/\epsilon^d) F_1(x/\epsilon). \end{aligned}$$

A simple calculation shows $h(F_\epsilon) = h(F_1) + \log \epsilon^d$. Since $h(f_\epsilon) = h(f_1) + \log \epsilon^d$, the result follows. \square

We denote by $T_x^{A,k}$ the affine Taylor approximation of T^k centred on x : $T_x^{A,k}(y) = T^k(x) + DT^k(x) \cdot (y - x)$, and denote by $FTE^A(x, k)$ the (ϵ -independent by Lemma 2.6) finite-time entropy of $T^{A,k}$ at x (i.e. over k steps).

The following theorem ensures the existence of the finite-entropy $FTE(x, k)$ under very general conditions on T :

Theorem 2.7. Let $T : X \rightarrow X$ be C^2 and $|\det DT(x)|$ be uniformly bounded away from zero. Fix $x \in X$. Then the limit $FTE(x, k) = \lim_{\epsilon \rightarrow 0} FTE_\epsilon(x, k)$ exists and equals $FTE^A(x, k)$.

Proof. See the Appendix. \square

In general it is difficult to write down closed form expressions for $FTE_\epsilon(x, k)$, so we provide two motivational examples to argue that $FTE_\epsilon(x, k)$ is a good measure of stretching.

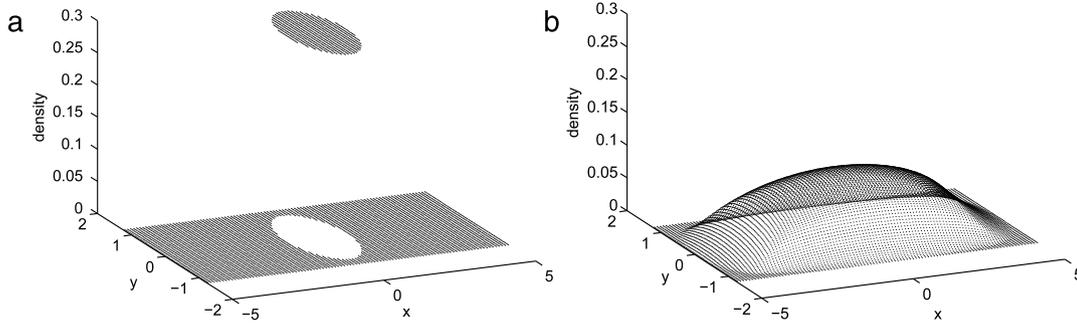


Fig. 2. (a) An initial density in \mathbb{R}^2 : Graph of density $f_{\epsilon,0}$ where $\epsilon = 1$. (b) Smoothed image of an initial density in \mathbb{R}^2 : Graph of $\mathcal{A}_\epsilon \mathcal{P} f_{\epsilon,0}$ where $T(x, y) = (4x, y/4)$ and $\epsilon = 1$.

Example 2.8 (Affine Map in 1D). Let $Tx = \alpha x + c$, $\alpha > 1$, $c \in \mathbb{R}$. Consider a density $f = \mathbf{1}_I/m(I)$ on an interval $I \subset \mathbb{R}$ of length L . Then it is straightforward to show that $h(f) = \log L$ and $h(\mathcal{A}_\epsilon f) = \epsilon/L + \log L$. In particular, for $L = 2\epsilon$ we obtain $h(f) = \log 2\epsilon$. As $\mathcal{P}^k f$ is a uniform density on an interval of length $\alpha^k L$ it follows that $h(\mathcal{A}_\epsilon \mathcal{P}^k f) = \epsilon/(2\alpha^k \epsilon) + \log(2\alpha^k \epsilon)$. Thus $FTE_\epsilon(x, k) = (1/k)(1/(2\alpha^k) + \log(2\alpha^k \epsilon) - \log(2\epsilon)) = \log \alpha + 1/(2k\alpha^k)$ for all $\epsilon > 0$, $x \in \mathbb{R}$ and $k \geq 1$. Thus, $FTE_\epsilon(x, k)$ captures the expansion rate α plus a small error term that decreases exponentially in k .

The following example for a simple linear map of \mathbb{R}^2 gives an indication of the approximate values of $FTE(x, k)$.

Example 2.9 (Affine Map in 2D). Consider the affine map $T(x, y) = (4x, y/4)$, with singular values $\sigma_1 = 4$ and $\sigma_2 = 1/4$. By Lemma 2.6 $FTE_\epsilon(x, k)$ is independent of ϵ and in this example we have chosen $\epsilon = 1$. Fig. 2(a) shows the initial uniform density $f_{1,0}$ supported on a disc of radius 1 centred at the origin, while Fig. 2(b) shows a graph of $\mathcal{A}_1 \mathcal{P} f_{1,0}$.

While the support of $\mathcal{P} f_{1,0}$ is a long, thin ellipse with radii 4 and $1/4$, the smoothing operator \mathcal{A}_1 spreads this support, with the dominant effect occurring in the y -coordinate. The support of $\mathcal{A}_1 \mathcal{P} f_{1,0}$ is contained in a region $[-(\sigma_1 + 1), \sigma_1 + 1] \times [-(\sigma_2 + 1), \sigma_2 + 1]$. For $\sigma_1 \gg 1 \gg \sigma_2$, this smoothed density is thus approximately a stretching of the initial density by a factor σ_1 and so $FTE(0, 1) \approx \log \sigma_1$.

In Section 6 we will demonstrate the strength of the finite-time entropy concept in several nonlinear example systems and also compare with the results obtained from finite-time Lyapunov exponent computations.

3. Set-oriented description of finite-time entropy

The concept of finite-time entropy fits naturally into the set-oriented framework for the global analysis of dynamical systems [31–33]. In particular, for reasons of computational efficiency we wish to utilise a discrete and finite rank approximation of the Perron–Frobenius operator \mathcal{P} . A standard approach for numerically approximating \mathcal{P} is Ulam’s method [34]; we will describe further computational aspects in Section 5.1. One partitions X into a large collection of connected sets $\{B_1, \dots, B_n\}$. In practice the partition is defined via a regular grid and each B_i is a d -dimensional box. One then considers a finite-rank approximation of \mathcal{P} , restricting the action of \mathcal{P} to the basis of characteristic functions $\text{sp}\{\mathbf{1}_{B_1}, \dots, \mathbf{1}_{B_n}\}$. The matrix representation of this action is

$$Q_{ij} = \frac{m(B_i \cap T^{-1}B_j)}{m(B_j)}. \quad (5)$$

It is usually more convenient to consider this discretised action on probability measures rather than densities. The row stochastic matrix

$$P_{ij} = \frac{m(B_i \cap T^{-1}B_j)}{m(B_j)} \quad (6)$$

represents the action on discrete probability measures. That is, if $v \geq 0$ is a probability vector ($\sum_i v_i = 1$) then the probability vector v' defined by $v' = vP$ is the push-forward of the probability vector v under the discretised action of T .

In this discrete setting, we represent the density functions of the previous section by discrete probability measures that assign mass to each partition element B_i , $i = 1, \dots, n$. Let μ be such a probability measure. We define the entropy of μ , denoted $H(\mu)$ by $H(\mu) = -\sum_{i=1}^n \mu(B_i) \log \mu(B_i)$. Writing $p_i = \mu(B_i)$ we have $H(p) = -\sum_{i=1}^n p_i \log p_i$.

The natural action of T on μ is $T^* \mu = \mu \circ T^{-1}$. So our discrete approximation of $T^* \mu(B_j)$ is $\sum_{i=1}^n p_i P_{ij} = (pP)_j$.

Definition 3.1. Assume that the elements of the partition $\{B_1, \dots, B_n\}$ have equal¹ volume.

Let δ_i denote the n -vector with a 1 in the i th position and 0 elsewhere. We define the discrete finite-time entropy on partition set B_i to be

$$FTE(B_i, k) = \frac{1}{|k|} H(\delta_i P^{(k)}) = -\frac{1}{|k|} \sum_{j=1}^n P_{ij}^{(k)} \log P_{ij}^{(k)} \quad (7)$$

where $P^{(k)}$ denotes the discretised transfer operator with respect to the map T^k .

In MATLAB code, one may calculate (7) as

```
[i, j, nonzeroPk]=find(Pk);
FTE=-(1/abs(k))*sum(sparse(i, j, nonzeroPk
.*log(nonzeroPk))');;
```

where $Pk = P^{(k)}$ is a sparse transition matrix, yielding an n -vector FTE containing $FTE(B_i, k)$, $i = 1, \dots, n$.

Eq. (7) mimics (3). The vector δ_i replaces $f_{\epsilon, x}$ where $x \in B_i$ and ϵ is of the order of the radius of the set B_i . The vector $\delta_i P^{(k)}$ replaces $\mathcal{A}_\epsilon \mathcal{P}^k f_{\epsilon, x}$ as the action of $P^{(k)}$ on measures is approximately that of $\mathcal{A}_\epsilon \mathcal{P}^k$ on densities, namely an application of $(T^k)^*$ followed by a smoothing of the order of the radii of the boxes B_j that are intersected by $T^k(B_i)$. We could also have written $FTE(B_i, k) = \frac{1}{|k|} (H(\delta_i P^{(k)}) - H(\delta_i))$ to make the analogy more obvious as $H(\delta_i) = 0$. For the purposes of estimating finite-time entropy, the sets B_i should be close to spherically symmetric to avoid bias related to

¹ The treatment of differing box volumes will be addressed in Section 5.3.

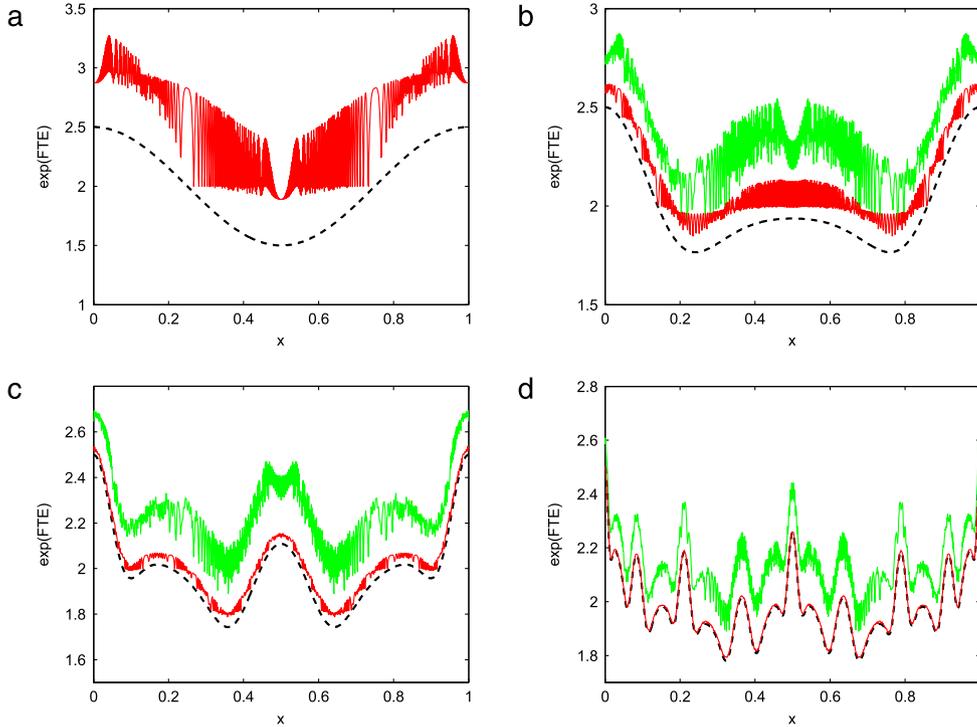


Fig. 3. Comparison of the FTE-based expansion rates with the analytical (dashed curves) ones for the nonlinear doubling map. (a) Graph of $FTE(x, 1)$ vs. x (dark solid (red, online)); (b) graph of $FTE(x, 2)$ vs. x (dark solid (red, online)), graph of $FTE_{\text{mult}}(x, 2)$ vs. x (light solid (green, online)); (c) graph of $FTE(x, 3)$ vs. x (dark solid (red, online)), graph of $FTE_{\text{mult}}(x, 3)$ vs. x (light solid (green, online)); (d) graph of $FTE(x, 5)$ vs. x (dark solid (red, online)), graph of $FTE_{\text{mult}}(x, 5)$ vs. x (light solid (green, online)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the orientation of the image set $T^k(B_i)$. In practice we have found that choosing B_i as cubes is perfectly adequate.

To illustrate the discrete setting we return to the examples (Examples 2.8 and 2.9) of the previous section as well as introduce a nonlinear doubling map.

