

Learning Dynamical Systems with the Spectral Exterior Calculus*

Suddhasattwa Das[†], Dimitrios Giannakis[‡], Yanbing Gu[‡], and Joanna Slawinska[‡]

Abstract. We present a data-driven framework for learning dynamical systems on compact Riemannian manifolds based on the spectral exterior calculus (SEC). This approach represents vector fields as linear combinations of frame elements constructed using the eigenfunctions of the Laplacian on smooth functions, along with their gradients. Such reconstructed vector fields generate dynamical flows that consistently approximate the true system, while being compatible with the nonlinear geometry of the manifold. The data-driven implementation of this framework utilizes embedded data points and tangent vectors as training data, along with a graph-theoretic approximation of the Laplacian. In this paper, we prove the convergence of the SEC-based reconstruction in the limit of large data. Moreover, we illustrate the approach numerically with applications to dynamical systems on the unit circle and the 2-torus. In these examples, the reconstructed vector fields compare well with the true vector fields, in terms of both pointwise estimates and generation of orbits.

Key words. reconstruction of dynamical systems, exterior calculus, vector fields, kernel methods, Laplace–Beltrami operators, frame theory

MSC codes. 37Mxx, 37Nxx, 53Z50, 62Jxx, 05C90

1. Introduction.

$$\dot{x}(t) = V|_{x(t)}$$

generated by a vector field $V: \mathcal{M} \rightarrow T\mathcal{M}$ on a manifold \mathcal{M} , a problem of considerable practical importance is to approximate (learn) the dynamics using samples of the state $x(t) \in \mathcal{M}$, sometimes indirectly via measurements. Over the years, a large number of parametric and non-parametric data-driven techniques have been developed to perform this task.

Parametric methods seek an optimal approximation to the generating vector field V of the dynamics within a parameterized family. In essence, the choice of parametric family amounts to assuming a specific form for the dynamical laws, and the learning task is mainly to determine parameters within the prescribed class. Parametric methods are particularly effective in applications with significant prior knowledge about the governing laws (which can constrain the class of the learned model), and may be combined with Bayesian inference techniques [70] to yield uncertainty-quantified parameter estimates. Some examples of the utility of parametric methods in specific domains are in the modeling of epidemics [3, 10] and forecasting and estimation of various aspects of geophysical fluid dynamics [60, 52, e.g.].

*Submitted to the editors DATE.

Funding: Dimitrios Giannakis acknowledges support from the U.S. Department of Defense, Basic Research Office under Vannevar Bush Faculty Fellowship grant N00014-21-1-2946 and the U.S. Office of Naval Research under MURI grant N00014-19-1-242. Yanbing Gu was supported as a PhD student from these grants. Dimitrios Giannakis and Joanna Slawinska acknowledge support from the U.S. Department of Energy grant DE-SC0025101.

[†]Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX (suddas@ttu.edu)

[‡]Department of Mathematics, Dartmouth College, Hanover, NH (dimitrios.giannakis@dartmouth.edu, yanbing.gu.gr@dartmouth.edu, joanna.m.slawinska@dartmouth.edu)

Our focus in this paper will be on non-parametric methods. These approaches do not rely on any foreknowledge about the explicit format of the dynamical laws, and instead leverage data to approximate V within a formally infinite-dimensional hypothesis space. This provides higher flexibility, sometimes with universal approximation guarantees [39, 56], which may overcome structural model errors associated with parametric approaches. At the same time, non-parametric methods tend to have higher training data requirements stemming from high dimensionality of the hypothesis spaces and ambient data spaces involved in real-world applications. Another major challenge is to ensure that the methods are refinable, i.e., they produce asymptotically consistent results with increasing amounts of training data.

