ESAIM: M2AN 54 (2020) 1777–1795
https://doi.org/10.1051/m2an/2020027

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

HIGHER-ORDER FINITE ELEMENT APPROXIMATION OF THE DYNAMIC LAPLACIAN

Nathanael Schilling[1,*], Gary Froyland[2] and Oliver Junge[1]

Abstract. The dynamic Laplace operator arises from extending problems of isoperimetry from fixed manifolds to manifolds evolved by general nonlinear dynamics. Eigenfunctions of this operator are used to identify and track finite-time coherent sets, which physically manifest in fluid flows as jets, vortices, and more complicated structures. Two robust and efficient finite-element discretisation schemes for numerically computing the dynamic Laplacian were proposed in Froyland and Junge [*SIAM J. Appl. Dyn. Syst.* **17** (2018) 1891–1924]. In this work we consider higher-order versions of these two numerical schemes and analyse them experimentally. We also prove the numerically computed eigenvalues and eigenvectors converge to the true objects for both schemes under certain assumptions. We provide an efficient implementation of the higher-order element schemes in an accompanying Julia package.

Mathematics Subject Classification. 37C30, 37C60, 37M99, 65P99.

Received July 24, 2019. Accepted April 13, 2020.

1. Introduction

The dynamic Laplacian is a second-order partial differential operator underlying a range of methods for computing finite-time coherent sets in finite-time non-autonomous dynamical systems. It was introduced in [12] in the context of defining sets that remain coherent in a Lagrangian sense *via* dynamic isoperimetry. Coherent sets are time-dependent families of sets whose boundaries remain small relative to the volume of the set as the family evolves according to the nonlinear dynamics; extensions to weighted, curved manifolds and non-volume-preserving dynamics were made in [15]. Coherent sets are captured by the eigenvectors of the dynamic Laplacian corresponding to the leading eigenvalues (*i.e.* those closest to 0). In particular, level sets of the eigenvector corresponding to the first nontrivial eigenvalue can be used to partition the domain into two coherent sets. A dynamic Cheeger inequality [12, 15] links this eigenvalue to the ratio of boundary size to volume. Moreover, if n eigenvalues are close to zero followed by a spectral gap, this forces the eigenvectors to be close to linear combinations of indicator functions on an n-partition [6]. We note that the dynamic Laplacian is the Laplace–Beltrami operator of a weighted manifold [20] and therefore can be used as the time-independent generator of a diffusion process approximating the given (time-dependent) advection-diffusion process. Eigenvectors corresponding to small eigenvalues decay slowest under this diffusion process and yield almost-invariant sets in the

Keywords and phrases. Dynamic Laplacian, finite-time coherent sets, finite elements, transfer operator.

[1] Center for Mathematics, Technical University of Munich, Garching 85747, Germany.
[2] School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.
*Corresponding author: schillna@ma.tum.de

sense of [7]. Algorithmically, coherent sets can be extracted from the eigenvectors of the dynamic Laplacian *via*, *e.g.* clustering techniques [11], optimising eigenbasis separation [8], optimising sublevel sets [16], or sparse eigenbasis approximation [18].

In most cases, the dynamic Laplacian eigenproblem must be solved numerically. To this end, a scheme based on radial basis functions had been proposed [13], which showed high order of convergence, but suffered from a number of drawbacks like high sensitivity with respect to the radius parameter, a non-real spectrum and non-sparseness of the discretized operator. In [14], two finite element schemes were proposed (the "Cauchy-Green" (CG) and the "Transfer Operator" (TO) approach), which eliminated each of these drawbacks.

Experimentally, only piecewise linear elements were considered in [14]. In this paper, we consider higher-order (and in particular quadratic) elements and analyse convergence properties both theoretically and experimentally. We provide an efficient implementation in the Julia package `CoherentStructures.jl`. We find that using P^2 elements can give a higher asymptotic order of convergence compared to P^1 elements in the "CG" approach. For the "CG" approach, classical theory concerning eigenproblems in FEM applies. We also provide some test cases where using P^2 elements can greatly reduce the amount of information needed to calculate partitions of the domain that show important dynamical features. The question of convergence in the "TO" approaches is more subtle. We prove convergence of eigenvalues and eigenvectors for a family of TO approaches with P^1 elements, but the proof does not give any insight into the convergence rates that should be expected. For P^2 elements we do not observe asymptotically higher orders of convergence even for the simple example of a one-dimensional shift-map on the torus. This suggests that using P^2 elements does not have substantial benefits when using the "TO" approach and that one should use the simpler and well-performing linear P^1 elements in the "TO" schemes.

This paper is organized as follows. In Section 2 we recall the definition of the dynamic Laplacian. In Section 3 we look at convergence rates for the "CG" approach which follow directly from the standard FEM theory for eigenproblems. Section 4 deals with convergence of the "TO" approach, with some of the proofs relegated to Appendix A.

2. The dynamic Laplacian

Let $\mathcal{I} \subset \mathbb{R}$ denote a finite subset of time and consider a finite family $(\Omega_t)_{t\in\mathcal{I}}$ of open bounded subsets of $\mathbb{R}^d$ with Lipschitz boundary. For each $t \in \mathcal{I}$ let $T_t : \Omega_0 \to \Omega_t$ be a volume-preserving diffeomorphism. We assume that T_t is sufficiently regular so that T_t and T_t^{-1} can be smoothly extended to the boundary, and that $0 \in \mathcal{I}$ with T_0 being the identity. A typical setting in which these conditions apply are T_t taken to be time-t flow maps of a divergence-free vector-field.

Denote the Laplace operator on Ω_t by Δ_t for each $t \in \mathcal{I}$. Then the *dynamic Laplacian* (an operator on Ω_0) is given by

$$\Delta^{\mathrm{dyn}} := \frac{1}{|\mathcal{I}|}\sum_{t\in\mathcal{I}} T_t^* \Delta_t T_{t,*}$$

where $T_t^* : L^2(\Omega_t) \to L^2(\Omega_0)$ denotes the pullback by T_t defined by $T_t^* f = f \circ T_t$, and $T_{t,*} : L^2(\Omega_0) \to L^2(\Omega_t)$ is the pushforward defined by $T_{t,*} f = f \circ T_t^{-1}$.

Standard PDE-theoretic arguments can be used to show that Δ^{dyn} is a uniformly elliptic second-order partial-differential operator [10,12], with weak form [14]

$$a(u,v) := \frac{1}{|\mathcal{I}|}\sum_{t\in\mathcal{I}} \int_\Omega \nabla u \cdot [DT_t]^{-1} \left([DT_t]^{-1}\right)^T \nabla v \,\mathrm{d}\ell^d, \tag{2.1}$$

where ℓ^d is the d-dimensional Lebesgue measure. The time set $\mathcal{I}$ can also be a compact interval, and the dynamic Laplacian can be defined by means of an integral over $\mathcal{I}$ [12,15]

$$\Delta^{\mathrm{dyn}} := \frac{1}{|\mathcal{I}|}\int_{\mathcal{I}} T_t^* \Delta_t T_{t,*} \,\mathrm{d}t,$$

where $|\mathcal{I}|$ now denotes the length of the interval $\mathcal{I}$, we do not consider this generalization further here.

The uniform ellipticity of the dynamic Laplacian ensures that the bilinear form in (2.1) is coercive under suitable boundary conditions. These are determined by the choice of the underlying space S on which (2.1) acts. For natural (Neumann) boundary conditions, we look at a on $\hat{H}^1 \times \hat{H}^1$, where $\hat{H}^1 = \hat{H}^1(\Omega_0) := \left\{v \in H^1(\Omega_0) : \int f \, \mathrm{d}\ell^d = 0\right\}$ is the space of mean-free H^1 functions. For homogeneous Dirichlet boundary conditions, a must be taken to act on $H^1_0 \times H^1_0$. With either choice of boundary conditions, the variational form of the eigenproblem $\Delta^{\mathrm{dyn}} u = \lambda u$ becomes

$$a(u, v) = \lambda \langle u, v \rangle_{L^2} \quad \text{for all} \quad v \in S. \tag{2.2}$$

In both cases, ellipticity ensures that there exists a countable sequence of pairwise orthogonal eigenvectors $u_0, u_1, \ldots$ corresponding to real eigenvalues $0 \geq \lambda_0 \geq \lambda_1 \geq \cdots$. Furthermore, the span of the eigenvectors is dense in L^2 (Dirichlet) and mean-free L^2 functions (Neumann) [10, 12].

2.1. Discretisation with finite elements

A natural discretisation of the eigenproblem (2.2) is by using a finite element method (FEM) [14]. In the standard Galerkin discretization of (2.2), $S' \subset S$ is taken to be a finite dimensional approximation space spanned by some basis $(\varphi_1, \ldots, \varphi_n)$. We now find approximate solutions $u \in S'$, $\lambda \in \mathbb{R}$, that satisfy (2.2) with S' taking the place of S. In matrix form, the coefficients $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ of u with respect to the basis $(\varphi_1, \ldots, \varphi_n)$ are found by solving the generalized eigenvalue problem

$$D\mathbf{u} = \lambda M \mathbf{u}. \tag{2.3}$$

Here, $D = (D_{i,j}) = a(\varphi_i, \varphi_j)$ is referred to as the *stiffness matrix* and $M = (M_{i,j}) = \langle \varphi_i, \varphi_j \rangle_{L^2}$ as the *mass matrix*. We use P^k Lagrange nodal basis functions on some (triangular or simplicial) mesh for $\varphi_1, \ldots, \varphi_n$, but other choices are also possible in general (see [9]).

Using a finite element method for approximating Δ^{dyn} has a number of advantages: The matrix formulation (2.3) inherits self-adjointness from the continuous problem and thus the computed eigenpairs are always real. With a suitably localized basis, D and M are both sparse. Sparse Hermitian generalized eigenproblems are well known in the literature, and a number of algorithms exist for efficiently solving them [1]. Also, finite element methods have been widely studied, and there are a range of theoretical results regarding convergence (*cf.* [9]) that are applicable to some solution approaches.