Example 3.2 (Affine Map in 1D). Consider a bounded interval $X \subset \mathbb{R}$ and a uniform 2ϵ -partition $\{B_1, \dots, B_n\}$ of X . The grid set $B_i \subset \mathbb{R}$ has length 2ϵ and we are concerned with the image of B_i under $Tx = \alpha x + c$, for $\alpha > 1$, $c \in \mathbb{R}$. Assume that α , c and $k \in \mathbb{N}$ are chosen so that the interval $T^k(B_i) \subset X$. The interval $T^k(B_i)$, $k \geq 1$, has length $2\epsilon\alpha^k$ and will cover either $\lfloor \alpha^k \rfloor$ or $\lfloor \alpha^k \rfloor - 1$ grid sets with “overhang” fractions of $0 \leq \eta_1, \eta_2 \leq 1$ on the two partially covered grid sets. The values of η_1, η_2 depend on the interplay of α , c , and the grid. The discrete measure given by $\delta_i P$ has either $\lfloor \alpha^k \rfloor$ or $\lfloor \alpha^k \rfloor - 1$ sets with measure $1/\alpha^k$. We treat the former case and the latter case follows similarly.

$$\begin{aligned} H(\delta_i P) &= -\lfloor \alpha^k \rfloor ((1/\alpha^k) \log(1/\alpha^k)) - (\eta_1/\alpha^k) \log(\eta_1/\alpha^k) \\ &\quad - (\eta_2/\alpha^k) \log(\eta_2/\alpha^k) \\ &= -((\lfloor \alpha^k \rfloor + \eta_1 + \eta_2)/\alpha^k) \log(1/\alpha^k) \\ &\quad - (\eta_1/\alpha^k) \log(\eta_1) - (\eta_2/\alpha^k) \log(\eta_2) \\ &= k \log \alpha + (\eta_1/\alpha^k) \log(1/\eta_1) + (\eta_2/\alpha^k) \log(1/\eta_2). \end{aligned}$$

Hence, for $1 \leq i \leq n$ and $k \geq 1$, one has

$$\begin{aligned} FTE(B_i, k) &= \log \alpha + (\eta_1/k\alpha^k) \log(1/\eta_1) \\ &\quad + (\eta_2/k\alpha^k) \log(1/\eta_2). \end{aligned} \quad (8)$$

Thus, the value $FTE(B_i, k)$ captures the logarithm of the stretching factor α with additional error terms due to the grid-based diffusion or smoothing that decrease exponentially with k . By differentiation, we determine that the largest error occurs when $\eta_1 = \eta_2 = e^{-1}$; then $FTE(B_i, k) = \log \alpha + 2/(k\alpha^k)$. Compare this with Example 2.8, where the additional term was $1/(2k\alpha^k)$.

In practice, we typically wish to estimate FTE values for large α corresponding to the stretching experienced under several iterations of a map. The error terms will then be small in comparison to $\log \alpha$.

Example 3.3 (Nonlinear Doubling Map). We illustrate the expansion estimates computationally for a nonlinear expanding map. Define a nonlinear doubling map $T : [0, 1] \rightarrow [0, 1]$ by

$$T(x) = 2x + \frac{\sin(2\pi x)}{4\pi} \pmod{1}.$$

We have modified the standard doubling map $x \mapsto 2x \pmod{1}$ by adding a nonlinear term. The expansion now ranges from $3/2$ to $5/2$ (see Fig. 3(a) (dashed)).

We partition the unit interval into $n = 1024$ equally sized intervals and use $R = 10000$ “inner grid” points per box for the computation of the transition matrix P . This large value of R has been deliberately chosen so that we can be sure that errors due to sparse sampling will be negligible in our studies.

Fig. 3 shows graphs of the exact derivatives $(T^k)'(x)$ vs. x (dashed) and graphs of $FTE(x, k)$ vs. x (where $FTE(x, k)$ is evaluated at the centre point x of each interval). The FTE estimates appear to converge very quickly to the exact derivatives, compatible with the estimates from Example 3.2.

Example 3.4 (Affine Map in 2D). Consider a bounded rectangle $X \subset \mathbb{R}^2$ and a uniform partition $\{B_1, \dots, B_n\}$ of X into squares of side length 2ϵ . The grid set $B_i \subset \mathbb{R}^2$ has area $4\epsilon^2$ and we are concerned with the image of B_i under $T(x, y) = (\alpha x, \beta y) + (c_1, c_2)$, for $\alpha > 1$ and $0 < \beta < 1$, $c_1, c_2 \in \mathbb{R}$. Assume that $\alpha, \beta, c_1, c_2, k$ are chosen so that the rectangle $T^k(B_i) \subset X$. The rectangle $T^k(B_i)$, $k \geq 1$ has

² See [32].

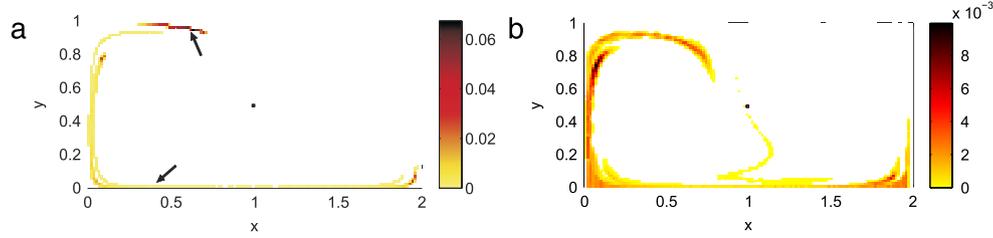


Fig. 4. Evolution of a uniform density initialised in the small black box B_{2048} by the discretised transfer operator. The boxes are coloured according to density (darker colour = higher density). (a) $FTE(B_{2048}, 5) = 0.6524$ which corresponds to a stretching factor of $\alpha^5 = 26.1$. The image density is shown as the top dark connected component supported on 31 boxes (indicated by upper arrow). The approximation of $FTE(B_{2048}, 10) \approx FTE_{\text{mult}}(B_{2048}, 10, 5) = 0.5260$ via $P^{(5)}$, corresponds to stretching factor $\alpha^{10} = 192.4$. The image density is supported on 331 boxes (light-coloured tangled structure close to the boundary as indicated by lower arrow). (b) $FTE_{\text{mult}}(B_{2048}, 10, 1) = 0.6243$, corresponding to a stretching of $\alpha^{10} = 514.5$. The image density is supported on 1692 boxes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

side lengths $2\epsilon\alpha^k$ and $2\epsilon\beta^k$. As $\beta < 1$ there are two main cases for the positioning of $T^k(B_i)$ relative to the grid.

Case 1: $T^k(B_i)$ fits entirely within a single horizontal strip of grid sets and has overhang fractions η_1, η_2 in the x -direction as in the one-dimensional example above. Here one has

$$FTE(B_i, k) = \log \alpha + (\eta_1/k\alpha^k) \log(1/\eta_1) + (\eta_2/k\alpha^k) \log(1/\eta_2). \quad (9)$$

Case 2: $T^k(B_i)$ straddles two horizontal strips of grid sets and has overhang fractions η_1, η_2 in the x -direction and overhang fractions γ_1, γ_2 in the y -direction (note $\gamma_1 + \gamma_2 = \beta^k$). Here one obtains

$$FTE(B_i, k) = \log \alpha + (\eta_1/k\alpha^k) \log(1/\eta_1) + (\eta_2/k\alpha^k) \times \log(1/\eta_2) + \log \beta + (\gamma_1/k(\gamma_1 + \gamma_2)) \times \log(1/\gamma_1) + (\gamma_2/k(\gamma_1 + \gamma_2)) \log(1/\gamma_2). \quad (10)$$

The derivation follows from repeated arguments of the type in [Example 3.2](#). In (9) the maximal error above $\log \alpha$ is $2/(k\epsilon\alpha^k)$ as in the one-dimensional example. In (10) the maximum additional error due to the terms involving $\gamma_1 + \gamma_2 = \beta^k$ occurs when $\gamma_1 = \gamma_2 = \beta^k/2$. Substituting these values into the last three terms of (10) yields $\log \beta + (1/2k) \log(2/\beta^k) + (1/2k) \log(2/\beta^k) = (1/k) \log 2$ so that the maximum error above $\log \alpha$ is $2/(k\epsilon\alpha^k) + (1/k) \log 2$.

From (9) and (10) it thus follows for this example that

$$\log \alpha \leq FTE(B_i, k) \leq \log \alpha + 2/(k\epsilon\alpha^k) + (1/k) \log 2, \quad (11)$$

where the upper bound converges to the lower bound $\log \alpha$ with increasing k .

We note that similar analyses can be carried out to study the effects of rotation (so that the sides of $T^k(B_i)$ are not parallel with grid boundaries). The maximum errors (including possible underestimates for some configurations) are of the same order as the errors discussed above.

Example 3.5 (Nonlinear Stretching in 2D). We now illustrate how FTE can quantify nonlinear stretching by studying the evolution of an initial density concentrated in a box of side length 2^{-6} under the double gyre flow.

The double gyre flow [10] is given by the time-periodic velocity field

$$\begin{aligned} \dot{x} &= -\pi A \sin(\pi f(x, t)) \cos(\pi y) \\ \dot{y} &= \pi A \cos(\pi f(x, t)) \sin(\pi y) \end{aligned} \quad (12)$$

where $f(x, t) = \epsilon \sin(\omega t)x^2 + (1 - 2\epsilon \sin(\omega t))x$. It has become a well-known test case for computations of FTLE fields and Lagrangian coherent structures, see e.g. [35,23] for recent analyses.

We choose parameter values $A = 0.25, \epsilon = 0.25, \omega = 2\pi$ and fix $t_0 = 0$. We obtain a flow of period $\tau = 1$ on the domain

$X = [0, 2] \times [0, 1]$. The time-1 map $T : X \rightarrow X$ is the Poincaré return map on the time-expanded domain at time slice $t = 0$. We compute the transition matrices for $k = 5$ and $k = 10$ iterates of T ; more details on the computational aspects will be discussed in [Sections 5 and 6.1](#).

[Fig. 4\(a\)](#) shows the evolution of an initially uniform density supported on the box B_{2048} shown in black at about $[1, 0.5]$ under the action of the discretised transfer operators $P^{(5)}$ and as well as an approximation of $P^{(10)}$ by $(P^{(5)})^2$, see [Section 5.1](#) for details on such an approximation. The side length of the box is 0.0156 , mimicking an ϵ -ball with $\epsilon = 0.0078$.

The image density under $P^{(5)}$ is shown as the dark connected component at the top of [Fig. 4\(a\)](#), the image density obtained by applying $(P^{(5)})^2$ is supported on the complicated light-coloured structure close to the boundary. The location of the two images is indicated by arrows.

We obtain $FTE(B_{2048}, 5) = 0.6524$ which corresponds to a stretching factor of $\alpha^5 = 26.1$. The image density is supported on 31 boxes, where visually the width of the support is for most parts only one box. Hence the FTE-value approximately measures – as desired – the length of the nonlinear strip (which appears to be well approximated by the value 26.1) and thus only the expansive behaviour. Similar estimates hold for the approximation of $FTE(B_{2048}, 10) \approx FTE_{\text{mult}}(B_{2048}, 10, 5) = 0.5260$, corresponding to a stretching factor of $\alpha^{10} = 192.4$ with the image density supported on 331 boxes. The width of the support is between one and two boxes so that 192.4 appears to be a good estimate of the length.

4. Finite-time entropy in nonautonomous systems

In this section we briefly show how the concept of finite-time entropy extends to nonautonomous dynamical systems. For this we assume that we are given a sufficiently smooth time-dependent vector field $f(x, t)$ where $x \in X, t \in \mathbb{R}$, and where X is again a compact subset of \mathbb{R}^d . The flow map $\Phi(x, t, \tau) : X \times \mathbb{R} \times \mathbb{R} \rightarrow X$ outputs the final position of a point x initialised at time t after flowing for duration τ . Similarly to (2) we can define a transfer operator $\mathcal{P}_{t,\tau} : L^1(X, m) \rightarrow L^1(X, m)$ related to the flow map $\Phi(\cdot, t, \tau)$ by

$$\mathcal{P}_{t,\tau} f(x) = \frac{f(\Phi(x, t + \tau, -\tau))}{|\det D\Phi(\Phi(x, t + \tau, -\tau), t, \tau)|}. \quad (13)$$

So if $f(x)$ is a density at time t , $\mathcal{P}_{t,\tau} f(x)$ describes the density at time $t + \tau$ evolved by the flow map.

Definition 4.1. Let $\epsilon > 0$. In analogy to (3) and (4) we define the finite-time entropy at x_0 by

$$FTE_\epsilon(x_0, t, \tau) := \frac{1}{|\tau|} [h(\mathcal{A}_\epsilon \mathcal{P}_{t,\tau} f_{\epsilon, x_0}) - h(f_{\epsilon, x_0})]. \quad (14)$$

We define the deterministic limit of finite-time entropy at x_0 by

$$FTE(x_0, t, \tau) := \lim_{\epsilon \rightarrow 0} FTE_\epsilon(x_0, t, \tau). \quad (15)$$

As in [22] we are typically not interested in considering all of X but in analysing some subset $U \subset X$ and some neighbourhood $V \supset \Phi(U, t, \tau)$ of the image of U . Within a set-oriented approach one therefore considers a fine partition $\{B_1, \dots, B_m\}$ of U as well as a fine partition $\{C_1, \dots, C_n\}$ of V , where we for now assume that all partition elements have equal size, in Section 5.3 this will be generalised to differing box volumes. In analogy to (6) the discretised action on probability measures can then be constructed as the (possibly rectangular) row-stochastic matrix

$$P^{(\tau)}(t)_{ij} = \frac{m(B_i \cap \Phi(C_j, t + \tau, -\tau))}{m(B_i)}, \quad (16)$$

and we obtain simply

$$\begin{aligned} FTE(B_i, t, \tau) &= \frac{1}{|\tau|} H(\delta_i P^{(\tau)}(t)) \\ &= -\frac{1}{|\tau|} \sum_{j=1}^n P^{(\tau)}(t)_{ij} \log P^{(\tau)}(t)_{ij}; \end{aligned} \quad (17)$$

compare with (7). Thus the previous definitions for autonomous systems easily extend to the case of nonautonomous systems.

5. Computational aspects

5.1. Construction of transition matrix P and fast computation of FTE

The numerical computation of $P^{(k)}$ is accomplished by numerical integration of trajectories of sample points.