There have been several approaches to non-parametric estimation of dynamical systems. They have diverse mathematical and computational foundations, including Fourier averaging [46, 23], non-parametric regression [48, 33, 69, 72, 1, 34, 43], neural network approximation [31, 78, 77, 9, 35, 51, 54, 61], reservoir computing [40, 32, 30], Markov state models [74, 21, 27], and linear evolution operators [62, 24, 68, 55]. Sparse identification of nonlinear dynamics (SINDy) [12] incorporates sparsity-inducing penalty terms in regression models to yield approximations of the dynamical vector field which are sparse in a user-supplied dictionary of functions. Physics-informed neural networks (PINNs) [65] and neural operators [45] are neural-network based approaches that have been successful in learning infinite-dimensional systems governed by partial differential equations. Meanwhile, methods based on Koopman and transfer operators leverage the intrinsically linear action of a dynamical flow on vector spaces of observables to build linear models for the evolution of the components of the state vector (treated in this context as observables) [67, 75, 8, 41, 28, 11, 22, 42, 44, 19]. These methods can be seen in a similar vein as the lift-and-learn approach [64], which seeks to embed the nonlinear state space evolution to a higher-dimensional space (e.g., a space of observables or measures) in which dynamical nonlinearities are brought into a standard form, or eliminated altogether. Recent work [18] has studied approximation barriers faced by data-driven techniques utilizing spectral decompositions of Koopman operators.

Many of the techniques outlined above effectively access the vector field V through its representation as a vector-valued function $\vec{V}: \mathcal{M} \rightarrow \mathbb{R}^d$ taking values in a Euclidean space \mathbb{R}^d in which the state space manifold \mathcal{M} is embedded. While every sufficiently smooth embedding $F: \mathcal{M} \hookrightarrow \mathbb{R}^d$ induces a vector-valued function $\vec{V} = F_*V$ through the pushforward map F_* on vector fields, such a representation is not intrinsic to the dynamical system as it depends on the embedding map. To put it differently, while it is true that every vector field V induces an \mathbb{R}^d -valued function \vec{V} under an embedding, not every \mathbb{R}^d -valued function on \mathcal{M} is realized by the pushforward of a vector field. Correspondingly, learning \vec{V} in a hypothesis space of \mathbb{R}^d -valued functions carries a risk of producing structurally inconsistent results.

1.1. Our approach. We take a distinct path of learning the generator of the dynamics in its intrinsic geometric and operator-theoretic forms:

- Geometrically, the dynamical vector field corresponds to a section $V: \mathcal{M} \rightarrow T\mathcal{M}$ of the tangent bundle of \mathcal{M} .
- From an operator-theoretic point of view, V acts as a linear operator on the space of continuously differentiable, real-valued functions on \mathcal{M} ,

$$(1.1) \quad Vf = \langle V, df \rangle, \quad \forall f \in C^1(\mathcal{M}),$$

obeying the Leibniz rule,

$$V(fg) = (Vf)g + f(Vg), \quad \forall f, g \in C^1(\mathcal{M}).$$

Here, $df: \mathcal{M} \rightarrow T^*\mathcal{M}$ is the exterior derivative or differential of f (and therefore a covector field), and $\langle \cdot, \cdot \rangle$ denotes the natural pairing between tangent vectors and dual vectors on \mathcal{M} . When \mathcal{M} is equipped with a Riemannian metric \mathbf{g} , we equivalently have

$$(1.2) \quad Vf = V \cdot \nabla f \equiv \langle V, \nabla f \rangle_{\mathbf{g}},$$

where ∇ and $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ are the gradient operator and Riemannian inner product on tangent vectors induced by \mathbf{g} , respectively.