While the entries of the mass matrix can be computed exactly, the integrand in (2.1) in the entries of the stiffness matrix can be of extremely high variation locally, making an accurate computation possibly expensive. Several approaches for approximating D were suggested in [14], which we investigate further in the following sections.

3. Convergence for the CG approach

In the "CG approach" of [14], a quadrature formula is used to approximate (2.1). This is the standard way of solving elliptic eigenproblems with finite elements.

We work on P^k-Lagrange finite element spaces for a family of quasi-uniform[3] triangulations (in 2 or 3 dimensions) $\left\{\mathcal{T}^0_h\right\}_{h>0}$ of Ω_0, where the family parameter h is related to the mesh size. Let $\varphi^1_h, \ldots, \varphi^{N(h)}_h$ be the associated nodal basis, $S^0_h := \operatorname{span}\left\{\varphi^1_h, \ldots, \varphi^{N(h)}_h\right\}$ the finite element approximation space, and $\hat{S}^0_h := \hat{H}^1 \cap S^0_h$. Moreover, let $\lambda_{i,h}$ be the Ritz-values with corresponding Ritz-vectors $v_{i,h}$ in $\hat{S}^0_h$, *i.e.* $\lambda_{i,h}$ is the minimum

[3] A family $\{\mathcal{T}_h\}_{h>0}$ of meshes is *quasi-uniform*, if it is shape-regular and there exists a constant $c > 0$ so that for any $h > 0$ and any element $K \in \mathcal{T}_h$ one has that $h_K \geq ch$, where h_K is the diameter of the element K. Shape regularity means that there is a constant σ_0 so that uniformly for all K it holds that $\sigma = \frac{h_K}{\rho_K} \leq \sigma_0$. Here ρ_K is the diameter of the largest ball that can be inscribed in K [9].

value of $a(v,v)$ over the set $\left\{v \in \hat{S}_h^0; \|v\|_{L^2} = 1, \langle v, v_j\rangle = 0 \text{ for } j = 1, \cdots, (i-1)\right\}$ and $v_{i,h}$ is the corresponding minimizer.

Under the assumption that the eigenvectors of $a : \hat{H}^1 \times \hat{H}^1 \to \mathbb{R}$ are in H^{k+1}, classical results from FEM theory [9, 25] give that (where C stands for a generic constant):

$$\frac{\lambda_{i,h} - \lambda_i}{\lambda_i} \leq Ch^{2k} \quad (h \to 0). \tag{3.1}$$

Moreover, it holds for simple (and L^2-normalised) eigenvectors that

$$\|v_i - v_{i,h}\|_{H^1} \leq Ch^k \quad (h \to 0). \tag{3.2}$$

In the case that the problem (2.1) satisfies the technical assumption of being *regularizing* [9] (*i.e.* $\|Pu\|_{H^2} \leq C\|u\|_{L^2}$), the "Aubin–Nitsche trick" gives

$$\|v_i - v_{i,h}\|_{L^2} \leq Ch^{k+1} \quad (h \to 0). \tag{3.3}$$

Note that not all elliptic problems are regularizing. We thus cannot expect (3.3) to hold in general.

If the entries of the stiffness and the mass matrix are approximated with a quadrature rule of order $2k-1$, the convergence orders are unaffected, provided that the quadrature points contain a $\mathcal{P}^k$ unisolvent set [2]. Moreover, the results generalize to non-simple eigenspaces and Kato's subspace distance used for the error [2].

3.1. Numerical experiments

We now aim at reproducing the predicted convergence rates in numerical experiments. As a reference, we compute (L^2 normalized) eigenpairs (λ_i, v_i) on a very fine mesh and estimate the error in the eigenvector $v_{i,h}$ by computing the L^2-distance of $v_{i,h}$ to the closest reference eigenvector, given by the expression

$$e_{i,h} := \sqrt{1 - |\langle v_i, v_{i,h}\rangle_{L^2}|^2}. \tag{3.4}$$

The L^2 inner-product is approximated by first interpolating to the fine grid, and then calculating the inner product there using quadrature.

To generalize this to higher-dimensional eigenspaces, let $\tilde{V}$ and V be two m-dimensional subspaces of L^2 with orthonormal bases given by $\{\tilde{v}_1, \ldots, \tilde{v}_m\}$ and $\{v_1, \ldots, v_m\}$. Let $\tilde{P}, P$ be orthogonal projections onto $\tilde{V}$ and V respectively. As a measure of the subspace error [21], we maximize $\|\tilde{v} - P\tilde{v}\|_{L^2}$ over $\tilde{v} \in \tilde{V}$ with $\|\tilde{v}\|_{L^2} = 1$. This is equivalent to maximizing $\sqrt{1 - \|P\tilde{v}\|^2}$, which has maximum given by $\sqrt{1 - \rho_{\min}(P)^2}$ where $\rho_{\min}$ is the smallest singular value of P on $\tilde{V}$ which is also that of the matrix $(\langle \tilde{v}_i, v_j\rangle_{L^2})_{i,j=1}^m$.

In general, one is not directly interested in the eigenvectors themselves, but in a partition of the domain obtained by a suitable post-processing of the eigenvectors (*e.g.* as mentioned, thresholding, clustering or sparse eigenbases). This motivates us making a qualitative comparison of such partitions based on eigenvectors for different approximation spaces. Here, we focus in particular on the question of how well Lagrangian coherent sets can be found with as little information about the flow as possible. Evidently, in this low data case some features may not be accurately resolved even if the eigenvectors clearly give some indication of their existence. We therefore also look at how coarse the grid can be made without affecting the topology of such a partition in some test cases.

In the sequel, mesh sizes of the form $n \times m$ refer to a regular triangular mesh with $n \times m$ (non-inner) nodes in each direction. For all time-integration done in this work we used a relative and absolute tolerance of 10^{-8} and the `DifferentialEquations.jl` Julia package [23] with the `BS5()` solver [3]. Derivatives of flow maps were approximated with second order central finite differences. The stiffness matrix was calculated by approximating the integral in (2.1) on each element with nodal basis functions for u and v using quadrature. The results from

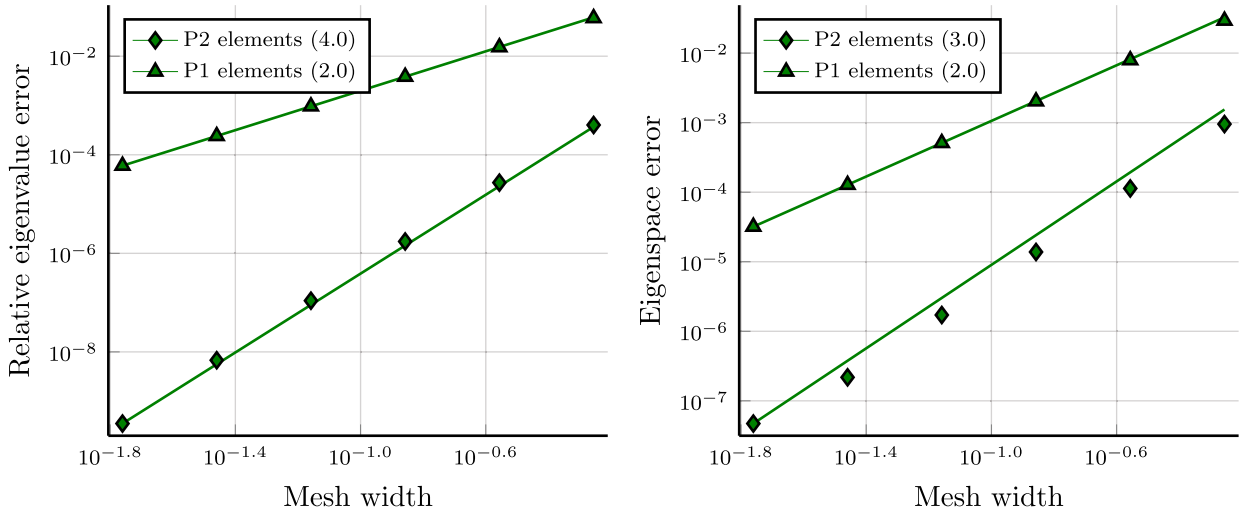

FIGURE 1. Standard map: errors in the first nontrivial eigenvalue (*left*) and 2-dimensional eigenspace corresponding to the two smallest nontrivial eigenvalues (*right*) of the dynamic Laplacian with the CG approach. The slopes of the corresponding lines are given in brackets in the legends.

this were then summed over all elements in the support of a basis function to give the corresponding entry in the stiffness matrix. Similarly, the integral in the mass matrix was also calculated element-wise and then additively combined to give the full mass matrix. The `CoherentStructures.jl` package internally uses the `JuAFEM.jl` package [5].

3.1.1. Standard map

As a simple first example, we consider two iterations of the standard map

$$T(x, y) = (x + y + a\sin(x), y + a\sin(x))$$

on the 2-torus $\mathbb{S}^1 \times \mathbb{S}^1$ with parameter $a = 0.971635$. This is the first example considered in [13] and is a weakly nonlinear map. Figure 1 shows the eigenvalue and eigenspace errors in dependence of the mesh width for triangular P^1 and P^2 Lagrange elements.

The reference solution was computed with P^2 elements on a regular 1025×1025 grid with quadrature order 5. In this case, the experimentally observed rates almost perfectly agree with the predictions (3.1) and (3.3) even for quadrature order 2 in the computation of the entries of stiffness- and mass-matrix. Improvements in errors ranged between one and two orders of magnitude when moving from P^1 to P^2 elements. To be able to directly compare these results with the corresponding plots for TO methods in Figure 5, we used a two-dimensional eigenspace in Figure 1.

3.1.2. Cylinder flow

As a second example, we consider the cylinder flow map used in [13] based on [17]. This is a time-dependent flow on the cylinder $\mathbb{S}^1 \times [0, \pi]$ defined by the non-autonomous ordinary differential equation

$$\begin{aligned}\dot{x}(t) &= c - A(t)\sin(x - \nu t)\cos(y) + \varepsilon G(g(x, y, t))\sin(t/2)\\ \dot{y}(t) &= A(t)\cos(x - \nu t)\sin(y)\end{aligned}$$

where $A(t) = 1 + 0.125\sin\left(2\sqrt{5}t\right)$, $G(\psi) = 1/\left(\psi^2 + 1\right)^2$ and $g(x, y, t) = \sin(x - \nu t)\sin(y) + y/2 - \pi/4$. Here the parameters $c = 0.5, \nu = 0.5, \varepsilon = 0.25$ were used, the time-interval was $[0, 40]$.