Autonomous case: First we consider the autonomous system $T : X \rightarrow X$. Within each box B_i , $i = 1, \dots, n$ of the partition of X choose uniformly distributed (for example, on a uniform grid) sample points $q_{i,r}$, $r = 1, \dots, R$. Then estimate

$$P_{ij}^{(k)} \approx \frac{\#\{r : T^k(q_{i,r}) \in B_j\}}{R}. \quad (18)$$

To simplify notation we set $P := P^{(1)}$. Repeated push-forwards of a probability vector v are achieved by matrix multiplication. Thus to push v forward by k iterates of our discretised action of T we compute $v^k = vP^k$. The matrix P is typically very sparse and the multiplication is very fast. Therefore, for the computation of the finite-time entropy we can exploit the approximation $v^k = vP^k \approx vP^{(k)}$; i.e. the repeated pushforward of the discretised transfer operator for T gives a similar result as a single pushforward with $P^{(k)}$. The probability measure described by v^k is slightly more diffuse and has larger support than the probability measure described by $vP^{(k)}$ due to repeated implicit perturbations at each matrix iteration. We define:

$$\begin{aligned} FTE_{\text{mult}}(B_i, k) &= \begin{cases} \frac{1}{k} H(\delta_i P^k), & \text{for } k > 1, \\ \frac{1}{|k|} H(\delta_i (P^{(-1)})^{|k|}), & \text{for } k < -1. \end{cases} \\ &= \begin{cases} -\frac{1}{k} \sum_{j=1}^n (P^k)_{ij} \log (P^k)_{ij}, & \text{for } k > 1, \\ -\frac{1}{|k|} \sum_{j=1}^n ((P^{(-1)})^{|k|})_{ij} \log ((P^{(-1)})^{|k|})_{ij}, & \text{for } k < -1. \end{cases} \end{aligned} \quad (19)$$

When treating flows, the underlying transition matrix may not correspond to an integer amount of time. In these cases we include a third parameter, i.e. $FTE_{\text{mult}}(\cdot, \text{totaltime}, \text{basetime})$, where *basetime* denotes the integration time for setting up the transition

matrix. For instance $FTE_{\text{mult}}(\cdot, 2, 0.5)$ would denote the FTE_{mult} -field using $(P^{(0.5)})^4$ as an approximation to $P^{(2)}$. If there are only two arguments of FTE_{mult} then *basetime* = 1 by convention.

The entries of $FTE_{\text{mult}}(\cdot, k)$ in Eq. (19) can be computed for positive k (and similarly for negative k) by directly considering the result of the respective matrix–matrix products of the sparse matrix P ; in MATLAB simply:

```
Pk=P^k;
[i,j,nonzeroPk]=find(Pk);
FTE_mult=-(1/k)*sum(sparse(i,j,nonzeroPk
.*log(nonzeroPk))';
```

where P is the one-step transition matrix, k the number of iterates and FTE_{mult} an n -vector containing the values of $FTE_{\text{mult}}(B_i, k)$, $i = 1, \dots, n$. However, most likely for large k the resulting matrix will no longer be sparse and one gains no memory saving from the sparse matrix representation.

In this case the rows $(P^k)_i$ can be obtained by computing the repeated push-forward of the vector δ_i by the matrix P , i.e. $\delta_i P^k = (\delta_i P^{k-1})P$. Here only k matrix–vector products with sparse P are needed. In MATLAB the respective FTE calculation would be realised by a for-loop over all boxes:

```
FTE_mult=zeros(1,n);
for i=1:n
    delta=zeros(1,n);
    delta(i)=1;
    for l=1:k
        delta=delta*P;
    end;
    I=find(delta);
    FTE_mult(i)=-1/k*dot(delta(I), log(delta(I)));
end;
```

Here P denotes again the one-step transition matrix, n the number of boxes, k the number of iterates and FTE_{mult} is the vector containing the values of $FTE_{\text{mult}}(B_i, k)$, $i = 1, \dots, n$. For illustration of this concept we briefly return to the nonlinear doubling map and the double gyre flow (Examples 3.3 and 3.5):

Example 5.1 (Nonlinear Doubling Map). In Fig. 3 the light solid curves correspond to the results obtained by considering P^k instead of $P^{(k)}$ (dark solid curves) for the computation of finite-time entropy. While the latter quickly converges to the analytical values, the FTE_{mult} -curves overestimate the true stretching—but in a consistent manner so that the qualitative stretching behaviour is picked up correctly. The overestimation is due to the creation of pseudo-orbits induced by the repeated matrix multiplication, leading to image measures with a larger support.

Example 5.2 (Nonlinear Stretching in 2d). We consider the image density under application of P^{10} as shown in Fig. 4(b), with $FTE_{\text{mult}}(B_{2048}, 10) = 0.6243$, corresponding to a stretching of $\alpha^{10} = 514.5$. The image density is supported on 1692 boxes, where the width the support appears to about three boxes on average. So $\alpha^{10} = 514.5$ is again a good approximation for the length of density support but due to the numerically induced diffusion the true stretching is overestimated. In Section 6.1 we will demonstrate that the overall overestimation is consistent and provides qualitatively convincing results.

Nonautonomous case: In nonautonomous systems the entries $P^{(\tau)}(t)_{ij}$, $i = 1, \dots, m$, $j = 1, \dots, n$, are approximated as

$$P^{(\tau)}(t)_{ij} \approx \frac{\#\{r : \Phi(q_{i,r}, t, \tau) \in C_j\}}{R} \quad (20)$$

where sample points $q_{i,r}$, $r = 1, \dots, R$ are chosen from a uniform distribution in each box B_i , $i = 1, \dots, m$.

In practice, if τ is large, one may estimate $P^{(\tau)}(t)$ as $P^{(s_1)}(t) \cdot P^{(s_2)}(t + s_1) \cdots P^{(s_N)}(t + s_1 + \cdots + s_{N-1})$ where $s_1 + \cdots + s_N = \tau$. Larger N will lead to greater numerical diffusion as discussed for FTE_{mult} .

5.2. Calculating FTE in backward time

While regions of high FTLE values in forward time indicate stable manifolds or repelling Lagrangian coherent structures (see e.g. [7,10,24]), unstable manifolds and attracting Lagrangian coherent structures will appear as ridges in the backward-time field. We will investigate the behaviour of finite-time entropy in both time directions in Section 6. In the following we will show how to efficiently compute FTE-fields in backward time.

Assume that we are given a row-stochastic matrix P that represents the dynamics of some map S , which either represents some mapping T^k or the dynamics of a flow map $\Phi(\cdot, t, \tau)$. We now develop a simple formula for computing a suitable backward-time transition matrix, thus avoiding a direct computation of a new matrix using S^{-1} in place of S .

We derive the formula for the backward-time transition matrix, denoted \tilde{P} , to approximate $P^{(-1)}$ in the general case of boxes with different volume and possibly where one has two distinct domains at the initial and final times, partitioned as $\{B_1, \dots, B_m\}$, $\{C_1, \dots, C_n\}$, respectively. If the initial domain is also the final domain, then $n = m$ and we identify $B_i = C_i$.

We are interested in volumes of intersection of sets of the form $B_i \cap S^{-1}C_j$; the points in B_i that land in C_j under one iteration of S . Note that $m(B_i \cap S^{-1}C_j) = m(B_i)P_{ij}$ by (16). In backward time, the Lebesgue mass in the set C_j is “sent back” to sets B_i for which $m(B_i \cap S^{-1}C_j) > 0$. The transition matrix

$$\tilde{P}_{ji} = \frac{m(B_i \cap S^{-1}C_j)}{m(S^{-1}C_j)}$$

is a row-stochastic matrix of conditional probabilities that consistently sends Lebesgue mass back from C_j to B_i ; the normalisation $m(S^{-1}C_j)$ in the denominator is the volume of the mass sent back (measured at the initial time t in the nonautonomous setup). Now,

$$\tilde{P}_{ji} = \frac{m(B_i \cap S^{-1}C_j)}{m(S^{-1}C_j)} = \frac{m(B_i)P_{ij}}{\sum_{i=1}^m m(B_i)P_{ij}},$$

so that $\tilde{P} \approx P^{(-1)}$ is computable directly from P and the individual box volumes. If $m(B_i) = m(C_j)$ for all i, j , then \tilde{P} is simply the transpose of P with the row sums normalised. In MATLAB this matrix-reversing reads:

```
Q=diag(sparse(m))*P;
Q=Q*diag(sparse(1./sum(Q)));
Ptilde=Q';
```

Here P is the original transition matrix, m is a vector containing the Lebesgue measure $m(B_i)$ for each box B_i , $i = 1, \dots, n$ and P_{tilde} is the resulting approximate backward-time matrix. We denote the finite-time entropy with respect to $\tilde{P}^{(k)}$ obtained from $P^{(k)}$ by $\tilde{FTE}(\cdot, k)$ (analogously $\tilde{FTE}_{mult}(\cdot, k)$ using \tilde{P}). In nonautonomous systems we obtain $\tilde{FTE}(\cdot, t, \tau) \approx FTE(\cdot, t + \tau, -\tau)$.

In Section 6.1 we will demonstrate in an example system that $\tilde{FTE}(\cdot, k)$ is a very good approximation to $FTE(\cdot, -k)$.

5.3. Computing FTE with boxes of differing volume

We derive extensions of (7) and (17) for the situation where the boxes $\{B_1, \dots, B_m\}$ are not of equal volume. Given a discrete probability measure μ described by the n -vector p we form a density function f for μ as $f = \sum_{i=1}^n \bar{f}_i \mathbf{1}_{B_i}$ where

$$\bar{f}_i = (1/m(B_i)) \int_{B_i} f \, dm = \mu(B_i)/m(B_i) = p_i/m(B_i).$$

Now,

$$\begin{aligned} h(f) &= - \int f \log f \, dm = - \sum_{i=1}^n \int_{B_i} \bar{f}_i \log \bar{f}_i \, dm \\ &= - \sum_{i=1}^n \bar{f}_i m(B_i) \log \bar{f}_i = - \sum_{i=1}^n p_i \log p_i / m(B_i). \end{aligned}$$

Thus we define for the autonomous case

$$\begin{aligned} FTE(B_i, k) &= - \frac{1}{|k|} \left(\sum_{j=1}^n (\delta_i P^{(k)})_j \log \frac{(\delta_i P^{(k)})_j}{m(B_j)} - \log \frac{1}{m(B_i)} \right) \\ &= - \frac{1}{|k|} \left(\sum_{j=1}^n (P^{(k)})_{ij} \log \frac{(P^{(k)})_{ij}}{m(B_j)} + \log m(B_i) \right). \end{aligned} \quad (21)$$

Example code for MATLAB reads as follows:

```
[i, j, s]=find(Pk);
s=log(s).*s-s.*log(m(j)');
FTE=-1/abs(k)*(sum(sparse(i, j, s)')+log(m)');
```

Here Pk is the k -step transition matrix and m is a column vector with entries corresponding to the Lebesgue measure of each box. Of course when $m(B_i) = m(B_j)$ for all i, j then (21) collapses to (7).

The ability to handle partition sets of different volumes enables the construction of adaptive algorithms that preferentially focus on boxes with high FTE values. This application is described further in Section 6.1 and illustrated in Fig. 8.

In the nonautonomous case if one has two distinct partitions for the initial and final domains $\{B_1, \dots, B_m\}$ and $\{C_1, \dots, C_n\}$, then one defines

$$\begin{aligned} FTE(B_i, t, \tau) &= - \frac{1}{|\tau|} \left(\sum_{j=1}^n (P^{(\tau)}(t))_{ij} \log \frac{(P^{(\tau)}(t))_{ij}}{m(C_j)} + \log m(B_i) \right). \end{aligned} \quad (22)$$

This latter approach is used in the polar vortex example in Section 6.5.

6. Examples

We demonstrate the strength of the new FTE concept by a number of example systems and discuss computational aspects. First we study the double gyre flow in detail before we turn to a numerical analysis of nonlinear stretching on the Lorenz attractor. By considering the ABC flow we demonstrate that our approach can also calculate nonlinear stretching in three dimensions. The idealised stratospheric flow serves as an example for a nonautonomous system. Finally we will consider the dynamics of the polar vortex to approximate transport barriers in a real-world system.

For most examples we will also show the results of FTLE computations for comparison. However, we note that the focus of this paper is to introduce a purely probabilistic stretching diagnostic, analyse its properties and demonstrate them in example systems.