In more detail, our approach is to learn V in a hypothesis space spanned by a dictionary of intrinsic vector fields. The training data consists of samples $\{y_n\}_{n=1}^N \subset \mathbb{R}^d$ which are the images $y_n = F(x_n)$ of points $\{x_n\}_{n=1}^N \subset \mathcal{M}$ under an unknown embedding $F: \mathcal{M} \rightarrow \mathbb{R}^d$, along with samples $\vec{V}(x_n) = F_*V|_{x_n}$ (“arrows”) of the corresponding pushforward vector field. We build this hypothesis space using the Spectral Exterior Calculus (SEC) [7]—a data-driven technique that derives frames (overcomplete bases) of Hilbert spaces of vector fields and differential forms on Riemannian manifolds using spectral data of the Laplace–Beltrami operator $\Delta = -\operatorname{div} \circ \nabla$ on functions. This leads to an approximation of the form

$$(1.3) \quad V \approx V^{(\ell)} = \sum_{ij} b_{ij} B_{ij}, \quad b_{ij} \in \mathbb{R}, \quad B_{ij} = \phi_i \nabla \phi_j,$$

where $B_{ij}: \mathcal{M} \rightarrow T\mathcal{M}$ are frame elements given by products of Laplace–Beltrami eigenfunctions $\phi_i \in C^\infty(\mathcal{M})$ and their gradients, and b_{ij} are the corresponding expansion coefficients. In (1.3), the (ℓ) superscript collectively represents various approximation parameters (such as the number of frame elements employed), which will be made precise in subsequent sections.

To build the frame elements B_{ij} , the SEC leverages (i) the *carré du champ* identity [4] obeyed by the Laplacian,

$$(1.4) \quad \Delta(fg) = (\Delta f)g + f(\Delta g) - 2\nabla f \cdot \nabla g, \quad \forall f, g \in C^\infty(\mathcal{M}),$$

which allows computing Riemannian inner products between gradient vector fields from the action of Δ on functions; (ii) the $C^\infty(\mathcal{M})$ -module structure of spaces of vector fields, which allows building generating sets of vector fields from products of the form $f_i \nabla f_j$ from a sufficiently rich collection of functions $f_i \in C^\infty(\mathcal{M})$; and (iii) graph-theoretic algorithms (e.g., [16, 73, 37, 76]) for approximation of the Laplacian induced by the embedding F in data space using the dataset $\{y_n\}_{n=1}^N$. Furthermore, the SEC provides a representation of vector fields as operators on functions, allowing approximation of the coordinate representation of \vec{V} in the embedding space,

$$\vec{V}(y) \approx \vec{V}^{(\ell)}(y) = F_*V^{(\ell)}(y) \equiv V^{(\ell)}F(x), \quad y = F(x),$$

where $V^{(\ell)}$ acts on F componentwise as a linear operator.

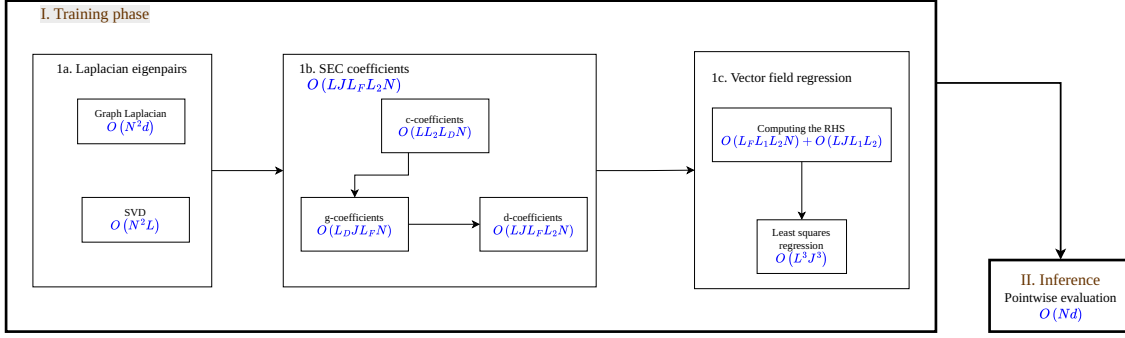


Figure 1. Schematic overview of the training and prediction (inference) phases of our numerical procedure. The training phase has three parts: eigendecomposition of the graph Laplacian (1a), computation of SEC coefficients (1b), and solution of the vector field regression problem (1c). Note that parts 1a and 1b use only the spatial locations of the training data on the manifold, and part 1c uses both locations and vector field data. As a result, parts 1a and 1b may be skipped if the vector field is changed but the sampled points on the manifold are not. The boxes show numerical time complexity estimates based on brute-force linear algebra computations. The actual cost may be improved if specialized numerical techniques are used such as low-rank kernel operator approximations in the inference phase.