The reference solution was computed on a regular mesh on 1025×1025 nodes with triangular P^2-Lagrange elements and quadrature order 8. We used quadrature order 5 for the numerical experiments (since we did not observe the same rates for smaller orders). The results for this numerical experiment are shown in Figure 6.

The orders of convergence observed in Figure 3 for the eigenvalue and eigenvector errors is surprisingly low; the experimental values are almost exactly half of the orders predicted in (3.1), (3.2) or (3.3). The observed order of the eigenvector errors for P^1 elements is particularly poor. We also note that the slopes shown in the figure do not remain consistent when varying the quadrature order.

For this flow (and other highly stretching and highly nonlinear systems), the mean diffusion tensor $A = [DT_t]^{-1}\left([DT_t]^{-1}\right)^T$ used in the weak formulation of the dynamic Laplacian (2.1) has extremely high variation locally (*cf.* Fig. 3). We suspect that even the reference grid of 1025×1025 nodes is not sufficiently resolved to be in a regime where the convergence orders predicted by the theory can be observed.

3.1.3. Bickley jet

The Bickley jet flow is a well-known test case first introduced in [24]. The flow is defined by a stream-function

$$\psi(x,y,t) = -U_0 L \tan h\,(y/L_0) + \sum_{i=1}^{3} A_i U_0 L \sec h^2(y/L)\cos\left(k_i\left(x - c_i t\right)\right) \tag{3.5}$$

with constants U_0, L_0, A_i, k_i, c_i on a cylindrical domain (see [24]). We considered the flow for a timespan of 40 days. 8-partitions were computed by k-means clustering on 200×60 values of the leading eigenvectors of the FEM approximation. The mesh widths in Figure 4 below are the lowest (at the aspect ratio $10:3$) for which the topology of the clustering result does not change. It is possible to obtain the same result (topologically) with P^2 elements at significantly reduced cost (in terms of the number of quadrature point used) to P^1 elements. In many cases, the computation requiring the most runtime (apart from the clustering) is the computation of diffusion tensors by time-integration at the quadrature points. As shown in Figure 4, in this case this is reduced by a factor of 25 by using P^2 elements.

4. Convergence for the TO approach

The CG approach has the disadvantage of requiring the numerical approximation of the derivative DT_t and a subsequent quadrature which might be challenging if DT_t has high variation locally.

An alternative approach [14] is to rewrite the weak form (2.1) of the dynamic Laplacian operator using the transfer operator T_t. In the sequel we discuss the case $t \in \mathcal{I} := \{0,1\}$, although the results hold for general finite $\mathcal{I}$. Letting a^t be the weak-form of the Laplacian on Ω_t, we have (for $u, v \in H^1(\Omega_0)$)

$$a(u,v) = \frac{1}{2}\left(a^0(u,v) + a^1(Tu,Tv)\right)$$

where we write $T := T_1$, and for brevity overload this notation so that T acts on elements of Ω_0 in the usual way *and* acts on functions in $L^2(\Omega_0)$ as $Tu := T_{1,*}u$. In the transfer operator (TO) approach, we approximate T on a suitable subspace of H^1 by an operator of the form $I_h T$, where I_h is some projection operator. Let $\left\{\mathcal{T}_h^1\right\}_{h>0}$ be a quasi-uniform family of triangulations on Ω_1, S_h^1 the approximation space with Lagrange $P^{k'}$ finite elements and $\hat{S}_h^1 := S_h^1 \cap \hat{H}^1(\Omega_1)$. Note that in general, the collection $\left\{\mathcal{T}_h^0\right\}_{h>0}$ (introduced at the beginning of Sect. 3) and $\left\{\mathcal{T}_h^1\right\}_{h>0}$ are unrelated. We consider the following options for I_h in the sequel:

(1) the L^2-orthogonal projection onto S_h^1 (L^2-*Galerkin*),
(2) the H^1-orthogonal projection onto S_h^1 (H^1-*Galerkin*),
(3) the (Lagrange P^k) nodal interpolation on S_h^1 (*collocation*).

Case 3 corresponds to the "TO approach" from [14], with the difference between the adaptive and non-adaptive approaches being the choice of the mesh $\{\mathcal{T}_h^1\}_{h>0}$.

We now define the (h-dependent) symmetric bilinear form

$$\tilde{a}_h(u,v) := \frac{1}{2}\left(a^0(u,v) + a^1\left(I_hTu, I_hTv\right)\right). \tag{4.1}$$

As in [14], eigenpairs are numerically approximated by calculating generalized eigenpairs of $(\tilde{D}_h, M_h)$ given by

$$\begin{aligned}\left(\tilde{D}_h\right)_{ij} &:= \tilde{a}_h\left(\varphi_h^i, \varphi_h^j\right)\\ \left(M_h\right)_{ij} &= \langle\varphi_h^i, \varphi_h^j\rangle_{L^2}.\end{aligned}$$

As a^0 is elliptic on $\hat{H}^1(\Omega_0)$ and a_1 is non-negative, it follows immediately from (4.1) that $\tilde{a}_h$ is elliptic uniformly in h on $\hat{S}_h^0$, with a lower bound of the ellipticity constant being half the ellipticity constant of a^0.

Hypotheses on I_h and the setup. In order to investigate the convergence of the numerically calculated eigenpairs to those of the dynamic Laplacian, we will require a regularity assumption on the meshes used:

(A0) $\{\mathcal{T}_h^0\}_{h>0}$ and $\{\mathcal{T}_h^1\}_{h>0}$ are families of quasi-uniform triangulations of Ω_0 and Ω_1 respectively that are nested in the sense that $S_{h'} \subset S_h$ for $h < h'$.

We will need to consider the following conditions for I_hT acting on S_h^0:

(A1) exactness for constant functions: $I_hT\chi_{\Omega_0} = \chi_{\Omega_1}$,
(A2) H^1-convergence: $\lim_{h\to 0}\|I_hTf - Tf\|_{H^1} = 0$ for any $f \in \cup_h S_h^0$,
(A3) L^2-stability: For all $h' > 0$ it holds that $\sup_{h<h'}\|I_hT|S_{h'}^0\|_{L^2\to L^2} \le C$, where $C > 0$ is independent of h'.

We now consider whether the conditions (A1)–(A3) are satisfied by the three variants of I_h:

(i) It is easy to see that condition (A1) is satisfied for all three variants, by the fact that $T := T_1$ is volume-preserving and I_h is exact on constant functions.
(ii) Condition (A2) is satisfied by the H^1-Galerkin approach. For the L^2-Galerkin approach, (A0) implies (A2) [9].
(iii) Condition (A3) is trivially satisfied by the L^2-Galerkin approach.

We will show (A2) and (A3) for the case that I_h comes from the collocation approach. In Appendix A, we give proofs of the following two lemmata for P^1-triangular elements under assumption (A0).

Lemma 4.1 (L^2-stability (A3) of P^1-nodal interpolation)**.** *Let $I_h : C\left(\overline{\Omega_1}\right) \to S_h^1$ be the P^1-Lagrange nodal interpolation operator corresponding to $\{\mathcal{T}_h^1\}_{h>0}$. If (A0) holds, then there exists a constant $C > 0$ so that $\|I_hTv\|_{L^2} \le C\|v\|_{L^2}$ for all $v \in S_{h'}^0$ and all $h < h'$.*

Lemma 4.2 (H^1-convergence (A2) of P^1-nodal interpolation)**.** *Let $I_h : C\left(\overline{\Omega_1}\right) \to S_h^1$ be the P^1-Lagrange nodal interpolation operator corresponding to $\{\mathcal{T}_h^1\}_{h>0}$. If (A0) holds, and if v is piecewise C^∞ on Ω_1 then*

$$\|I_hv - v\|_{H^1} \to 0 \quad \text{as} \quad h \to 0.$$

Convergence for homogeneous Neumann Boundary. Let $r(v) = a(v,v)/\langle v,v\rangle_{L^2}$ and $\tilde{r}_h(v) = \tilde{a}_h(v,v)/\langle v,v\rangle_{L^2}$ be the Rayleigh quotients for a and $\tilde{a}_h$ respectively. Assuming (A0), we know that both $r(v)$ and $\tilde{r}_h(v)$ are trivially minimized by a constant function. Going to mean-free functions, we therefore define λ_1 and $\tilde{\lambda}_{h,1}$ to be the corresponding smallest Ritz values on $\hat{H}^1$ and $\hat{S}_h^0$, respectively and v, $\tilde{v}_h$ the corresponding Ritz vectors (normalized to have unit L^2-norm).