The FTE calculations obviously depend, like FTLE, on the number of iterations of the underlying map. As the support of the densities $\mathcal{A}_\epsilon \mathcal{P}^k f_\epsilon$ have a minimum width of ϵ , in the presence of hyperbolicity, this support will fill the entire space for large enough k . For such large k , one can no longer expect an increase in $h(\mathcal{A}_\epsilon \mathcal{P}^k f_\epsilon)$ that is linear in k . Eventually the value of $h(\mathcal{A}_\epsilon \mathcal{P}^k f_\epsilon)$ will asymptote to the maximum possible value of $\log m(X)$. The smaller ϵ is, the longer it will take the push-forward densities to “fill” the space X and linear growth of $h(\mathcal{A}_\epsilon \mathcal{P}^k f_\epsilon)$ will be observed for longer. The same arguments hold for the discrete versions

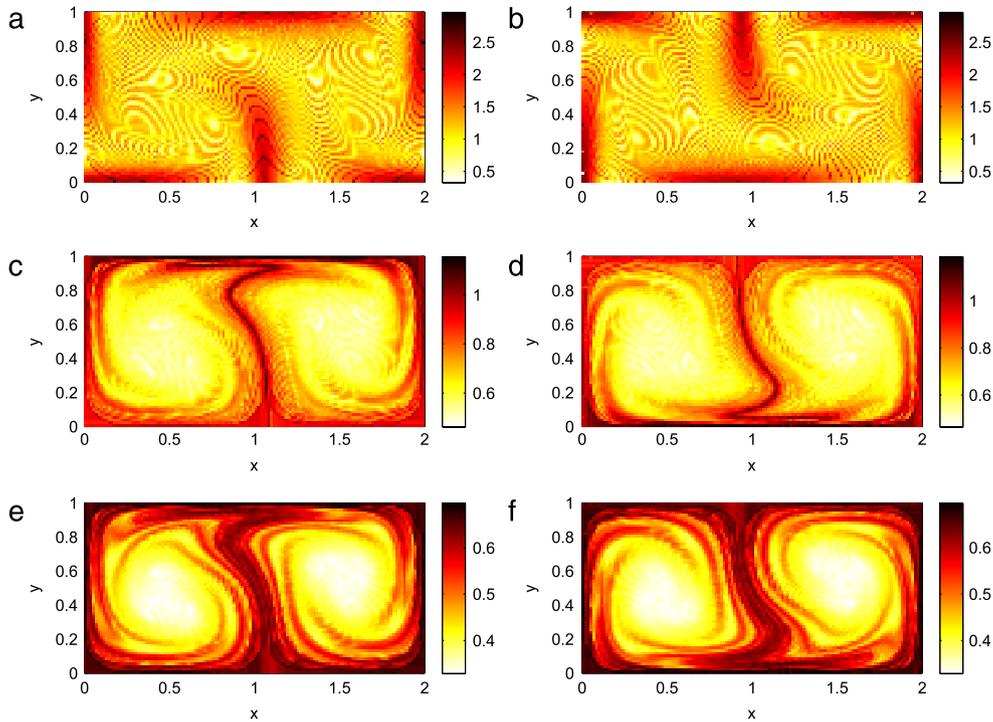


Fig. 5. FTE fields for the double gyre flow. The domain is partitioned in 8192 boxes and $R = 100$ inner grid points are iterated by the flow map T to obtain P and the backward-time approximation \tilde{P} . (a) $FTE(\cdot, 1)$ (b) $FTE(\cdot, 1)$; (c) $FTE_{\text{mult}}(\cdot, 5)$; (d) $FTE_{\text{mult}}(\cdot, 5)$; (e) $FTE_{\text{mult}}(\cdot, 10)$ (f) $FTE_{\text{mult}}(\cdot, 10)$.

of FTE. In discrete approximations to both FTLEs and FTE one may monitor how long one observes exponential growth (see e.g. [24, p. 50ff], and references therein), but we have not pursued this in this paper. An advantage of FTE_{mult} is that many values of k may be tested extremely cheaply using a single transition matrix.

For all our computations we use the software packages GAIO [32,33] and MATLAB.

6.1. Double gyre flow

We return to the famous double gyre test case as introduced in Example 3.5. Throughout we will fix $t_0 = 0$ and parameters as described above and consider the system on the domain $X = [0, 2] \times [0, 1]$. For this choice the time-1 flow map $T : X \rightarrow X$ (corresponding to the Poincaré return map on time-expanded space at time slice $t = 0$) has, in particular, two nontrivial unstable fixed points at $x_0 \approx (1.0808, 0)$ and $x_1 \approx (0.9192, 1)$. The stable manifold of x_0 and the unstable manifold of x_1 form a striking heteroclinic tangle, which has been the focus of many computational studies [10,35,23] and which we expect to pick up by our new FTE approach.

To create a numerical approximation of the transfer operator the domain X is partitioned in $n = 8192 = 2^{13}$ boxes. We estimate the entries of the transition matrix P for the flow map T by numerically integrating the system (12) on $[0, 1]$ with respect to 100 inner grid points per box using a Runge–Kutta scheme with constant step size 0.01. From P we form $\tilde{P} \approx P^{(-1)}$ for an approximation of the time-reversed dynamics. In Fig. 5 we show the results of the approximate FTE-field computations for different powers k of the one-step transition matrices P and \tilde{P} . While for $k = 1$ neighbourhoods of the local stable and unstable manifolds of the unstable fixed points are highlighted, $FTE_{\text{mult}}(\cdot, k)$ for $k = 5$ and $k = 10$ highlight increasingly large parts of the global manifolds. The results are very similar to results of the FTLE computations \mathcal{E}_5 and \mathcal{E}_{10} ; see Fig. 6.

To demonstrate the efficiency of our approach we compare the CPU times of the different computations. While the computation

of the transition matrix (8192 boxes, 100 test points per box) for a time-1 flow map is computationally demanding with about 110 s, the $FTE_{\text{mult}}(\cdot, 5)$ -field takes then less than a second and $FTE_{\text{mult}}(\cdot, 10)$ about 5 s, the respective backward time computations to obtain the FTE_{mult} -fields take about the same time. The FTLE computations based on solving the variational equation need about 22 s for the time-5 map and 40 s for the time-10 map. For this FTLE run-time test only the centre point in each box is considered, resulting in very spurious FTLE fields (not shown). Averaging over the FTLE values over four points per box takes already as much as 85 s for the time-5 and 168 s for the time-10 map (see Fig. 6(a), (b)) giving qualitatively similar results to the FTE computations. To sum up, we argue that *once* the transition matrix is set up *all* the respective FTE-fields of interest (forward in time, backward in time, and for any duration that is an integer multiple of *basetime*) can be approximated very efficiently.

To study the influence of the number of test points and the integration time we also computed the transition matrix $P^{(5)}$ from T^5 using $R = 100$ and $R = 400$ inner grid points per box. Visually the resulting FTE fields $FTE(\cdot, 5)$ as well as $FTE_{\text{mult}}(\cdot, 10, 5)$ agree very well; see Fig. 7. However due to the sparser sampling the $R = 100$ test point computations (Fig. 7(b) and (d)) achieve slightly smaller FTE values than the higher resolution computations (Fig. 7(a) and (c)).

Qualitatively the results also match nicely the corresponding approximate FTE fields in Fig. 5(c) and (e) respectively. Because of the numerical diffusion the FTE_{mult} values obtained by using powers of P are larger than the ones based on $P^{(5)}$ and are very similar to the respective FTLE values in Fig. 6.

We also studied the error in the backward-time FTE-field induced by using \tilde{P} as an approximation for $P^{(-1)}$. With the matrices P and $P^{(-1)}$ computed using only $R = 100$ test points per box, we observe that the relative error for $FTE(B_i, 1)$, $i = 1, \dots, 8192$ is about 2% on average and there are only a few boxes with very high error. But the errors drops quickly increasing k , for instance for $FTE_{\text{mult}}(\cdot, 10)$ the errors takes values smaller than 0.1% altogether. Also an increase in the number of test points for

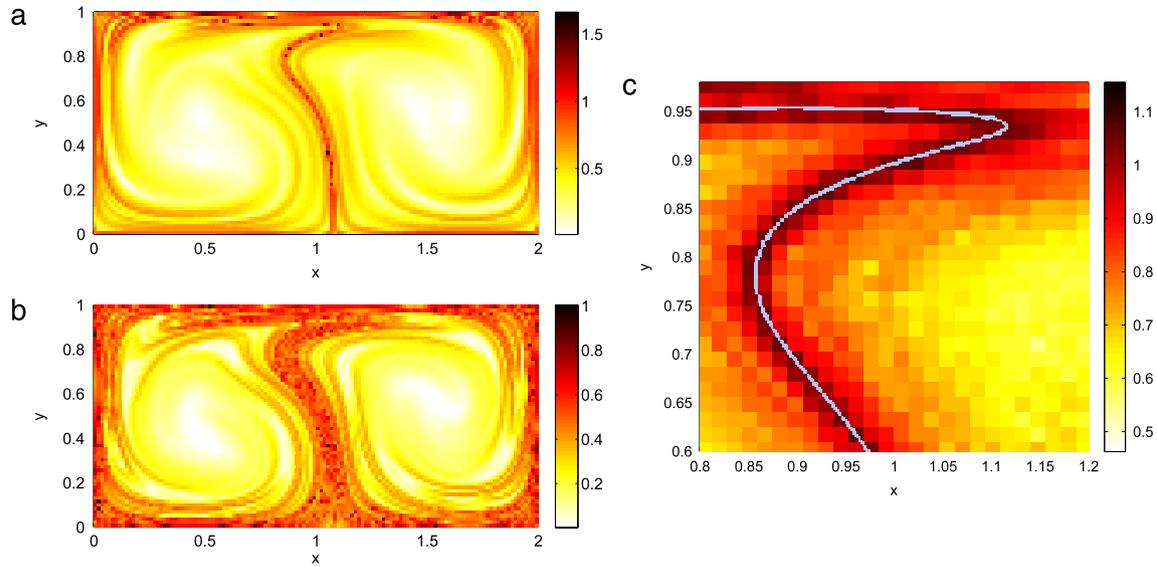


Fig. 6. Finite-time Lyapunov exponents for double gyre system. (a) FTLE values per box $\mathcal{E}_5(B_i)$, $i = 1, \dots, 8192$ obtained as average FTLE value over 4 inner grid points per box; (b) same as (a) but \mathcal{E}_{10} . (c) Ridges of a very fine representation of the FTLE-field \mathcal{E}_5 (light coloured; based on a box covering of 2^{19} boxes) match the structures highlighted by the $FTE_{\text{mult}}(\cdot, 5)$ -field from Fig. 5(c). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

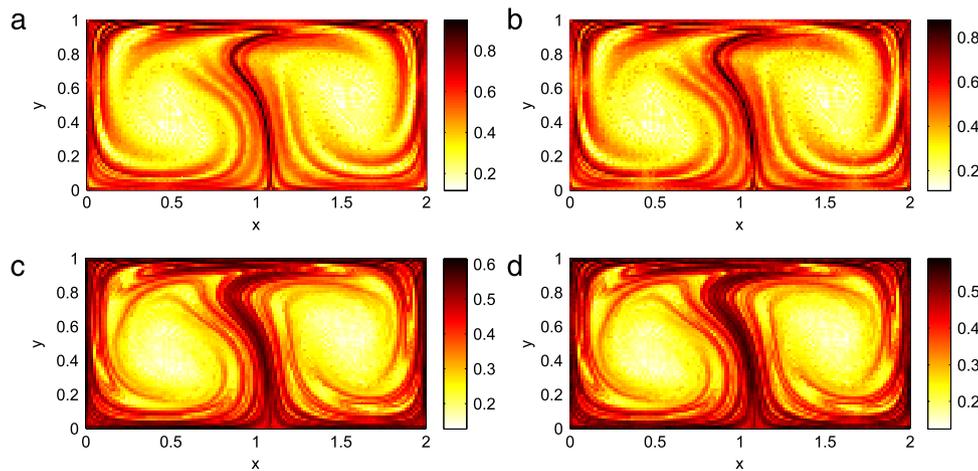


Fig. 7. Influence of number of test points for the map T^5 of the double gyre system. (a) $FTE(\cdot, 5)$ based on transition matrix computed using $R = 400$ inner grid points per box; (b) $FTE(\cdot, 5)$ based on transition matrix computed using $R = 100$ inner grid points per box; (c) $FTE_{\text{mult}}(\cdot, 10, 5)$, same set-up as (a); (d) $FTE_{\text{mult}}(\cdot, 10, 5)$, same as set-up as (b).

computing P helps reducing the error. Overall we can conclude, that in this test case, the backward-time approximation of the dynamics using \tilde{P} gives reliable results.

In Section 5.3 we have introduced how FTE computations can be extended to discretisations with differing box volumes. This allows us to develop adaptive algorithms for the refinement of those boxes that feature high FTE values. Fig. 8 shows the result of such a computation. We started with a box covering consisting of $2^{11} = 2048$ boxes and successively refined those boxes B_i , $i = 1, \dots, n$, for which $FTE_{\text{mult}}(B_i, 5) > 0.5(\frac{1}{n} \sum_i FTE_{\text{mult}}(B_i, 5) + \max_i FTE_{\text{mult}}(B_i, 5))$, where n denotes the current number of boxes in the covering. This rule ensures that the threshold for box refinement is above the arithmetic average and below the maximum. After six steps we obtain an FTE-field with respect to 8938 boxes, which has a high resolution in the vicinity of the stable manifold of x_0 (see Fig. 8(c) for an enlarged view) and low resolution where FTE values are low. Note that a regular covering would consist of as many as 131 072 boxes.

Finally let us comment on the artefacts observed in Fig. 5(a) and (b). Here the FTE field displays little “ripples”. This effect is related to the computations in Example 3.4. The formula (9) is

for the case where the thin image of a box fits within the original box partition elements in the contracting direction. The formula (10) is for the case where the box image straddles the original box partition elements in the contracting direction, leading to a higher FTE value. For the map in Example 3.4, on average, every $1/\beta$ horizontal strip will have the higher than usual $FTE(B_i, 1)$ values, leading to horizontal “ripples”. For general maps, the structure of these ripples will depend on the interplay between the map T and the grid. As k increases, the ripples disappear because (i) when using $FTE(\cdot, k)$ the ripple-causing terms in (10) decay to zero, and (ii) when using $FTE_{\text{mult}}(\cdot, k)$ as the images of boxes become longer and less structurally aligned with the grid, it is more likely that the FTE computation picks up a smoothing combination of terms from (9) and (10). Usually one is not interested in computing FTE for very small k , however one way to reduce this ripple effect for small k is to estimate the discrete entropy of a box image in a spatially consistent way. For example, one can align the image grid so that the image of the original box centre is a centre point of an image grid box.