In this approximation, the components of $\vec{V}^{(\ell)}$ lie in a space of C^1 functions created by a kernel integral operator on \mathcal{M} induced by the training data. The vector-valued function $\vec{V}^{(\ell)}$ thus generates a smooth dynamical system on \mathbb{R}^d ,

$$(1.5) \quad \dot{y}(t) = \vec{V}^{(\ell)}|_{y(t)},$$

that was learned in a hypothesis space of intrinsic vector fields on the manifold \mathcal{M} . See Figure 1 for a schematic overview of the training and prediction phases of our approach.

Our main result, [Theorem 5.1](#), establishes the precise nature of the convergence of the SEC-based approximation of vector fields described in this paper. The statement of the theorem makes some general assumptions on the vector field V and manifold \mathcal{M} supporting it ([Assumption 1](#)), the embedding of the manifold in Euclidean space ([Assumption 2](#)), and the nature of its sampling ([Assumption 3](#)). The numerical approximations employed are based on kernel methods that make some general assumptions on the kernel ([Assumption 4](#)), but otherwise do not assume prior knowledge of the structure of the manifold, the dynamical vector field, or the sampling distribution of the training data. In [Proposition 5.6](#), we give an error estimate for dynamical orbits (solutions of (1.5)) generated by vector field approximations (including but not limited to SEC-based approximations) that converge in a Hodge norm employed in [Theorem 5.1](#).

We test our approach with numerical experiments involving dynamical systems on manifolds embedded in Euclidean space, including rotation systems on circles and tori and a Stepanoff flow on the 2-torus [59]. The examples demonstrate accurate reconstruction of the dynamical vector field on the embedded manifold as well as stable out-of-sample evaluation away from the manifold. We also show comparisons between the solutions to initial-value problems associated with the true vector fields and the SEC approximations. The results

illustrate the efficacy of the SEC in reproducing a variety of dynamical behaviors, including metastable fixed points and periodic, quasiperiodic, and aperiodic orbits.

1.2. Plan of the paper. In [section 2](#), we introduce the class of systems under study and establish basic notation. In [section 3](#), we describe our SEC-based scheme for approximation of the dynamical vector field assuming prior knowledge of the spectrum of the Laplace–Beltrami operator and its eigenfunctions. In [section 4](#), we present a data-driven formulation of that scheme, which casts the approximation framework of [section 3](#) in a basis of approximate Laplace–Beltrami eigenfunctions learned using graph-theoretic techniques. We analyze the convergence properties of these approximations in [section 5](#). [Section 6](#) presents our numerical experiments. Our primary conclusions and a discussion on future directions are included in [section 7](#). The paper contains appendices on numerical methods ([Appendix A](#)) and the proof of [Proposition 5.6](#) ([Appendix B](#)). A Python implementation of our approach, reproducing the numerical results presented in this paper, can be found at the repository <https://github.com/ygu626/SEC>.

2. Preliminaries and notation. We consider a dynamical flow $\Phi^t: \mathcal{M} \rightarrow \mathcal{M}$, $t \in \mathbb{R}$, on a differentiable manifold \mathcal{M} , generated by a vector field $V: \mathcal{M} \rightarrow T\mathcal{M}$. The system is observed by means of a map $F: \mathcal{M} \rightarrow \mathbb{R}^d$ into a d -dimensional Euclidean data space. We shall make the following standing assumptions on the manifold, dynamical system, and observation map.