We now derive a convergence result for $\tilde{v}_h$ to v in two parts:

Theorem 4.3. *Provided that (A0) holds and I_h satisfies the assumptions (A1) and (A2), then*

$$\limsup_{h\to 0} \tilde{\lambda}_{h,1} \leq \lambda_1.$$

Proof. Pick any sequence h_n with $h_n \to 0$. For arbitrary $w \in \hat{S}^0_{h_n}$, have

$$\begin{aligned}\limsup_{n\to\infty} \tilde{\lambda}_{h_n,1} &= \limsup_{n\to\infty} \tilde{r}_{h_n}(\tilde{v}_{h_n}) \quad \text{(definition of } \tilde{v}_{h_n}\text{)} \\ &\leq \limsup_{n\to\infty} \tilde{r}_{h_n}(w) \qquad \text{(Ritz-values minimize the Rayleigh-quotient).}\end{aligned}$$

Now for arbitrary $\epsilon > 0$, pick[4] some $w \in \hat{S}^0_{h'}$ where $h' = h'(\epsilon)$ with $\|w\|_{L^2} = 1$ so that $r(w) < \lambda_1 + \epsilon$. Then by the assumption that our meshes are nested, w is in the domain of $\tilde{a}_{h_n}$ for sufficiently large n and we have:

$$\begin{aligned}\limsup_{n\to\infty} \tilde{\lambda}_{h_n,1} &\leq \limsup_{n\to\infty} \tilde{r}_{h_n}(w) \\ &= \limsup_{n\to\infty} \tilde{a}_{h_n}(w,w) \quad \text{(definition of Rayleigh Quotient)} \\ &= \limsup_{n\to\infty} \left(a(w,w) + \tilde{a}_{h_n}(w,w) - a(w,w)\right) \\ &\leq \limsup_{n\to\infty} \left(\lambda_1 + \epsilon + a^1\left(I_{h_n}Tw, I_{h_n}Tw\right) - a^1(Tw,Tw)\right).\end{aligned}$$

For this fixed v we know that $a^1\left(I_{h_n}Tw, I_{h_n}Tw\right) - a^1(Tw,Tw) \to 0$ by the H^1 convergence assumption, and by H^1-continuity of a^1. This gives the result. □

Theorem 4.4. *Provided that (A0) holds and I_h satisfies (A1) and (A3), then*

$$\liminf_{h\to 0} \tilde{\lambda}_{h,1} \geq \lambda_1.$$

Proof. We use an approach reminiscent of the *direct method* in the calculus of variations (*cf.* [19]). Assume that I_h satisfies both properties, but (for the sake of contradiction) that there exists a sequence $h_n \to 0$ monotonically with $\tilde{\lambda}_{h_n,1} \to \beta < \lambda_1$. Note that by positivity of $\tilde{a}_h$ the sequence cannot diverge to $-\infty$, and by Theorem 4.3 it cannot diverge to $+\infty$. Hence, using coercivity of a^0 on $\hat{H}^1$ we get that $|\tilde{v}_{h_n}|_{H^1} \leq C$ for some $C > 0$, otherwise the a^0 term would go to infinity. Here and below, C represents a generic constant independent of h_n. The Poincaré-Friedrichs inequality (*cf.* [9], Lem. B.66) for mean-free functions states that the L^2 norm is uniformly bounded by a constant times the H^1 seminorm, thus it follows that

$$\|\tilde{v}_{h_n}\|_{H^1} \leq C. \tag{4.2}$$

Recall that the Banach–Alaoglu theorem states that the closed unit ball in H^1 is weakly sequentially compact, while the Rellich–Kondrachev theorem asserts that norm-bounded H^1 sets are L^2-precompact. Hence we can assume that (going to a subsequence) $\tilde{v}_{h_n} \rightharpoonup \tilde{v}$ in H^1 and $\tilde{v}_{h_n} \to \tilde{v}$ in L^2 for some $\tilde{v} \in H^1$. Note that the two limits are indeed the same since $\cdot \mapsto \langle \cdot, v\rangle_{L^2}$ is a continuous linear functional on H^1 for all $v \in L^2$.

As $\tilde{v}_{h_n} \in \hat{H}^1$, it follows that $\tilde{v} \in \hat{H}^1$, the L^2-convergence gives us further that $\|\tilde{v}\|_{L^2} = 1$.

Without loss of generality, assume that $\ell^d(\Omega_1) = 1$ and define $J_n : \Omega_0 \to \mathbb{R}$ as the constant function taking the value $\int_{\Omega_1} I_{h_n} T\tilde{v}_{h_n} \mathrm{d}\ell^d$, and observe that

$$a^1\left(I_{h_n}T\tilde{v}_{h_n}, I_{h_n}T\tilde{v}_{h_n}\right) = a^1\left(I_{h_n}T\tilde{v}_{h_n} - J_n, I_{h_n}T\tilde{v}_{h_n} - J_n\right). \tag{4.3}$$

[4]That we can do this follows the classical theory of density of finite-element spaces in Sobolev spaces ([9], Cor. 1.110), or from convergence of eigenvalues of the Galerkin approximation such as equation (3.1) in the case that the eigenvectors are sufficiently smooth at the boundary.

We know from the positivity of a^0 that $a^1\left(TI_{h_n}\tilde{v}_{h_n}, TI_{h_n}\tilde{v}_{h_n}\right) \leq \tilde{r}_{h_n}\left(\tilde{v}_{h_n}\right)$ and as the right hand side converges, the left hand side is bounded by a constant that does not depend on n. Hence, by coercivity of a^1 on mean-free H^1 functions, it follows from (4.3) and the Poincaré–Friedrichs inequality that $\|I_{h_n}T\tilde{v}_{h_n} - J_n\|_{H^1} < C$, yielding that $I_{h_n}T\tilde{v}_{h_n}$ is H^1-bounded as long as J_n is – which is the case by L^2-stability, the L^2–L^2 continuity of T, and the fact that $\|\tilde{v}_{h_n}\|_{L^2} = 1$. Hence $I_{h_n}T\tilde{v}_{h_n}$ is (uniformly in n) H^1-bounded. Going to a subsequence, it is therefore weakly H^1- and strongly L^2-convergent (as above). As the weak-H^1 and L^2-limits coincide, to show that the sequence converges to $T\tilde{v}$ weakly in H^1, it is enough to show L^2-convergence.

Let C_1 be the stability constant from the L^2-stability condition. We know that as $v_{h_n} \to \tilde{v}$ in L^2, it holds that $T\tilde{v}_{h_n} \to T\tilde{v}$. Therefore given $\varepsilon > 0$, pick m so that $\|T\tilde{v}_{h_m} - T\tilde{v}\|_{L^2} \leq \varepsilon/3$ and and $C_1\|\tilde{v}_{h_n} - \tilde{v}_{h_m}\|_{L^2} \leq \varepsilon/3$ for all $n \geq m$. Also, by the nesting property of the meshes $\tilde{v}_{h_n} - \tilde{v}_{h_m} \in \hat{S}^0_{h_n}$ and hence

$$\begin{aligned}\|I_{h_n}T\tilde{v}_{h_n} - T\tilde{v}\|_{L^2} &\leq \|I_{h_n}T\tilde{v}_{h_n} - I_{h_n}T\tilde{v}_{h_m}\|_{L^2} + \|I_{h_n}T\tilde{v}_{h_m} - T\tilde{v}_{h_m}\|_{L^2} + \|T\tilde{v}_{h_m} - T\tilde{v}\|_{L^2}\\ &\leq C_1\|\tilde{v}_{h_n} - \tilde{v}_{h_m}\|_{L^2} + \|I_{h_n}T\tilde{v}_{h_m} - T\tilde{v}_{h_m}\|_{L^2} + \|T\tilde{v}_{h_m} - T\tilde{v}\|_{L^2}\\ &\leq 2\varepsilon/3 + \|I_{h_n}T\tilde{v}_{h_m} - T\tilde{v}_{h_m}\|_{L^2}.\end{aligned}$$

Now using the H^1-convergence property, pick $n_0 \geq m$ so that for $n \geq n_0$, we have

$$\|I_{h_n}T\tilde{v}_{h_m} - T\tilde{v}_{h_m}\|_{L^2} \leq \varepsilon/3.$$

It follows that for all $n \geq n_0$

$$\|I_{h_n}T\tilde{v}_{h_n} - T\tilde{v}\|_{L^2} \leq \epsilon.$$

Summarizing, we now have proved that

(1) $\tilde{v}_{h_n} \rightharpoonup \tilde{v}$ in H^1,
(2) $I_{h_n}T\tilde{v}_{h_n} \rightharpoonup T\tilde{v}$ in H^1, and
(3) $\tilde{v} \in \hat{H}^1$ with $\|\tilde{v}\|_{L^2} = 1$.

The functions $u \mapsto \sqrt{a^0(u,u)}$ and $u \mapsto \sqrt{a^1(u,u)}$ are norms that are equivalent to the $\hat{H}^1$-norm by the Poincaré–Friedrichs inequality[5]. By the Banach–Steinhaus theorem, norms are weakly sequentially lower-semicontinuous. The L^2 convergence of $I_{h_n}Tv_{h_n}$ also gives that $J_n \to 0$ in L^2, and therefore also in H^1 as J_n is constant). Thus we have:

$$\begin{aligned}a^0\left(\tilde{v},\tilde{v}\right) &\leq \liminf_{n\to\infty} a^0\left(\tilde{v}_{h_n},\tilde{v}_{h_n}\right)\\ a^1\left(T\tilde{v},T\tilde{v}\right) &\leq \liminf_{n\to\infty} a^1\left(I_{h_n}T\tilde{v}_{h_n} - J_n, I_{h_n}T\tilde{v}_{h_n} - J_n\right)\\ &= \liminf_{n\to\infty} a^1\left(I_{h_n}T\tilde{v}_{h_n}, I_{h_n}T\tilde{v}_{h_n}\right)\end{aligned}$$

yielding

$$\begin{aligned}a\left(\tilde{v},\tilde{v}\right) &= a^0\left(\tilde{v},\tilde{v}\right) + a^1\left(T\tilde{v},T\tilde{v}\right)\\ &\leq \liminf_{n\to\infty}\left(a^0\left(\tilde{v}_{h_n},\tilde{v}_{h_n}\right) + a^1\left(I_{h_n}T\tilde{v}_{h_n}, I_{h_n}T\tilde{v}_{h_n}\right)\right)\\ &= \liminf_{n\to\infty} \tilde{r}_{h_n}\left(\tilde{v}_{h_n}\right)\\ &= \lim_{n\to\infty} \tilde{\lambda}_{h_n,1} = \beta < \lambda_1.\end{aligned}$$

This is a contradiction to the definition of λ_1, which requires that $a\left(\tilde{v},\tilde{v}\right) \geq \lambda_1$ and thus the claim is proved. □

[5]The weak topologies also coincide, as for any $f \in \hat{H}^1$, the Riesz representation theorem gives a $\tilde{f} \in \hat{H}^1$ so that $\langle\cdot, f\rangle_{H^1} = a_0(\cdot,\tilde{f})$.