In Fig. 9(a) we show the result of such an image grid adapted computation. The ripples are largely removed and the resulting

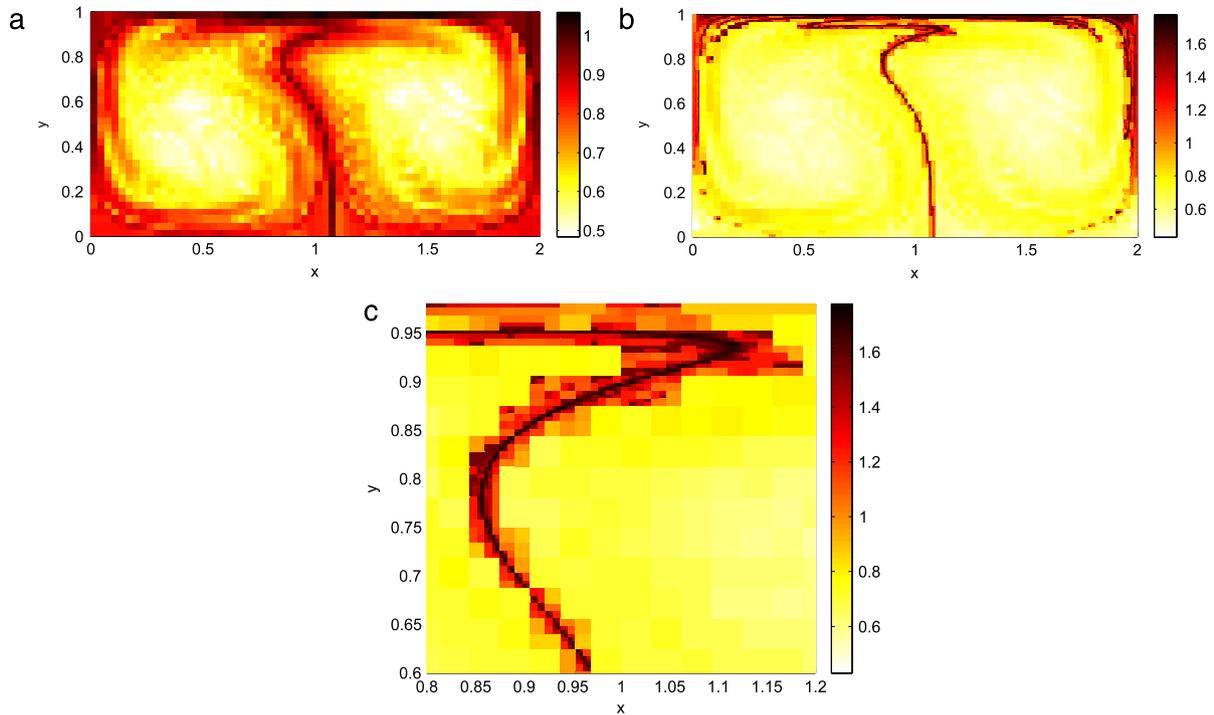


Fig. 8. Adaptive refinement in regions with high FTE—covering with boxes of differing volumes is obtained. The transition matrix P is obtained by integrating the time-1 flow map for $R = 100$ inner grid points per box. We consider the $FTE_{\text{mult}}(\cdot, 5)$ fields. Boxes B_i are refined when $FTE_{\text{mult}}(B_i, 5) > 0.5(\frac{1}{n} \sum_i FTE_{\text{mult}}(B_i, 5) + \max_i FTE_{\text{mult}}(B_i, 5))$, where n denotes the current number of boxes in the covering. (a) Initial covering with 2048 boxes. (b) Final covering after six steps with 8938 boxes (regular covering would consist of 131072 boxes). (c) Enlarged view of highlighted parts of (b).

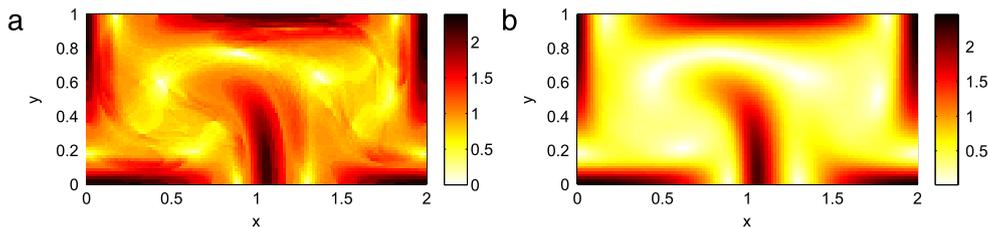


Fig. 9. (a) Reducing “ripples” due to occasional special alignment of the image densities with the box grid set as visible in Fig. 5(a). Here we adapt the image boxes in such a way that the image of a box centre will be the centre of a box again. (b) FTLE field $\mathcal{E}_1(B_i)$, $i = 1, \dots, 8192$ obtained as averaging the FTLE values over $R = 100$ inner grid points per box.

FTE field is similar to the respective FTLE computation shown in Fig. 9(b).

6.2. Lorenz flow

The Lorenz flow [36]

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\beta z + xy \end{aligned} \quad (23)$$

with classical parameters $\sigma = 10$, $\beta = \frac{8}{3}$, $\rho = 28$, is known to possess a chaotic attractor Λ with an SBR (Sinai–Bowen–Ruelle) measure μ [37].

First, we approximate the chaotic attractor observed in the Lorenz system using a set-oriented subdivision scheme; for details on such computations we refer to [31–33]. We obtain an approximation of the attractor Λ consisting of $n = 96\,244$ equally sized boxes. These boxes will also be the basis for approximating the transfer operator. For this we consider the time-1 map for $R = 1000$ inner grid points per box to estimate P . Our Runge–Kutta scheme constant step size is 0.01. We then compute the FTE fields using $k = 1$ and $k = 2$, i.e. $FTE(\cdot, 1)$ and $FTE_{\text{mult}}(\cdot, 2)$. As the

maximum side length of each box is 0.2344, we are mimicking $\epsilon = 0.1172$ in the definition of FTE_ϵ . The resulting fields are shown in Fig. 10. The FTE fields strikingly highlight regions which are related to the two-dimensional stable manifold of the hyperbolic fixed point at the origin, i.e. high values of FTE are observed where the manifold would intersect the numerically approximated attractor. The influence of this invariant manifold is also known to determine the formation of almost-invariant sets; see [23] for more details. Note that the corresponding FTLE field \mathcal{E}_1 obtained by considering the average FTLE value over $R = 1000$ inner grid points per box gives very comparable results (Fig. 10(c)). In Fig. 10(d) an approximation of the $FTE(\cdot, 1)$ field by $FTE_{\text{mult}}(\cdot, 1, 0.2)$ is shown. Here the transition matrix $P^{(0.2)}$ results from considering the time-0.2 flow map. The corresponding FTE field is affected by the numerical diffusion but still picks up nicely the relevant structures.

6.3. ABC flow

To demonstrate the applicability of our methodology in three dimensions we consider the following system of ordinary differential equations:

$$\dot{x} = A \sin z + C \cos y$$

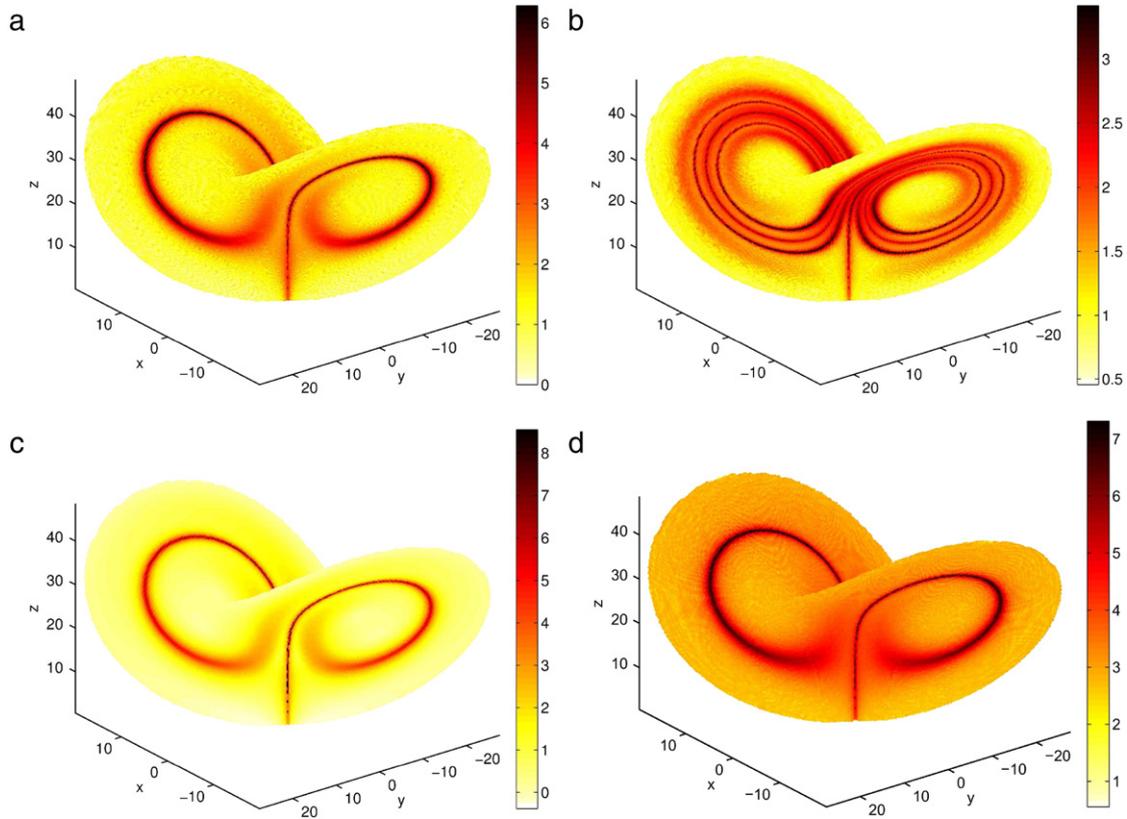


Fig. 10. FTE fields with respect to an approximation of the chaotic attractor in the Lorenz system and comparison with FTLE computation. The regions of high stretching are related to the two-dimensional stable manifold of the hyperbolic fixed point at the origin. (a) $FTE(\cdot, 1)$; (b) $FTE_{\text{mult}}(\cdot, 2)$; (c) FTLE field ε_1 , computed as the average FTLE with respect to 1000 inner grid points per box for the time-1 map (compare to (a)); (d) $FTE_{\text{mult}}(\cdot, 1, 0.2)$.

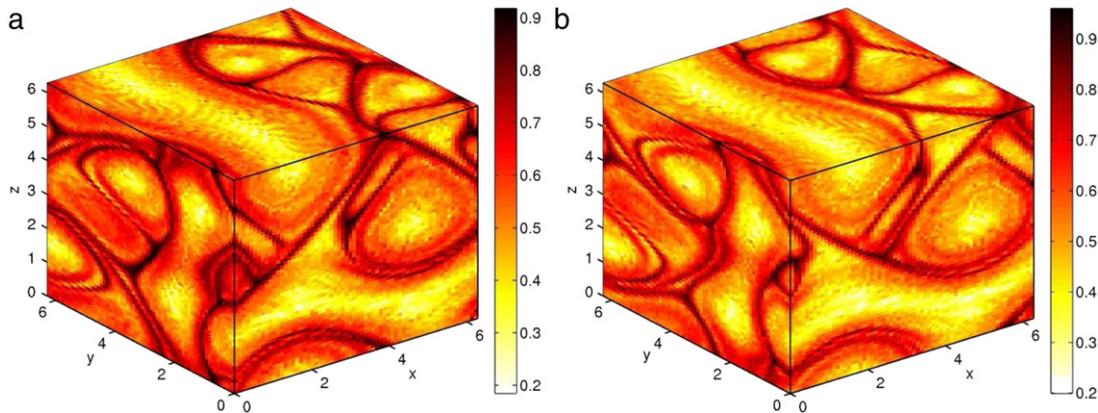


Fig. 11. FTE fields for ABC flow. (a) $FTE(\cdot, 5)$; (b) $\widetilde{FTE}(\cdot, 5)$.

$$\begin{aligned} \dot{y} &= B \sin x + A \cos z \\ \dot{z} &= C \sin y + B \cos x. \end{aligned} \tag{24}$$

This class of flows is known as ABC (Arnold-Beltrami-Childress) flows and the system (24) is notable for being an exact steady solution of Euler’s equation, exhibiting a nontrivial streamline geometry; see [8,38]. We consider the ABC flow on the torus \mathbb{T}^3 , i.e. $0 \leq x < 2\pi$, $0 \leq y < 2\pi$, $0 \leq z < 2\pi$ in (24) and choose parameters $A = \sqrt{3}$, $B = \sqrt{2}$, $C = 1$. For the numerical approximation of the transfer operator we partition $M = [0, 2\pi]^3$ into $n = 262\,144 = 2^{18}$ boxes and compute the transition matrix $P^{(5)}$ via the time-5 flow map using a classical Runge–Kutta scheme with constant step size $h = 0.01$. Each box is sampled uniformly with 1000 inner grid points. The respective $FTE(\cdot, 5)$ -field as well as an approximation to the backward-time field $\widetilde{FTE}(\cdot, 5)$ are shown in Fig. 11. As the side

length of each box is 0.0982, we are mimicking $\epsilon = 0.0491$ in the definition of FTE_ϵ . Here two-dimensional stable and unstable manifolds of hyperbolic periodic orbits are highlighted (cf. [23, Fig. 5]).