Assumption 1. *The manifold \mathcal{M} is closed (i.e., smooth, compact, and boundaryless) and orientable, and the vector field V is continuous.*

Assumption 2. *The observation map $F: \mathcal{M} \rightarrow \mathbb{R}^d$ is a C^∞ embedding.*

We will let $F_{*x}: T_x\mathcal{M} \rightarrow T_{F(x)}\mathbb{R}^d \cong \mathbb{R}^d$ denote the pushforward map on tangent vectors at $x \in \mathcal{M}$ into Euclidean vectors in data space, and $F_*: T\mathcal{M} \rightarrow T\mathbb{R}^d \cong \mathbb{R}^{2d}$ the corresponding map on the tangent bundle of \mathcal{M} . We will also let $Y = F(\mathcal{M}) \subset \mathbb{R}^d$ denote the image of the manifold in data space under the embedding F .

Within this setup, our training data consists of:

1. Samples $\{y_n \in \mathbb{R}^d\}_{n=1}^N$ in data space given by the images $y_n = F(x_n)$ of a collection of points $\{x_n\}_{n=1}^N$ on the manifold for some $N \in \mathbb{N}$.
2. Corresponding samples $\{\vec{v}_n \in \mathbb{R}^d\}_{n=1}^N$ of the pushforward of the dynamical vector field in data space, where $\vec{v}_n = F_{*x_n}V(x_n)$.

Note that we do not assume any explicit knowledge of the manifold points x_n , the vector field V , or the observation map F . In what follows, $C^r(\mathcal{M}, T\mathcal{M})$, $r \in \mathbb{N}_0$, will denote the space of r -times continuously differentiable vector fields on \mathcal{M} , i.e., the space of r -times continuously differentiable sections of the tangent bundle $T\mathcal{M}$.

Let \mathbf{g} be the Riemannian metric tensor on \mathcal{M} induced by the embedding F , i.e., $\mathbf{g}_x(u, v) = F_{*x}u \cdot F_{*x}v$ for $u, v \in T_x\mathcal{M}$, where $\cdot: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ is the standard Euclidean inner product. The metric \mathbf{g} induces a gradient operator $\nabla: C^1(\mathcal{M}) \rightarrow C(\mathcal{M}, T\mathcal{M})$, a finite volume measure ν on \mathcal{M} , and a divergence operator $\text{div}: C^1(\mathcal{M}, T\mathcal{M}) \rightarrow C(\mathcal{M})$ such that

$$\int_{\mathcal{M}} f \text{div} W \, d\nu = - \int_{\mathcal{M}} \langle \nabla f(x), W(x) \rangle_{\mathbf{g}(x)} \, d\nu(x), \quad \forall f \in C^1(\mathcal{M}), \quad \forall W \in C^1(\mathcal{M}, T\mathcal{M}).$$

The Laplace–Beltrami operator $\Delta: C^2(\mathcal{M}) \rightarrow C(\mathcal{M})$ induced by \mathbf{g} is defined as $\Delta = -\text{div} \circ \nabla$,

and has an associated orthonormal basis $\{\phi_j\}_{j \in \mathbb{N}_0}$ of $L^2(\nu)$ consisting of eigenfunctions,

$$(2.1) \quad \Delta \phi_j = \lambda_j \phi_j, \quad \phi_j \in C^\infty(\mathcal{M}), \quad \lambda_j \geq 0,$$

where the corresponding eigenvalues are ordered as $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots \nearrow \infty$. In what follows, $\varphi_j \in C^\infty(\mathbb{R}^d)$ will be extensions of the eigenfunctions ϕ_j ; i.e., $\varphi_j(y) = \phi_j(x)$ for $y = F(x)$. In [subsection 4.1](#), we will show how these extensions φ_j are constructed using kernel integral operators.

Next, we consider the sampling distribution of our training data.