Theorem 4.5. *The conclusion of the previous theorem also holds if (A0) is satisfied and I_h comes from the H^1-Galerkin approach.*

Proof. The previous proof required the L^2-stability only in two places. The first was in bounding J_n, this is trivially bounded as $\|I_h\|_{H^1 \to H^1} = 1$ and $\tilde{v}_{h_n}$ is H^1-bounded. The second was in showing that $I_{h_n} T \tilde{v}_{h_n} \rightharpoonup T\tilde{v}$. But this follows from the fact that for $f \in H^1$:

$$\langle I_{h_n} T \tilde{v}_{h_n}, f \rangle_{H^1} = \langle T \tilde{v}_{h_n}, I_{h_n} f \rangle_{H^1}.$$

As $I_h f \to f$ in H^1 and $T\tilde{v}_{h_n} \rightharpoonup T\tilde{v}$, this gives the claim. □

Theorem 4.6 (Eigenvector convergence)**.** *Assume that (A0)–(A3) hold, and that the eigenspace corresponding to λ_1 is one-dimensional. Then $|\langle \tilde{v}_h, v \rangle_{L^2}| \to 1$ for $h \to 0$.*

Proof. The proof of Theorem 4.4 shows that for any monotone sequence $h_n \to 0$ for which λ_{h_n} converges to λ_1, every in L^2 convergent subsequence of $\tilde{v}_{h_n}$ converges to some $\tilde{v} \in H^1$ with $r(\tilde{v}) = \lambda_1$. Theorems 4.4 and 4.3 show that $\lambda_{h_n} \to \lambda_1$ for any $h_n \to 0$. As $\tilde{v}_{h_n}$ is bounded in H^1 and therefore pre-compact in L^2, this means that $\tilde{v}_{h_n} \to \tilde{v}$ in L^2 with $r(\tilde{v}) = \lambda_1$, from which it follows that $\tilde{v} = \pm v$, giving the result for a subsequence. We get the final result by noting that if $\tilde{v}_{h_n}$ has a subsequence $\tilde{v}_{h_{n_k}}$ for which $|\langle \tilde{v}_{h_{n_k}}, v \rangle| \to c \neq 1$, by the above it must have a (further) subsequence that converges to a different value, a contradiction. □

Theorem 4.7 (Higher eigenpairs)**.** *Assuming (A0)–(A3), the results of the previous theorems generalize to the m-th eigenpair for $m > 1$.*

Proof. The generalization of Theorem 4.3 can be proven by the same method as the case $m = 1$: instead of choosing a w with $r(w) < \lambda_1 + \epsilon$, choose a series of pairwise-orthogonal $(w_i)_{i=1\ldots m} \in \hat{S}_h^0$ with the property that $r(w_i) \leq \lambda_i + \epsilon$.

The proof of Theorem 4.4 must only be modified to show that the limit point $\tilde{v}_m$ is L^2-orthogonal to the previous eigenvectors span $\{v_1, \ldots, v_{m-1}\}$. Writing $\tilde{v}_{m,h}$ for the m-th Ritz-vector of $\tilde{a}_h$ on $\hat{S}_h^1$, this readily follows (by induction) from the fact that $\tilde{v}_{m,h_n} \perp \operatorname{span}\{\tilde{v}_{1,h_n}, \ldots, \tilde{v}_{m-1,h_n}\}$ and thus also $\langle \tilde{v}_{i,h_n}, T v_i \rangle \to 0$ in L^2 for $i = 1, \ldots, m-1$.

The proof of the generalization of Theorem 4.6 then follows from the generalization of Theorem 4.4. □

Convergence for homogeneous Dirichlet boundary. For the case of a homogeneous Dirichlet boundary, we must replace $\hat{H}^1$ with H_0^1. Here we require the additional condition:

(A1′) Boundary preservation: for $v \in H_0^1(\Omega_1) \cap S_h^0$ we have that $I_h T v \in H_0^1(\Omega_1)$.

We note that the I_h from collocation and H^1-Galerkin have this property.

Theorem 4.8. *The convergence results of Theorems 4.3, 4.4 and 4.6 hold also for the Ritz-values for homogeneous Dirichlet boundary conditions, provided that (A0) holds and I_h satisfies (A1′) in addition to satisfying (A2) and (A3).*

Proof. Same as the proofs for the Neumann case, except that coercivity of the bilinear forms involved on $\hat{H}^1$ is replaced with that on H_0^1. The boundary preservation property ensures we do not need J_n in the proof of Theorem 4.4, the proof otherwise goes on exactly the same lines. □

4.1. Numerical experiments

The three TO approaches yield a stiffness-matrix D of the form

$$D = D_0 + A^T D_1 A$$

where A is the representation matrix of $I_h T$ and D_0, D_1 result from the bilinear form of the static Laplacian. In the case of the L^2- or H^1-Galerkin approach, the matrix A has the form

$$A = G^{-1}\tilde{A},$$

where $\tilde{A} = \langle \varphi_i, T\varphi_j \rangle$ and $G_{i,j} = \langle \varphi_i, \varphi_j \rangle$. Note that sparsity of G and $\tilde{A}$ does not necessarily imply sparsity of A. A naive calculation of the full matrix A therefore renders the computation of (some) eigenvalues of D too expensive for larger problems. We therefore do not include the Galerkin–TO approaches in the numerical experiments below, further work is needed to determine if the Galerkin–TO methods can be modified to overcome this issue.

All TO results shown below are therefore the collocation-based ones. For the "non-adaptive TO" experiments, we used identical regular meshes for the initial and final triangulations $\mathcal{T}_h^0$ and $\mathcal{T}_h^1$. For the "adaptive TO" experiments, the initial mesh was regular but the final mesh was a Delaunay-triangulation (with the `VoronoiDelaunay.jl` package [22]) of the images of the initial mesh points under the flow map T. The theory outlined above requires (i) nested meshes, and (ii) uniform bounds on shape-regularity and quasi-uniformity constants of the meshes as they are refined; each of these properties is difficult to guarantee in general for the image triangulations. We nevertheless include the adaptive TO in the plots for comparison.

We note that the finest grid used for the TO experiments is 257×257. This is not as fine as the finest grid used in the CG approach (513×513). The reason for this is that the stiffness matrix D from the TO approach does not have the same banded structure as that coming from the CG approach, making a solution of the eigenproblem more expensive (for the example of the cylinder flow: 1 h 28 min for the non-adaptive TO using a P^1 513×513 mesh *vs.* 14 min for the CG approach with P^2 elements on a 1025×1025 mesh). The reference solution here is the same as that used for the CG experiments.

Our numerical experiments appear to be consistent with our theoretical results proven above.

4.1.1. Standard map

We investigate the convergence of collocation-based TO approaches on the Standard map test case.

The results for this numerical experiment are shown in Figure 2. We see an improvement by a full order of magnitude in the errors in both the eigenvalues and the eigenspaces for P^2 elements over P^1 elements. Concerning the convergence order, this experiment suggests that the order for P^2 elements is not higher than for P^1 elements in the TO approach, which match those obtained for eigenvectors based on P^1 elements in the CG approach.

4.1.2. Cylinder flow

We observe a smaller error reduction in the eigenvalue and eigenvector errors for P^2 *vs.* P^1 elements for the cylinder flow (Fig. 6) compared to the standard map example. The reduction in convergence orders compared to the standard map (*i.e.* comparing Fig. 6 to Fig. 5) mirrors the reductions observed in the corresponding CG experiments (*i.e.* comparing Fig. 2 to Fig. 1), where the orders are around half those seen for the standard map. Similarly to the standard map results, the P^2 elements did not perform asymptotically better than P^1 elements.

4.1.3. Bickley jet

We repeat the experiment from Section 3.1.3 with the (non-adaptive) TO approach, see Figure 7. Here we note that this experiment falls outside of the theory described so far as the computational domain is not exactly mapped to itself by the flow, but our attempts to work around this (by *e.g.* having a larger mesh at the final

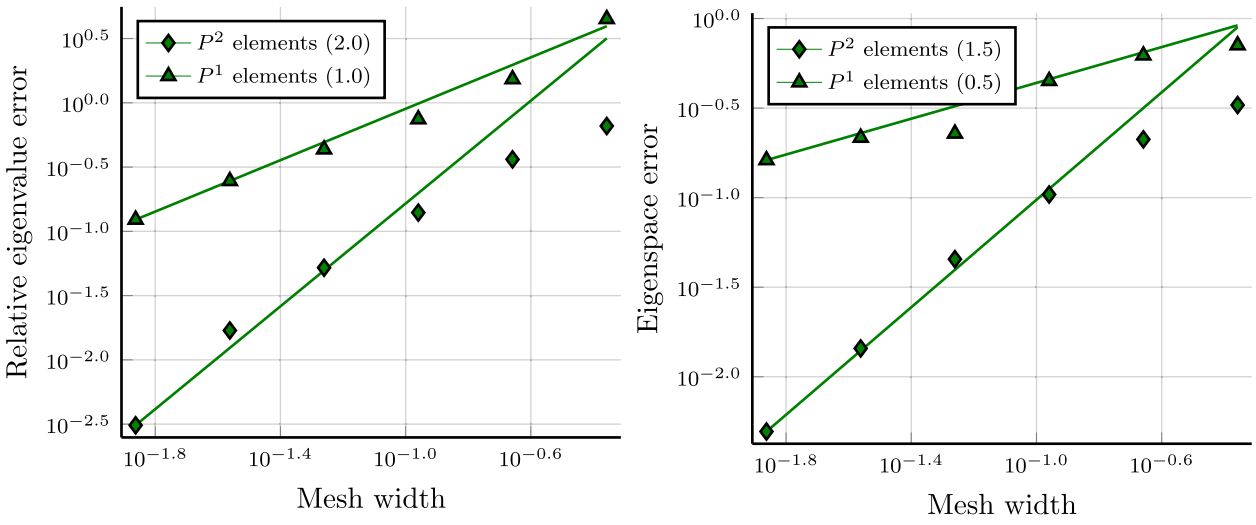

FIGURE 2. Cylinder flow: errors in the first nontrivial eigenvalue (*left*) and corresponding 1-dimensional eigenspace (*right*) of the dynamic Laplacian discretized with CG approach. The slopes of the corresponding lines are given in brackets in the legends.