6.4. Idealised stratospheric flow

The quasi-periodically forced flow system as discussed in [39] is given by

$$\begin{aligned} \frac{dx}{dt} &= -\frac{\partial \Psi}{\partial y} \\ \frac{dy}{dt} &= \frac{\partial \Psi}{\partial x} \end{aligned}$$

with streamfunction

$$\Psi(x, y, t) = c_3 y - U_0 L \tanh(y/L) + A_3 U_0 L \operatorname{sech}^2(y/L) \cos(k_1 x)$$

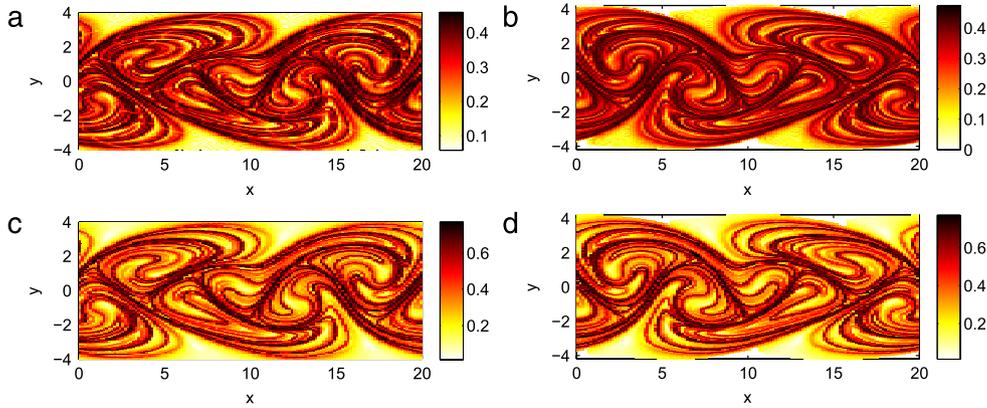


Fig. 12. FTE fields for the idealised stratospheric flow. (a) $FTE(\cdot, 10, 10)$; (b) $\widetilde{FTE}(\cdot, 10, 10) \approx FTE(\cdot, 20, -10)$; (c) approximation of FTLE field for same time span [10, 20] (d) approximation of FTLE field for backward time span [20, 10].

$$+ A_2 U_0 L \operatorname{sech}^2(y/L) \cos(k_2 x - \sigma_2 t) \\ + A_1 U_0 L \operatorname{sech}^2(y/L) \cos(k_1 x - \sigma_1 t).$$

We use the parameter values as in [22], i.e. $U_0 = 5.41$, $A_1 = 0.075$, $A_2 = 0.4$, $A_3 = 0.2$, $L = 1.770$, $c_2/U_0 = 0.205$, $c_3/U_0 = 0.7$, $k_1 = 2/r_e$, $k_2 = 4/r_e$, $k_3 = 6/r_e$ where $r_e = 6.371$ as well as $\sigma_2 = k_2(c_2 - c_3)$, $\sigma_1 = \sigma_2(1 + \sqrt{5})/2$, where we have dropped the physical units for brevity. The physical assumptions underlying the model equations and the parameters are described in detail in [39].

For the estimation of $P^{(10)}(10)$ we consider the flow map $\Phi(10, 10, \cdot)$ and a partition of the set $S^1 \times [-4, 4]$ is a circle parameterised from 0 to 20, in $m = 16384$ sets B_i , $i = 1, \dots, m$. We integrate the flow map for 100 inner grid points per box, choose a covering of the image by $n = 15763$ boxes C_j , $j = 1, \dots, n$, in $S^1 \times [-4.2, 4.2]$ and approximate the respective transition probabilities. We compute the FTE-field $FTE(\cdot, 10, 10)$ using the rectangular transition matrix $P^{(10)}(10) \in \mathbb{R}^{16384, 15763}$, see Fig. 12(a). Moreover we consider an approximation to the backward-time field $FTE(\cdot, 20, -10)$ via $\widetilde{FTE}(\cdot, 10, 10)$ as outlined in Section 5.2 (Fig. 12(b)). The results compare nicely to the respective approximate finite-time Lyapunov fields obtained via a relative dispersion approach; see [24] for details on such a computation.

6.5. Stratospheric polar vortex

For the numerical analysis of transport related to the stratospheric polar vortex we use two-dimensional velocity data from the ECMWF Interim data set.³ The global ECMWF data is given at a temporal resolution of 6 h and a spatial resolution of a 121×240 grid in longitude and latitude directions respectively. We consider the flow over a two-week period on a 475 K isentropic surface. As in [22] we focus on the stratosphere over the southern hemisphere south of 30 degrees latitude. There are strong persistent transport barriers in that region that determine the Antarctic polar vortex. These transport barriers have been studied using FTLE-related stretching diagnostics (finite-size Lyapunov exponents—FSLE) [40,41]; the FSLE fields were found to highlight an area (the so-called “surf zone”) outside the vortex boundary as determined by the maximum gradient in potential vorticity. We expect the FTE approach to highlight this zone just beyond the vortex boundary in a similar way.

We use the same setup as in [22], to which we refer the reader for more details. We start with an initial covering of $X = S^1 \times [-90^\circ, -30^\circ]$, where S^1 is a circle parameterised from 0° to 360° ,

by $m = 13471$ boxes B_i , $i = 1, \dots, m$ of approximately equal volume. Using 100 uniformly distributed sample points in each box the approximate image $\Phi(X, t, \tau)$ is computed and covered with $n = 14395$ boxes C_j , $j = 1, \dots, n$. The initial time t is chosen to be September 1, 2008 and the flow time τ as 14 days. For the numerical approximation of the flow a Runge–Kutta scheme with step size 45 min is used, with linear interpolation in space and time between data points. We construct $P := P^{(t)}(\tau)$ using the same sample points. From this matrix we compute the forward time FTE-field $FTE(\cdot, t, \tau)$, see Fig. 13(a). In addition we construct the approximate-backward time transition matrix \tilde{P} and obtain $\widetilde{FTE}(\cdot, t, \tau)$ as an approximation to $FTE(\cdot, t + \tau, -\tau)$; see Fig. 13(b). Regions of high value in the resulting FTE-fields indicate the existence of isolated transport barriers. In [22] an approximation of the polar vortex as a finite-time coherent set is obtained using the same transition matrix as our FTE calculations. As expected, some of the structures highlighted in Fig. 13 by the FTE approach delineate a region that contains the polar vortex but do not provide a tight boundary (see Fig. 3(c) and (d) in [22] where it is found that the maximally coherent sets (in red) on the 475 K isentropic surface almost exactly coincide with the maximum PV gradient curve (in green)).

7. Discussion and conclusion

We have introduced a new stretching diagnostic, finite-time entropy, that is applicable to autonomous and nonautonomous dynamical systems. It considers the growth rate of uncertainty generated by an initial small random ϵ -perturbation of the original nonlinear dynamics, followed by a final small ϵ -perturbation. This uncertainty growth is captured by measuring the entropy increase experienced by a localised probability density of resolution $\epsilon > 0$ under the action of the perturbed dynamical system. We have obtained (i) a nonlinear probability-based estimate of stretching for every ϵ -resolution, (ii) robustness w.r.t. the ϵ -perturbation size, with a convergence result in the zero-perturbation limit, and (iii) continuity of the FTE field for every scale $\epsilon > 0$.

FTE fits very well into the set-oriented numerical framework for the global analysis of dynamical systems. This allows for the development of extremely simple and numerically efficient methods of constructing an estimate of the FTE field. Within this framework the FTE field is instantaneously calculable from a sparse transition matrix generated from the dynamics. In particular, in the context of autonomous systems it suffices to set up a transition matrix using short-time integrations only, and to compute results for longer time spans by sparse matrix multiplication. The set-oriented framework also enables adaptive

³ <http://data.ecmwf.int/data/index.html>.

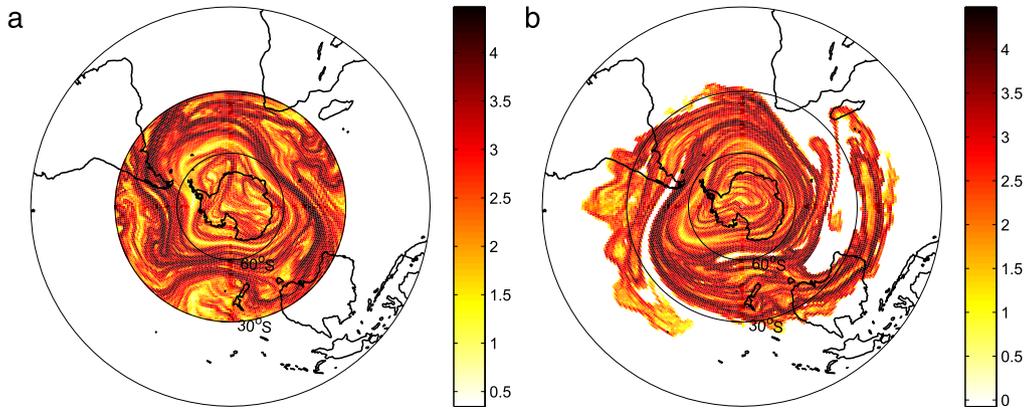


Fig. 13. FTE fields in the region of the Antarctic polar vortex. (a) Forward-time field $FTE(\cdot, t, \tau)$ ($t = \text{September 1, 2008}$, $\tau = 14$ days) indicates repelling transport barriers; (b) backward-time field $FTE(\cdot, t, \tau) \approx FTE(\cdot, t + \tau, -\tau)$ highlights attracting transport barriers.

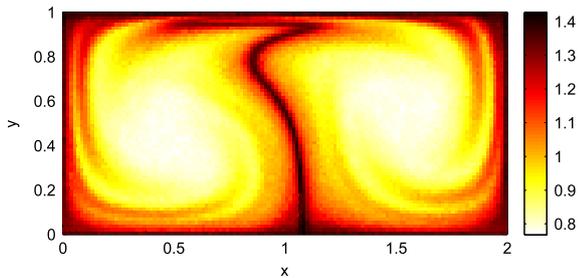


Fig. 14. $FTE_{\text{mult}}(\cdot, 5)$ -field for a stochastic double gyre flow $dX_t = g(X_t, t)dt + \delta dW_t$, with g corresponding to the right-hand side of (12) and $\delta = 0.001$. For this choice of δ the stochastic noise dominates the ε -perturbation used in the FTE approach. For the approximation of the transition matrix (2^{13} boxes as above) we consider the time-1 flow map with periodic boundary conditions, approximated by an Euler–Maruyama scheme with stepsize $h = 0.001$.

algorithms to be employed to more quickly resolve high-value FTE regions.

We have demonstrated that the FTE approach reliably and efficiently measures stretching in a completely derivative-free way and is capable of extracting the backward-time dynamics from the forward-time transition matrix *without* any additional trajectory integration.

For the latter we note that there is a recent related result to compute backward-time FTLE fields from the forward-time information [42]. However, from the computational perspective the backward-time FTLE method of [42] is currently confined to two-dimensional flows. Our approach only manipulates a sparse transition matrix—so the time-reversing operation is independent of the dimension of phase space.

The results of FTE computations in our examples are very similar to the corresponding FTLE fields because we focussed on deterministic dynamical systems. The FTE approach may also easily be applied to systems with both deterministic and stochastic parts by replacing the Perron–Frobenius operator with a noisy Perron–Frobenius operator or general Markov operator; see Fig. 14 for an example computation and [27] for related work. In particular, when estimating FTLE and FTE from an embedded time series we would expect the FTE approach to be more robust with respect to noise because we do not need to estimate derivatives but work directly with probability flow. We will pursue these aspects further in future work.

Acknowledgements

The authors thank Naratip Santitissadeekorn for providing the transition matrix and related code for the polar vortex example. KPG is grateful to the School of Mathematics and Statistics at the

University of New South Wales for their hospitality. The research of GF has been partially supported by an ARC Discovery Project (DP110100068).

Appendix. Proof of existence of $FTE(x)$

The norm $\|\cdot\|$ is the $L^1(m)$ norm throughout and m is non-normalised Lebesgue measure. We present Lemma A.1 here as we have been unable to find the result in the literature. Some arguments in the proof have been drawn from the proof of Theorem 9.3.2 in [28].

Lemma A.1. *Let $m(X) < \infty$ and let $\{f_n\}, f_n : X \rightarrow \mathbb{R}^+$ be a sequence of uniformly essentially bounded densities ($\|f_n\|_{L^\infty} \leq \Lambda' \forall n \geq 0$, $\Lambda' < \infty$) converging in L^1 to an essentially bounded density f . Then $h(f_n) \rightarrow h(f)$.*

Proof. The overall strategy is to break all functions f_n, f into two pieces: one with low values and one with high values; the latter will carry the bulk of the L^1 norms. In part 1 of the proof we show that the entropy of the first pieces is small for f_n and f separately. In part 2 of the proof we estimate the entropy difference of the remainder; the individual entropies are nontrivial, but their difference will be small.