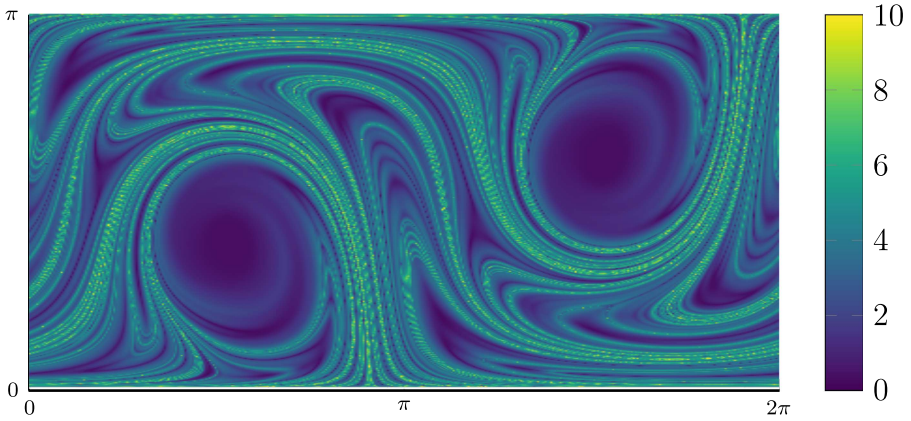

FIGURE 3. Cylinder flow: log of the trace of the mean diffusion tensor, $\log_{10} \operatorname{tr}(A(x))$.

time) did not significantly affect the topology of the clustering result in the low data cases we looked at. We note that on even finer meshes, we also saw clusterings different to those in Figure 7, such as those where the central jet was a separate cluster. We did not concern ourselves with these further as our aim was merely to compare the different discretisations in the low data limit, and it is not clear what the "correct" clustering should be.

As in Section 3.1.3, the mesh widths in Figure 7 below are the lowest at that aspect ratio for which the topology of the clustering result does not change; nodes at the boundaries that are identified due to the periodic boundary conditions are counted twice. Here, the majority of the computational cost (apart from the clustering) is given by the evaluation of the (inverse) flow map, which has to be evaluated once for each basis function of the finite element space. Note that there are three basis functions per element for P^1 and six per element for P^2 elements. While for P^2 elements, a coarser mesh was sufficient, the number of basis functions was comparable (even a little larger) than for the P^1 case – which is in contrast to the corresponding CG experiment.

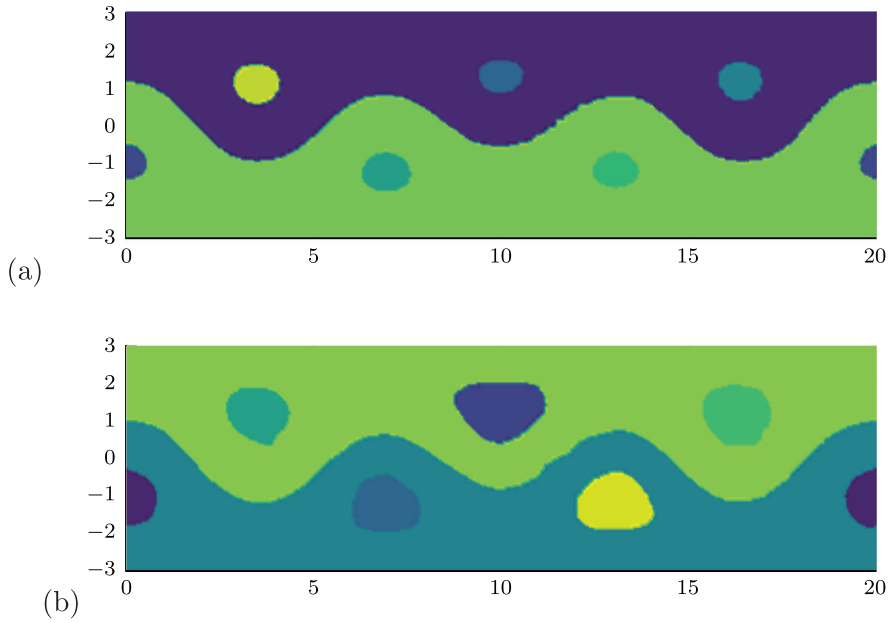

FIGURE 4. Bickley jet: comparison of coherent sets obtained by the CG approach with P^1 *vs.* P^2 elements. (a) P^1 elements on 101×31 mesh (18 000 quadrature points, 6000 triangles). (b) P^2 elements on 21×7 mesh (720 quadrature points, 240 triangles).

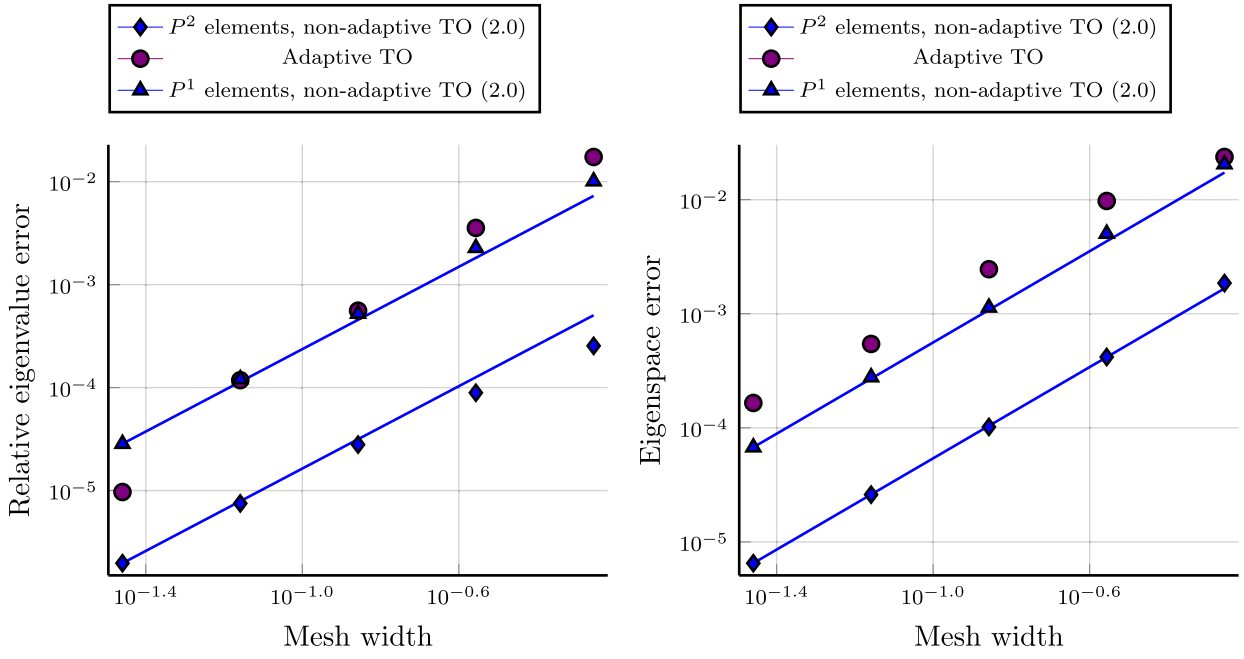

FIGURE 5. Standard map: errors in the first nontrivial eigenvalue (*left*) and 2-dimensional eigenspace spanned by the smallest eigenvectors (*right*) of the dynamic Laplacian discretized with the TO approach. The slopes of the corresponding lines are given in brackets in the legends.

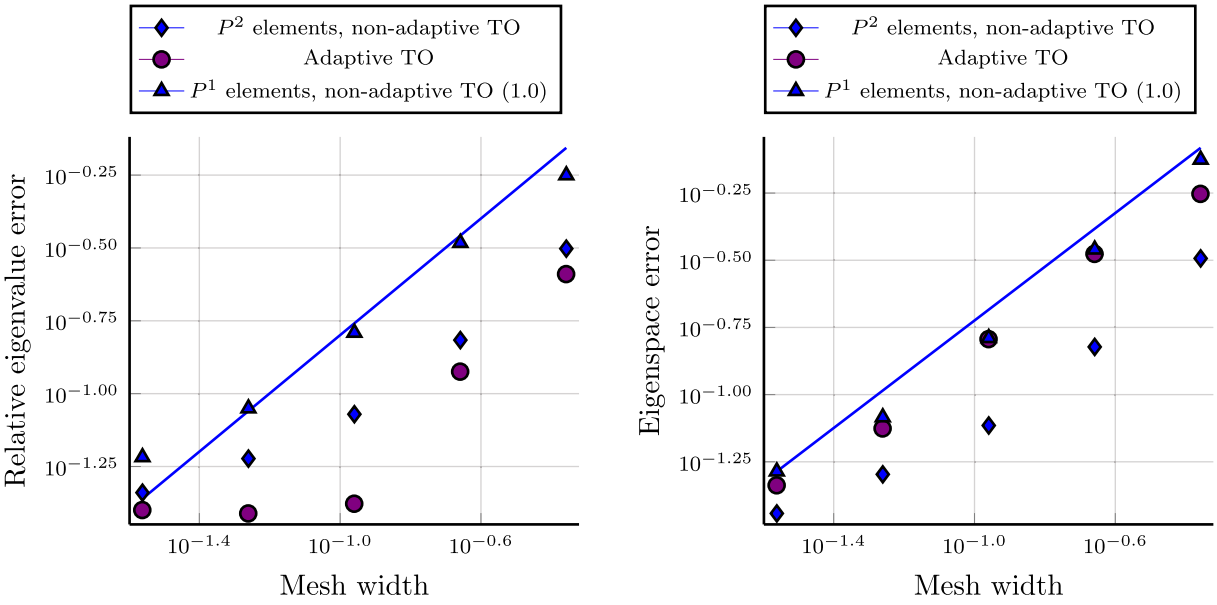

FIGURE 6. Cylinder flow: errors in the first nontrivial eigenvalue (*left*) and corresponding eigenspace (*right*) of the dynamic Laplacian for the TO approaches. The slopes of the corresponding lines are given in brackets in the legends.

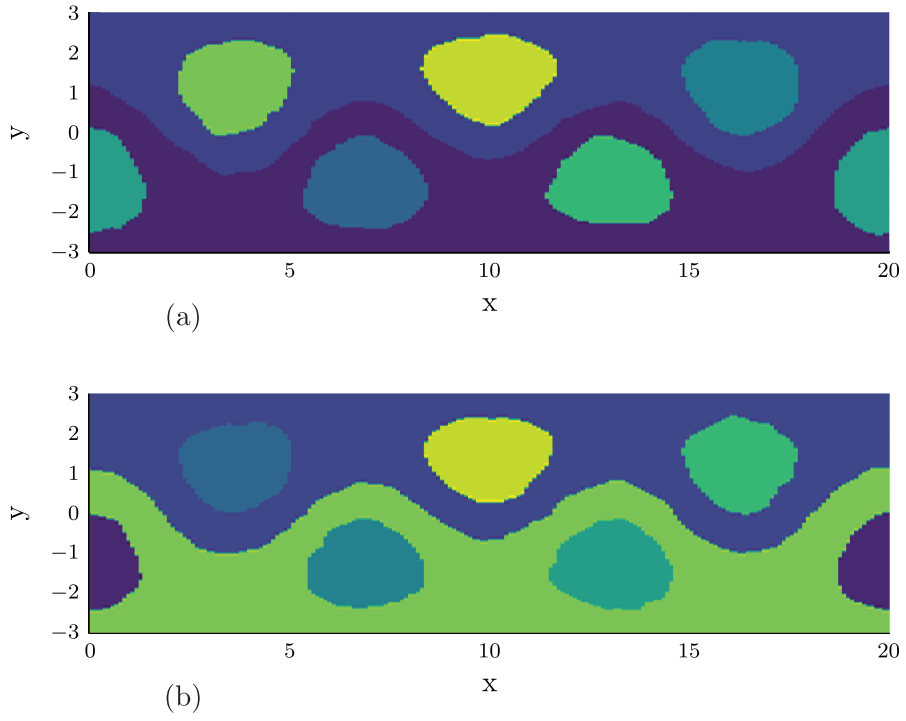

FIGURE 7. Bickley jet: comparison of TO method with P^1 and P^2 elements. (a) P^1 elements on 71×22 mesh (1540 nodes). (b) P^2 elements on 41×13 mesh (2000 nodes).

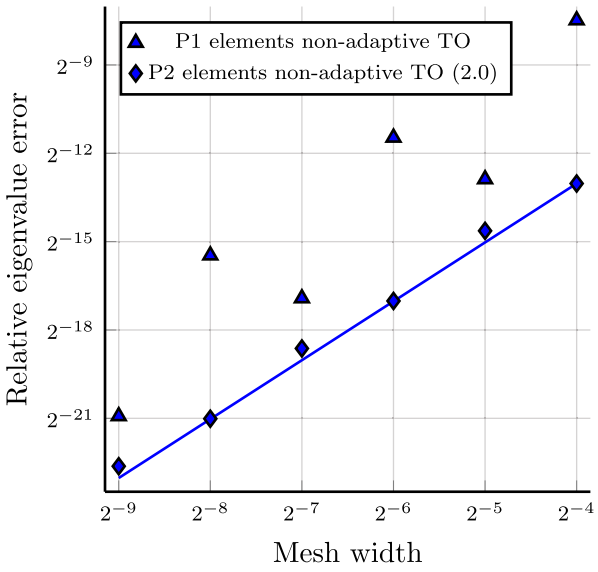

FIGURE 8. 1D shift map: errors in the first calculated eigenvalue. The slope of the corresponding line is given in brackets in the legend.

4.2. TO convergence rates

Determining the theoretical convergence orders of the TO-methods is an outstanding task. In order to investigate whether we can expect higher-order convergence, we consider the collocation-based non-adaptive TO approach on a simple example in one dimension. In this example we can largely isolate the dynamics from the errors to focus on errors arising from translations of the basis functions.

The only volume-preserving diffeomorphisms of the circle $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ to itself are rigid rotations. Let $T : \mathbb{S}^1 \to \mathbb{S}^1$ be given by $T(x) = (x + \alpha) \mod 1$. Rigid rotations commute with the Laplace operator; thus $\Delta^{\text{dyn}} = (\Delta + T_1^* \Delta T_{1,*})/2 = \Delta$, and the dynamic Laplacian is equal to the static Laplacian. The first nontrivial eigenspace of the dynamic Laplacian is therefore spanned by $v(x) = \sqrt{2}\sin(2\pi x)$ and $u(x) = \sqrt{2}\cos(2\pi x)$ which are orthonormal eigenvectors for the eigenvalue $\lambda_1 = 4\pi^2$.

Any errors in the corresponding discrete bilinear form

$$\tilde{a}_h(u, v) = \frac{1}{2}\left(a(u, v) + a\left(I_h T u, I_h T v\right)\right)$$

therefore arise solely from discretisation errors related to the rotation of the φ_i by α. In our experiments, we consider $\alpha = 0.15$. For the non-adaptive collocation TO approach, denote the leading nontrivial eigenvalue of Δ^{dyn} by $\tilde{\lambda}_{h,1}$. We now look at the order of convergence for the error $\frac{|\tilde{\lambda}_{h,1} - \lambda_1|}{\lambda_1}$ as $h \to 0$.

The convergence rates in Figure 8 are identical to those found for the standard map in Figure 5. The CG approach in this example is simply computing the eigenvectors and eigenvectors of the standard Laplace operator and therefore unsurprisingly one recovers the theoretical orders of convergence. These numerical results suggest that when using the non-adaptive TO method, we cannot expect an asymptotically higher convergence rate for P^2 elements in comparison to P^1-elements even for very simple flow maps.

5. CONCLUSION

We compared the use of P^1 and P^2 elements in collocation-based CG and TO approximations of the dynamic Laplacian. In the CG approach applied to weakly nonlinear dynamics, P^2 elements can significantly reduce the

computational cost by providing an asymptotically higher order of convergence. A benefit of P^1 elements for the CG approach is that a first-order method of quadrature can be used in the discretization, whereas using first-order quadrature for P^2 elements results in a singular mass matrix. This does not affect the asymptotic order of convergence, but nevertheless introduces a constant factor which may be relevant in some low-data cases. It seems promising to investigate how to locally adapt the element order and/or the element diameter (in the spirit of p- or hp-adaptive finite element methods).

In the non-adaptive TO approach, there seems to be little benefit gained by using P^2 as opposed to P^1 elements. The adaptive TO is inherently P^1-based, and thus does not benefit from a P^2 discretization either. In general, numerical experiments suggest that collocation-based TO approaches have equal (or in some cases better) rates of convergence compared to the P^1 CG discretisation.

The P^2 CG discretisation generally had a higher order of convergence, though the numerically observed convergence rates varied significantly when the quadrature order was changed or when applied to more nonlinear and hyperbolic dynamics. It is difficult to compare the CG and TO approaches in general as the former relies on being able to calculate derivatives of the flow map, whereas the latter is purely data based. This makes the TO approach applicable to some cases where the CG approach cannot be used.

A hindrance to using the TO approach for finely resolved meshes is the fact that here the call to `eigs` takes much longer compared to when one uses the CG discretisation on the same mesh. We suspect that this is due to the fact that unlike in the TO approach, the CG discretisation preserves the banded structure of the stiffness matrix. More work is needed to determine how the eigenproblem can be solved efficiently in this case, or whether it is possible to avoid the eigenproblem completely but still be able to compute coherent sets. More work is also needed to determine the true rates of convergence of eigenvalues and eigenvectors. We proved that they do converge for the non-adaptive TO approach, but were only able to conjecture what the true rates are based on numerical experiments. It also remains to be seen whether Galerkin TO approaches can be modified to be computationally efficient.

Appendix A. Proofs

Throughout all proofs, C refers to a constant depending only on the mesh and dimension.

A.1. Proof of Lemma 4.1

The following proof of Lemma 4.1 uses ideas from the proof of Lemma 2.1 from [4]. We start with a helper lemma:

Lemma A.1. *Let $\left\{\mathcal{T}_{h'}^0\right\}_{h'>0}$ and $\left\{\mathcal{T}_h^1\right\}_{h>0}$ be quasi-uniform meshes on d-dimensional open sets Ω_0 and Ω_1 respectively. Let $F:\Omega_0\to\Omega_1$ be smooth on Ω_0 with smooth extension to the boundary. Then there exists $C>0$ so that for all $h,h'>0$ and $\tau'\in\mathcal{T}_{h'}^0$ it holds that:*

$$\mathcal{N}(\tau') := \left|\left\{\tau\in\mathcal{T}_h^1 : F(\tau')\cap\tau\neq\emptyset\right\}\right| \leq C\left(\frac{h'+h}{h}\right)^d.$$

Proof. By quasi-uniformity of the meshes, we can find positive constants a', A' so that for any triangle $\tau'\in\mathcal{T}_{h'}^0$ there is a point τ_x' so that:

$$B\left(\tau_x', a'h'\right) \subset \tau \subset B\left(\tau_x', A'h'\right), \tag{A.1}$$

where $B(x,r)$ is the open ball of radius r centered at x. This gives that where $y := F\left(\tau_x'\right)$:

$$F(\tau') \subset F\left(B\left(\tau_x', A'h'\right)\right) \subset B\left(y, \underbrace{\|DF\|_\infty A'}_{:=E} h'\right) \tag{A.2}$$

where the last inclusion follows from the mean-value theorem, and $\|DF\|_\infty$ exists as F is smoothly extensible to the boundary.

Now write a, A for the shape-regularity/quasi-uniformity constants of $\{\mathcal{T}_h^1\}_{h>0}$, so that a formula like (A.1) holds for $\{\mathcal{T}_h^1\}_{h>0}$ also.