Part 1: For some $0 < \Lambda \leq \min\{1, \Lambda'\} < \infty$ essentially decompose $f_n = f_n^1 + f_n^2$ where

$$f_n^1(x) = \begin{cases} f_n(x), & \text{if } f_n(x) < \Lambda; \\ 0, & \text{otherwise.} \end{cases}$$

$$f_n^2(x) = \begin{cases} f_n(x), & \text{if } \Lambda \leq f_n(x) \leq \Lambda'; \\ 0, & \text{otherwise.} \end{cases}$$

Decompose $f = f^1 + f^2$ in exactly the same way using the same Λ, Λ' . Let $\lambda_n = \|f_n^1\|$, $\lambda = \|f^1\|$. Note that $\lambda_n \leq \|f_n^1 - f^1\| + \|f^1\| \leq \|f_n - f\| + \lambda$. Let $D = \{f \in L^1 : f \geq 0, \int f dm = 1\}$ be the space of densities. It is easy to check that $h(cf) = ch(f) - c \log c$ for $f \in D$ and $c \in \mathbb{R}^+$. Assume $\lambda_n > 0$. If $\lambda_n = 0$, then $h(f_n^1) = 0$ and there will be nothing to prove. We now bound $h(f_n^1)$ above and below in terms of λ_n . Denote the entropy function by $\eta(u) = -u \log u$, $u \geq 0$. First, since $\Lambda \leq 1$ and $\eta(u) \geq 0$ for $0 \leq u \leq 1$, one has $h(f_n^1) \geq 0$. Second, since $f_n^1/\lambda_n \in D$, one has $h(\lambda_n \cdot (f_n^1/\lambda_n)) = \lambda_n h(f_n^1/\lambda_n) - \lambda_n \log \lambda_n$. Noting that $h(f_n^1/\lambda_n) \leq \log m(X)$ (see Remark 9.2.1 in [28] or note that the maximum entropy occurs for uniform densities $f^* = 1/m(X)$ where $h(f^*) = \log m(X)$), one has $h(f_n^1) \leq \lambda_n \log m(X) - \lambda_n \log \lambda_n$. Thus, λ_n goes to zero, the (nonnegative) RHS goes to zero and so $|h(f_n^1)|$ can be made arbitrarily small by making λ_n sufficiently small. Completely analogously, one has $0 \leq |h(f^1)| \leq \lambda \log m(X) - \lambda \log \lambda$.

Part 2:

$$\begin{aligned}
 |h(f_n^2) - h(f^2)| &= \left| \int_{\text{supp} f_n^2} \eta(f_n^2) dm - \int_{\text{supp} f^2} \eta(f^2) dm \right| \\
 &= \left| \int_{\text{supp} f_n^2 \setminus \text{supp} f^2} \eta(f_n^2) dm \right. \\
 &\quad \left. + \int_{\text{supp} f^2 \cap \text{supp} f_n^2} \eta(f_n^2) - \eta(f^2) dm \right. \\
 &\quad \left. - \int_{\text{supp} f^2 \setminus \text{supp} f_n^2} \eta(f^2) dm \right|. \tag{25}
 \end{aligned}$$

To handle the first term above we note that

$$\begin{aligned}
 \Lambda \cdot m(\text{supp} f_n^2 \setminus \text{supp} f^2) &= \int_{\text{supp} f_n^2 \setminus \text{supp} f^2} \Lambda dm \\
 &\leq \int_{\text{supp} f_n^2 \setminus \text{supp} f^2} |f_n^2| dm \\
 &= \int_{\text{supp} f_n^2 \setminus \text{supp} f^2} |f_n^2 - f^2| dm \\
 &\leq \|f_n^2 - f^2\|. \tag{26}
 \end{aligned}$$

Thus, $m(\text{supp} f_n^2 \setminus \text{supp} f^2) \leq \|f_n^2 - f^2\|/\Lambda$. Now

$$\begin{aligned}
 \left| \int_{\text{supp} f_n^2 \setminus \text{supp} f^2} \eta(f_n^2) dm \right| &\leq \max_{\Lambda \leq u \leq \Lambda'} |\eta(u)| \cdot (\|f_n^2 - f^2\|/\Lambda) \\
 &=: S_{\Lambda, \Lambda'} \|f_n^2 - f^2\|. \tag{27}
 \end{aligned}$$

The third term in (25) is bounded similarly. For the second term in (25),

$$\begin{aligned}
 &\left| \int_{\text{supp} f^2 \cap \text{supp} f_n^2} \eta(f_n^2) - \eta(f^2) dm \right| \\
 &\leq \max_{\Lambda \leq u \leq \Lambda'} |\eta'(u)| \int |f_n^2 - f^2| dm \\
 &=: S_{\Lambda, \Lambda'} \|f_n^2 - f^2\|, \tag{28}
 \end{aligned}$$

where the inequality uses the mean value theorem applied pointwise on the integrand and $S_{\Lambda, \Lambda'} := \max_{\Lambda \leq u \leq \Lambda'} |\eta'(u)|$.

To finish note that as the supports of f_n^1 and f_n^2 are disjoint (as are those of f^1, f^2) we may write $|h(f_n) - h(f)| \leq |h(f_n^1)| + |h(f^1)| + |h(f_n^2) - h(f^2)|$.

Given $\varepsilon > 0$ we choose Λ small enough so that $|h(f^1)| \leq |\lambda(\Lambda) \log m(X) - \lambda(\Lambda) \log(\lambda(\Lambda))| < \varepsilon/6$. With Λ , and hence λ , now fixed, we find an N_1 large enough so that using the bound $\lambda_n \leq \lambda + \|f_n - f\|$ one has $|h(f_n^1)| \leq \lambda_n \log m(X) - \lambda_n \log \lambda_n < \varepsilon/4$ for $n \geq N_1$. Finally, we find an N_2 large enough so that one has $(2S_{\Lambda, \Lambda'}/\Lambda + S_{\Lambda, \Lambda'})\|f_n - f\| < \varepsilon/2$ for $n \geq N_2$. Thus, for $n \geq \max\{N_1, N_2\}$, $|h(f_n) - h(f)| < \varepsilon$. \square

Definition A.2. Denote the tangent space of $X \subset \mathbb{R}^d$ at $x \in X$ by $T_x X$; throughout, we will identify $T_x X$ with \mathbb{R}^d . For $x \in X$, define $\Phi_{x, \varepsilon} : B_\varepsilon(x) \rightarrow T_x X$ by $\Phi_{x, \varepsilon}(z) = (z - x)/\varepsilon$. The map $\Phi_{x, \varepsilon}$ is a “flat” version of the inverse exponential map followed by a linear expansion by factor $1/\varepsilon$. In particular, $\Phi_{x, \varepsilon} B_\varepsilon(x) = B_1(0) := \{y \in T_x X : \|y\| < 1\}$.

We now construct by conjugation with $\Phi_{x, \varepsilon}$ the “expanded” action of T on $T_x X$. Define $S_{x, \varepsilon} : T_x X \circlearrowleft$ as $S_{x, \varepsilon} = \Phi_{x, \varepsilon} \circ T \circ \Phi_{x, \varepsilon}^{-1}$.

This construction is repeated for Perron–Frobenius operators. First, we construct the Perron–Frobenius operator for $\Phi_{x, \varepsilon}$, $\mathcal{P}_{\Phi_{x, \varepsilon}} : L^1(X, m) \rightarrow L^1(T_x X, m)$ and then via conjugacy we form $\mathcal{P}_{S_{x, \varepsilon}} : L^1(T_x X, m) \circlearrowleft$. We also construct $\mathcal{A}_{1,0} : L^1(T_x X, m) \circlearrowleft$ as $\mathcal{P}_{\Phi_{x, \varepsilon}} \circ \mathcal{A}_{\varepsilon, x} \circ \mathcal{P}_{\Phi_{x, \varepsilon}}^{-1}$.

Lemma A.3. Let $f \in L^1(X, m)$. Then $h(\mathcal{P}_{\Phi_{x, \varepsilon}} f) = h(f) - \log \varepsilon^d \int f dm$.

Proof.

$$\begin{aligned}
 h(\mathcal{P}_{\Phi_{x, \varepsilon}} f) &= - \int_{\text{supp} f \circ \Phi_{x, \varepsilon}^{-1}} \varepsilon^d f \circ \Phi_{x, \varepsilon}^{-1} \log(\varepsilon^d f \circ \Phi_{x, \varepsilon}^{-1}) dm \\
 &= - \int_{\text{supp} f \circ \Phi_{x, \varepsilon}^{-1}} \varepsilon^d f(x + \varepsilon w) \log(\varepsilon^d f(x + \varepsilon w)) dw \\
 &= - \int_{\text{supp} \Phi^{-1}(\text{supp} f \circ \Phi_{x, \varepsilon}^{-1})} f(z) \log(\varepsilon^d f(z)) dz \\
 &\quad \text{letting } z = \Phi^{-1}(w) = x + \varepsilon w \\
 &= - \int_{\text{supp} f} f(z) \log f(z) dz - \log \varepsilon^d \int f dz. \quad \square
 \end{aligned}$$

Definition A.4. Define an affine approximation to T on $B_\varepsilon(x)$ as $T_x^A(y) = T(x) + DT(x) \cdot (y - x)$. Via conjugation with $\Phi_{x, \varepsilon}$, define $S_x^A(y) = \Phi_{x, \varepsilon} \circ T_x^A \circ \Phi_{x, \varepsilon}^{-1}$.

Lemma A.5. Given $\varepsilon_0 > 0$ there is a constant $C = C(x, \varepsilon_0) < \infty$ so that $\|S_{x, \varepsilon}(y) - S_x^A(y)\| \leq C\varepsilon$ for all $0 < \varepsilon < \varepsilon_0$ and all $y \in B_1(0)$.

Proof. By Taylor’s theorem, $\|T(z) - T_x^A(z)\| \leq C(x, \varepsilon_0)\|z - x\|^2$. In particular, if $z \in B_\varepsilon(x)$, then $\|T(z) - T_x^A(z)\| \leq C\varepsilon^2$. We now have

$$\begin{aligned}
 \|S_{x, \varepsilon}(y) - S_x^A(y)\| &= \|\Phi_{x, \varepsilon} \circ T \circ \Phi_{x, \varepsilon}^{-1}(y) - \Phi_{x, \varepsilon} \circ T_x^A \circ \Phi_{x, \varepsilon}^{-1}(y)\| \\
 &= \|(T(x + \varepsilon y) - x)/\varepsilon - (T_x^A(x + \varepsilon y) - x)/\varepsilon\| \\
 &\leq C\varepsilon \|y\|. \quad \square
 \end{aligned}$$

Lemma A.6. Given $\varepsilon_0 > 0$ there is a constant $C = C(x, \varepsilon_0) < \infty$ so that $\|DS_{x, \varepsilon}(y) - DS_x^A(y)\| \leq C\varepsilon$ for all $0 < \varepsilon < \varepsilon_0$ and all $y \in B_1(0)$.

Proof. Note that $DS(y) = D(\Phi_{x, \varepsilon} \circ T \circ \Phi_{x, \varepsilon}^{-1})(y) = DT(\Phi_{x, \varepsilon}^{-1}(y))$ by the chain rule and $D\Phi_{x, \varepsilon} \equiv \varepsilon \cdot \text{Id}$. Now,

$$\|DS_{x, \varepsilon}(y) - DS_x^A(y)\| = \|DT(\Phi_{x, \varepsilon}^{-1}(y)) - DT_x^A(\Phi_{x, \varepsilon}^{-1}(y))\|$$

and if $y \in B_1(0)$ then $\Phi_{x, \varepsilon}^{-1}(y) \in B_\varepsilon(x)$. By Taylor’s theorem, there exists a $C = C(x, \varepsilon_0)$ such that $\|DT(\Phi_{x, \varepsilon}^{-1}(y)) - DT_x^A(\Phi_{x, \varepsilon}^{-1}(y))\| \leq C\varepsilon$. \square

Lemma A.7.

$$FTE_\varepsilon(x, 1) - FTE_\varepsilon^A(x, 1) = h(\mathcal{A}_1 \mathcal{P}_{S_{x, \varepsilon}} f_{0,1}) - h(\mathcal{A}_1 \mathcal{P}_{S_x^A} f_{0,1}),$$

where $f_{0,1} = 1/m(B_1(0)) \mathbf{1}_{B_1(0)}$.

Proof.

$$\begin{aligned}
 FTE_\varepsilon(x, 1) - FTE_\varepsilon^A(x, 1) &= [h(\mathcal{A}_\varepsilon \mathcal{P}_T f_{x, \varepsilon}) - h(f_{x, \varepsilon})] - [h(\mathcal{A}_\varepsilon \mathcal{P}_{T_x^A} f_{x, \varepsilon}) - h(f_{x, \varepsilon})] \\
 &= h(\mathcal{P}_{\Phi_{x, \varepsilon}} \mathcal{A}_\varepsilon \mathcal{P}_T f_{x, \varepsilon}) - h(\mathcal{P}_{\Phi_{x, \varepsilon}} \mathcal{A}_\varepsilon \mathcal{P}_{T_x^A} f_{x, \varepsilon}) \quad \text{by Lemma A.3} \\
 &= h(\mathcal{A}_1 \mathcal{P}_{S_{x, \varepsilon}} \mathcal{P}_{\Phi_{x, \varepsilon}} f_{x, \varepsilon}) - h(\mathcal{A}_1 \mathcal{P}_{S_x^A} \mathcal{P}_{\Phi_{x, \varepsilon}} f_{x, \varepsilon}) \\
 &\quad \text{by defn of } \mathcal{A}_1, \mathcal{P}_{S_{x, \varepsilon}}, \mathcal{P}_{S_x^A} \\
 &= h(\mathcal{A}_1 \mathcal{P}_{S_{x, \varepsilon}} f_{0,1}) - h(\mathcal{A}_1 \mathcal{P}_{S_x^A} f_{0,1}). \quad \square
 \end{aligned}$$

Lemma A.8. Let T be C^2 and $|\det DT|$ be uniformly bounded away from zero. Then $\|\mathcal{P}_{S_{x,\epsilon}}f_{0,1} - \mathcal{P}_{S_x^A}f_{0,1}\| \rightarrow 0$ as $\epsilon \rightarrow 0$.

Proof. We write

$$\begin{aligned} & \int |\mathcal{P}_{S_{x,\epsilon}}f_{0,1} - \mathcal{P}_{S_x^A}f_{0,1}| \, dm \\ &= \int \left| \left(\frac{f_{0,1}}{|\det DS_{x,\epsilon}|} - \frac{f_{0,1}}{|\det DS_x^A|} \right) \circ (S_x^A)^{-1} \right| \, dm \\ & \quad + \int \left| \frac{f_{0,1}}{|\det DS_x^A|} \circ S_{x,\epsilon}^{-1} - \frac{f_{0,1}}{|\det DS_x^A|} \circ (S_x^A)^{-1} \right| \, dm. \end{aligned}$$

Treating the first term: a change of variables yields

$$\begin{aligned} & \int \left| \left(\frac{f_{0,1}}{|\det DS_{x,\epsilon}|} - \frac{f_{0,1}}{|\det DS_x^A|} \right) \circ S_{x,\epsilon}^{-1} \right| \, dm \\ &= \int \left| f_{0,1} \left(1 - \frac{|\det DS_{x,\epsilon}|}{|\det DS_x^A|} \right) \right| \, dm \\ &\leq \sup_{y \in B_1(0)} \left| 1 - \frac{|\det DS_{x,\epsilon}|}{|\det DS_x^A|} \right| \int f_{0,1} \, dm. \end{aligned}$$

Invoking Lemma A.6 we see this first term goes to zero as $\epsilon \rightarrow 0$ since $f_{0,1}$ is a density, and $|\det DS| = |\det DT(\Phi_{x,\epsilon}^{-1})|$ is uniformly bounded away from zero.

For the second term, we note that $f_{0,1}/|\det DS_x^A|$ is a constant function on its support, thus

$$\begin{aligned} & \int \left| \frac{f_{0,1}}{|\det DS_x^A|} \circ S_{x,\epsilon}^{-1} - \frac{f_{0,1}}{|\det DS_x^A|} \circ (S_x^A)^{-1} \right| \, dm \\ &= \frac{1}{m(B_1(0))|\det DS_x^A|} \int |\mathbf{1}_{B_1(0)} \circ S_{x,\epsilon}^{-1} - \mathbf{1}_{B_1(0)} \circ (S_x^A)^{-1}| \, dm \\ &= \frac{m(S_{x,\epsilon}B_1(0)\Delta S_x^AB_1(0))}{m(B_1(0))|\det DS_x^A|}, \end{aligned}$$

which tends to 0 as $\epsilon \rightarrow 0$ by the following Lemma A.9. \square

Lemma A.9. Assume $\det DT(x) \neq 0$. Then $m(S_{x,\epsilon}B_1(0)\Delta S_x^AB_1(0)) \rightarrow 0$ as $\epsilon \rightarrow 0$, where Δ denotes the symmetric difference.

Proof. As $\det DT(x) \neq 0$, and S_x^A is affine, one has that $S_x^AB_1(0)$ is an ellipsoid of full dimension. The set $S_x^AB_1(0)$ thus has a well-defined topological boundary, which is a connected $(d-1)$ -dimensional hypersurface of bounded $d-1$ -dimensional volume.

There is a constant $C = C(x, \epsilon_0)$ such that $\sup_{y \in B_1(0)} \|S_x^A(y) - S_{x,\epsilon}(y)\| \leq C\epsilon$ for all $0 < \epsilon < \epsilon_0$ by Lemma A.5. Define $\Theta_\epsilon = \{y \in T_{T(x)}X : d(y, \partial(S_x^AB_1(0))) \leq C\epsilon\}$. Clearly, $S_x^AB_1(0)\Delta S_{x,\epsilon}B_1(0) \subset \Theta_\epsilon$. Further, we may compute the volume of Θ_ϵ as $m(\Theta_\epsilon) = m_{d-1}(\partial S_x^AB_1(0)) \cdot 2C\epsilon$.

Thus $m(S_x^AB_1(0)\Delta S_{x,\epsilon}B_1(0)) \leq m(\Theta_\epsilon) = m_{d-1}(\partial S_x^AB_1(0)) \cdot 2C\epsilon$ approaches zero as linearly in ϵ . \square

Remark. Note that Lemma A.8 also applies to the discrete approximation induced by the Ulam construction: $f_\epsilon = \mathbf{1}_{B_{n,i}}/m(B_{n,i})$, a density supported on a single box on an n -grid, and $\mathcal{A}_\epsilon = \pi_n$, the canonical L^1 projection onto the span of $\mathbf{1}_{B_{n,i}}$, $i = 1, \dots, n$.

Existence of FTE. We conclude with proving the existence of finite-time entropy as $\epsilon \rightarrow 0$:

Proof of Theorem 2.7. We prove the theorem for $k = 1$; the proof for higher k follows in exactly the same way. By Lemma A.7 we see that the difference $|FTE_\epsilon(x, 1) - FTE_\epsilon^A(x, 1)|$ may be studied via the difference $|h(\mathcal{A}_1\mathcal{P}_{S_{x,\epsilon}}f_{0,1}) - h(\mathcal{A}_1\mathcal{P}_{S_x^A}f_{0,1})|$. The functions $\mathcal{A}_1\mathcal{P}_{S_{x,\epsilon}}f_{0,1}$, $\mathcal{A}_1\mathcal{P}_{S_x^A}f_{0,1}$ are essentially bounded, and so we may apply Lemma A.1 to conclude the result, once we show that $\|\mathcal{P}_{S_{x,\epsilon}}f_{0,1} - \mathcal{P}_{S_x^A}f_{0,1}\| \rightarrow 0$ (we may remove \mathcal{A}_1 since $\|\mathcal{A}_1\| \leq 1$). By Lemma A.8 we are done. \square

References

- [1] H. Aref, The development of chaotic advection, *Phys. Fluids* 14 (4) (2002) 1315–1325.
- [2] J.D. Meiss, Symplectic maps, variational principles, and transport, *Rev. Mod. Phys.* 64 (3) (1992) 795–848.
- [3] S. Wiggins, *Chaotic Transport in Dynamical Systems*, Springer, 1992.
- [4] S. Wiggins, The dynamical systems approach to Lagrangian transport in oceanic flows, *Annu. Rev. Fluid Mech.* 37 (2005) 295–328.
- [5] V. Rom-Kedar, A. Leonard, S. Wiggins, An analytical study of transport, mixing and chaos in an unsteady vortical flow, *J. Fluid Mech.* 214 (1990) 347–394.
- [6] V. Rom-Kedar, S. Wiggins, Transport in two-dimensional maps, *Arch. Ration. Mech. Anal.* 109 (1990) 239–298.
- [7] G. Haller, Finding finite-time invariant manifolds in two-dimensional velocity fields, *Chaos* 10 (2000) 99–108.
- [8] G. Haller, Distinguished material surfaces and coherent structures in three-dimensional fluid flows, *Physica D* 149 (2001) 248–277.
- [9] G. Haller, A variational theory of hyperbolic Lagrangian coherent structures, *Physica D* 240 (2011) 574–598.
- [10] S.C. Shadden, F. Lekien, J.E. Marsden, Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows, *Physica D* 212 (2005) 271–304.
- [11] D. Bohl, A. Mehta, N. Santitissadeekorn, E. Bollt, Characterization of mixing in a simple paddle mixer using experimentally derived velocity fields, *J. Fluids Eng.* 133 (6) (2011) 061202.
- [12] M. Dellnitz, O. Junge, On the approximation of complicated dynamical behaviour, *SIAM J. Numer. Anal.* 36 (2) (1999) 491–515.
- [13] M. Dellnitz, O. Junge, Almost invariant sets in Chua's circuit, *Internat. J. Bifur. Chaos* 7 (11) (1997) 2475–2485.
- [14] P. Deuffhard, W. Huisinga, A. Fischer, C. Schütte, Identification of almost invariant aggregates in nearly uncoupled Markov chains, *Linear Algebra Appl.* 315 (2000) 39–59.
- [15] C. Schütte, W. Huisinga, P. Deuffhard, Transfer operator approach to conformational dynamics in biomolecular systems, in: B. Fiedler (Ed.), *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, Springer, 2001, pp. 191–223.
- [16] G. Froyland, M. Dellnitz, Detecting and locating near-optimal almost-invariant sets and cycles, *SIAM J. Sci. Comput.* 24 (6) (2003) 1839–1863.
- [17] G. Froyland, Statistically optimal almost-invariant sets, *Physica D* 200 (2005) 205–219.
- [18] G. Froyland, Unwrapping eigenfunctions to discover the geometry of almost-invariant sets in hyperbolic maps, *Physica D* 237 (2008) 840–853.
- [19] G. Froyland, S. Lloyd, A. Quas, Coherent structures and isolated spectrum for Perron–Frobenius cocycles, *Ergodic Theory Dynam. Systems* 30 (2010) 729–756.
- [20] G. Froyland, S. Lloyd, N. Santitissadeekorn, Coherent sets for nonautonomous dynamical systems, *Physica D* 239 (2010) 1527–1541.
- [21] G. Froyland, N. Santitissadeekorn, A. Monahan, Optimally coherent sets in geophysical flows: A new approach to delimiting the stratospheric polar vortex, *Phys. Rev. E* 82 (2010) 056311.
- [22] G. Froyland, N. Santitissadeekorn, A. Monahan, Transport in time-dependent dynamical systems: finite-time coherent sets, *Chaos* 20 (2010) 043116.
- [23] G. Froyland, K. Padberg, Almost-invariant sets and invariant manifolds—connecting probabilistic and geometric descriptions of coherent structures in flows, *Physica D* 238 (2009) 1507–1523.
- [24] K. Padberg, Numerical analysis of transport in dynamical systems, Ph.D. Thesis, Universität Paderborn, Paderborn, 2005.
- [25] K. Padberg, B. Thiere, R. Preis, M. Dellnitz, Local expansion concepts for detecting transport barriers in dynamical systems, *Commun. Nonlinear Sci. Numer. Simul.* 14 (12) (2009) 4176–4190.
- [26] P. Tallapragada, S.D. Ross, A geometric and probabilistic description of coherent sets (2011) (submitted for publication).
- [27] D. Schneider, J. Fuhrmann, W. Reich, G. Scheuermann, A variance based FTLE-like method for unsteady uncertain vector fields, in: R. Peikert, H. Hauser, H. Carr, R. Fuchs (Eds.), *Topological Methods in Data Analysis and Visualization II, Mathematics and Visualization*, Springer, 2012, pp. 255–268.
- [28] A. Lasota, M.C. Mackay, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, second ed., Springer, 1994.
- [29] M. Moreau, Note on the entropy production in a discrete Markov system, *J. Math. Phys.* 19 (1978) 2494.
- [30] J. Voigt, Stochastic operators, information, and entropy, *Comm. Math. Phys.* 81 (1981) 31–38.
- [31] M. Dellnitz, A. Hohmann, A subdivision algorithm for the computation of unstable manifolds and global attractors, *Numer. Math.* 75 (1997) 293–317.
- [32] M. Dellnitz, G. Froyland, O. Junge, The algorithms behind GAIO—set oriented numerical methods for dynamical systems, in: B. Fiedler (Ed.), *Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems*, Springer, 2001, pp. 145–174.
- [33] M. Dellnitz, O. Junge, Set oriented numerical methods for dynamical systems, in: B. Fiedler, G. Iooss, N. Kopell (Eds.), *Handbook of Dynamical Systems II: Towards Applications*, World Scientific, 2002, pp. 221–264.
- [34] S. Ulam, *Problems in Modern Mathematics*, Interscience, 1964.
- [35] F. Lekien, C. Coulliette, Chaotic stirring in quasi-turbulent flows, *Phil. Trans. R. Soc. A* 365 (1861) (2007) 3061–3084.
- [36] E.N. Lorenz, Deterministic nonperiodic flow, *J. Atmospheric Sci.* 20 (2) (1963) 130–141.

- [37] W. Tucker, The Lorenz attractor exists, *C. R. Acad. Sci. Paris, Sér. I* (1999) 1197–1202.
- [38] T. Dombre, U. Frisch, J.M. Greene, M. Hénon, A. Mehr, A.M. Soward, Chaotic streamlines in the ABC flows, *J. Fluid Mech.* 167 (1986) 353–391.
- [39] I. Rypina, M. Brown, F. Beron-Vera, H. Koçak, M. Olascoaga, I. Udovydchenkov, On the Lagrangian dynamics of atmospheric zonal jets and the permeability of the stratospheric polar vortex, *J. Atmospheric Sci.* 64 (2007) 3595.
- [40] T. Koh, B. Legras, Hyperbolic lines and the stratospheric polar vortex, *Chaos* 12 (2) (2002) 382–394.
- [41] B. Joseph, B. Legras, Relation between kinematic boundaries, stirring, and barriers for the antarctic polar vortex, *J. Atmospheric Sci.* 59 (2002) 1198–1212.
- [42] G. Haller, T. Sapsis, Lagrangian coherent structures and the smallest finite-time Lyapunov exponent, *Chaos* 21 (2011) 023115.