Then if $\tau \in \mathcal{T}_h^1$ intersects $B(y, Eh')$, it must hold that $B(\tau_x, ah) \subset B(\tau_x, Ah) \subset B(y, Eh' + 2Ah) \subset B(y, c(h' + h))$ with $c := \max\{E, 2A\}$. As for different τ, the sets $B(\tau_x, ah)$ are disjoint, we get

$$\mathcal{N}(\tau') \leq \left(\frac{c(h' + h)}{ah}\right)^d$$

which gives the claim. □

Lemma A.2. *Let $\{\mathcal{T}_h\}_{h>0}$ be a family of quasi-uniform meshes on an open subset of d-dimensional space. Let S_h be the space of functions representable by P^k-Lagrange elements on the mesh. Then there exist $C, C' > 0$ so that for all $v \in S^h$:*

$$C_1\|v\|_{L^2}^2 \leq h^d \sum_p |v(p)|^2 \leq C_2\|v\|_{L^2}^2$$

where p ranges over the nodes of the triangulation.

Proof. This follows directly from well-known results about the spectrum of mass-matrices, see p. 386 of [9] □

Proof of Lemma 4.1. Throughout the proof, C refers to a constant that does not depend on h or h' and whose exact value can change from line to line. Let $h' > h > 0$ and $v \in S_{h'}^0$. Then by Lemma A.2, and as F is volume-preserving,

$$\|I_h T v\|_{L^2}^2 \leq Ch^d \sum_p |v\left(F^{-1}(p)\right)|^2 =: (*)$$

where p ranges over the nodes of $\mathcal{T}_h^1$. Using Lemma A.1, we know that for any triangle $\tau' \in \mathcal{T}_h^0$, at most $C\left(\frac{h'+h}{h}\right)^d$ triangles from $\mathcal{T}_{h'}^1$ can intersect with $F(\tau')$; up to a constant factor this therefore bounds the number of vertices p in $\mathcal{T}_h^1$ for which $F^{-1}(p)$ lies in a given triangle. Moreover, as we are using P^1-Lagrange elements, $|v\left(F^{-1}(p)\right)|$ is bounded by $|v(p')|$ for some vertex p' of the triangle that contains $F^{-1}(p)$. This gives, using Lemma A.2 and the fact that $h' > h$, that:

$$\begin{aligned}(*) &\leq Ch^d \left(\frac{h' + h}{h}\right)^d \sum_{p'} |v(p')|^2 \\ &\leq C(h')^d \sum_{p'} |v(p')|^2 \\ &\leq C\|v\|_{L^2}^2\end{aligned}$$

where p' ranges over the nodes of $\mathcal{T}_{h'}^0$. □

A.2. Proof of Lemma 4.2

Proof. Without loss of generality, look at triangles in dimension 2; similar arguments apply to higher-dimensional meshes. Since $\|I_h v - v\|_{L^2} \leq C\|I_h v - v\|_{L^\infty}$ and v is uniformly continuous, we have that

$$\|I_h v - v\|_{L^2} \to 0 \quad \text{as} \quad h \to 0.$$

What remains to be shown is $\|\nabla(I_h v - v)\|_{L^2} \to 0$.

We first show that $\|\nabla I_h v\|_{L^\infty} \leq C\|\nabla v\|_{L^\infty}$ (where C does not depend on h). It is enough to prove this for any triangle τ in the mesh $\mathcal{T}_h^1$. Assume first that the triangle τ has vertices $0, e_1$ and e_2, where e_1, e_2 are the canonical unit vectors in $\mathbb{R}^2$. Without loss of generality also $v(0) = 0$. Then $\nabla I_h v = (v(e_1), v(e_2))^T$.

By a mean value inequality[6], $\max\{|v(e_1)|, |v(e_2)|\} \leq \|\nabla v\|_{L^\infty(\tau)}$. Thus $\|\nabla I_h v\|_{L^\infty(\tau)} \leq \|\nabla v\|_{L^\infty(\tau)}$. Shape regularity immediately gives $\|\nabla I_h v\|_{L^\infty(\tau)} \leq C\|\nabla v\|_{L^\infty(\tau)}$ for general triangles, taking suprema over all triangles gives the claim.

By assumption v is C^2 except for on a nowhere dense set of measure zero. By standard FEM theory, $\|\nabla(I_h v - v)\|_{L^2(\Omega')} \to 0$ on all sub-meshes Ω' on which v is C^∞, and with $h \to 0$ we can choose Ω' so that $\ell^d(\Omega') \to \ell^d(\Omega)$, recalling that ℓ^d was the notation used for d-dimensional Lebesgue measure. Moreover, $|\nabla I_h v| < C$ almost everywhere, and hence discontinuities do not cause problems. It follows that $\|\nabla(I_h v - v)\|_{L^2} \to 0$. □

Acknowledgements. OJ and NS acknowledge support by the Priority Programme SPP 1881 Turbulent Superstructures of the Deutsche Forschungsgemeinschaft and a Universities Australia/DAAD travel grant. GF is partially supported by an ARC Discovery Project and a joint Universities Australia/DAAD travel award. GF thanks the Faculty of Mathematics at the Technical University Munich for hospitality during his research visits. We would like thank Christian Ludwig for the reference to [4] used in the proof of Lemma 4.1. We also thank Alavaro de Diego and Daniel Karrasch for helpful comments and their contributions to the `CoherentStructures.jl` package.

References

[1] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe and H. van der Vorst, Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide. SIAM (2000).
[2] U. Banerjee and J.E. Osborn, Estimation of the effect of numerical integration in finite element eigenvalue approximation. *Numer. Math.* **56** (1989) 735–762.
[3] P. Bogacki and L.F. Shampine, An efficient Runge–Kutta (4, 5) pair. *Comput. Math. App.* **32** (1996) 15–28.
[4] X.-C. Cai, The use of pointwise interpolation in domain decomposition methods with nonnested meshes. *SIAM J. Sci. Comput.* **16** (1995) 250–256.
[5] K. Carlsson, KristofferC/JuAFEM.jl: finite element toolbox for Julia. Available from: https://github.com/KristofferC/JuAFEM.jl (2020).
[6] E.B. Davies, Metastable states of symmetric Markov semigroups II. *J. London Math. Soc.* **s2–26** (1982) 541–556.
[7] M. Dellnitz and O. Junge, On the approximation of complicated dynamical behavior. *SIAM J. Numer. Anal.* **36** (1999) 491–515.
[8] P. Deuflhard and M. Weber, Robust Perron cluster analysis in conformation dynamics. *Linear Algebra Appl.* **398** (2005) 161–184.
[9] A. Ern and J. Guermond, Theory and Practice of Finite Elements. In Vol 159 of *Applied Mathematical Sciences*. Springer Science & Business Media (2004).
[10] L. Evans, Partial Differential Equations. In Vol 19 of *Graduate Studies in Mathematics*. American Mathematical Society (1997).
[11] G. Froyland, Statistically optimal almost-invariant sets. *Phys. D: Nonlinear Phenom.* **200** (2005) 205–219.
[12] G. Froyland, Dynamic isoperimetry and the geometry of Lagrangian coherent structures. *Nonlinearity* **28** (2015) 3587–3622.
[13] G. Froyland and O. Junge, On fast computation of finite-time coherent sets using radial basis functions. *Chaos: Interdiscip. J. Nonlinear Sci.* **25** (2015) 087409.
[14] G. Froyland and O. Junge, Robust FEM-based extraction of finite-time coherent sets using scattered, sparse, and incomplete trajectories. *SIAM J. Appl. Dyn. Syst.* **17** (2018) 1891–1924.
[15] G. Froyland and E. Kwok, A dynamic Laplacian for identifying Lagrangian coherent structures on weighted Riemannian manifolds. To appear in: *J. Nonlinear. Sci.* https://doi.org/10.1007/s00332-017-9397-y (2017).
[16] G. Froyland and K. Padberg, Almost-invariant sets and invariant manifolds-connecting probabilistic and geometric descriptions of coherent structures in flows. *Phys. D: Nonlinear Phenom.* **238** (2009) 1507–1523.
[17] G. Froyland, S. Lloyd and N. Santitissadeekorn, Coherent sets for nonautonomous dynamical systems. *Phys. D: Nonlinear Phenom.* **239** (2010) 1527–1541.
[18] G. Froyland, C. Rock and K. Sakellariou, Sparse eigenbasis approximation: multiple feature extraction across spatiotemporal scales with application to coherent set identification. *Commun. Nonlinear Sci. Numer. Simul.* **77** (2019) 81–107.

[6] As v is only piecewise C^∞, we note that we must apply a piecewise version of the mean value inequality. But it can be seen that this holds by considering $f(t) := v(te_1)$ and noting that f is piecewise smooth and $|f'(t)| \leq \|\nabla v\|_{L^\infty}$. Thus from the fundamental theorem of calculus we get $|f(1) - f(0)| \leq \int_0^1 |f'(t)|\mathrm{d}t$, implying $|v(e_1)| \leq \|v\|_{L^\infty}$ (and similarly for other unit vectors).

[19] I.M. Gelfand and S.V. Fomin, Calculus of Variations. Prentice-Hall, Inc. (1963).
[20] D. Karrasch and J. Keller, A geometric heat-flow theory of Lagrangian coherent structures. *J. Nonlinear Sci.* **30** (2020) 1849–1888.
[21] T. Kato, Perturbation Theory for Linear Operators, reprint of the 2nd edition. In: *Classics in Mathematics*. Springer (1995).
[22] A. Keselman *et al.*, JuliaGeometry/VoronoiDelaunay.jl: fast and robust Voronoi & Delaunay tesselation creation with Julia. Available from: https://github.com/JuliaGeometry/VoronoiDelaunay.jl (2020).
[23] C. Rackauckas and Q. Nie, DifferentialEquations.jl – a performant and feature-rich ecosystem for solving differential equations in Julia. *J. Open Res. Softw.* **5** (2017) 15.
[24] I.I. Rypina, M.G. Brown, F.J. Beron-Vera, H. Koçak, M.J. Olascoaga and I.A. Udovydchenkov, On the Lagrangian dynamics of atmospheric zonal jets and the permeability of the stratospheric polar vortex. *J. Atmos. Sci.* **64** (2007) 3595–3610.
[25] G. Strang and G.J. Fix, An Analysis of the Finite Element Method. In: *Prentice Hall Series in Automatic Computation*. Prentice-Hall Englewood Cliffs, NJ (1973).