Assumption 3. *The points $x_1, x_2, \dots \in \mathcal{M}$ underlying our training data lie in an equidistributed sequence with respect to a probability measure μ with a strictly positive density $\sigma := \frac{d\mu}{d\nu} \in C^\infty(\mathcal{M})$; that is, $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_{\mathcal{M}} f \sigma d\nu$, for all $f \in C(\mathcal{M})$.*

Let δ_x denote the Dirac δ -measure supported at $x \in \mathcal{M}$. [Assumption 3](#) is equivalent to weak convergence of the empirical sampling measures $\mu_N := \frac{1}{N} \sum_{n=1}^N \delta_{x_n}$ to μ . For the rest of the paper, $\nu_* := F_*\nu$ will be the pushforward of the volume measure of the manifold to data space \mathbb{R}^d . We also define the Hilbert space $L^2(\nu_*; \mathbb{R}^d)$ of equivalence classes of \mathbb{R}^d -valued functions on data space \mathbb{R}^d with respect to ν_* to have inner product $\langle f, g \rangle_{L^2(\nu_*; \mathbb{R}^d)} = \int_{\mathbb{R}^d} f(y) \cdot g(y) d\nu_*(y)$. Note that this space includes the pushforwards of elements of \mathbb{H} under F , as well as other square-integrable \mathbb{R}^d -valued functions which are not tangent to the embedded manifold Y .

3. Vector field approximation framework. In this section, we describe the analytical aspects of our approach for approximating dynamical vector fields using the SEC.

3.1. Frames for Hilbert spaces of vector fields. We equip the space of continuous vector fields $C(\mathcal{M}, T\mathcal{M})$ with the Hodge inner product

$$(3.1) \quad \langle V, W \rangle_{\mathbb{H}} := \int_{\mathcal{M}} \langle V(x), W(x) \rangle_{\mathfrak{g}(x)} d\nu(x).$$

This is finite for all $V, W \in C(\mathcal{M}, T\mathcal{M})$ by compactness of \mathcal{M} . Taking the closure with respect to the norm $\|\cdot\|_{\mathbb{H}}$ induced by this inner product leads to a Hilbert space of vector fields which we denote by \mathbb{H} . Equivalently, elements $V \in \mathbb{H}$ may be characterized by the requirement that the ν -a.e. defined map $x \mapsto \|V(x)\|_{\mathfrak{g}_x}$ lies in $L^2(\nu)$.

As noted in [section 1](#), every $V \in C(\mathcal{M}, T\mathcal{M})$ acts as linear operator $V: C^1(\mathcal{M}) \rightarrow C(\mathcal{M})$ on continuously differentiable functions on the manifold via [\(1.1\)](#), or, equivalently, [\(1.2\)](#). In what follows, we will sometimes use the notation $V \triangleright f \equiv Vf$ to distinguish application of V as an operator on a function $f \in C^1(\mathcal{M})$ from evaluation $V(x) \in T_x\mathcal{M}$ as a section $V: \mathcal{M} \rightarrow T\mathcal{M}$ of the tangent bundle at a point $x \in \mathcal{M}$.

Over the years, a variety of numerical techniques for approximation of vector fields (as well as more general tensor fields) on manifolds have been developed, including methods based on simplicial complexes [\[38, 25\]](#), finite elements [\[2\]](#), wavelets [\[47\]](#), and radial basis functions [\[36\]](#). The SEC scheme [\[7\]](#) employed in this work performs vector field approximation in *frames* for the Hilbert space of vector fields \mathbb{H} , constructed from eigenfunctions of the Laplace–Beltrami operator Δ . We first recall the definition of a frame of a Hilbert space [\[53\]](#).

Definition 3.1. A countable subset $\{u_i\}_{i \in \mathbb{I}}$ of a Hilbert space H is said to be a frame if there exist constants $C_1, C_2 > 0$ (called frame constants) such that

$$(3.2) \quad C_1 \|v\|_H^2 \leq \sum_{i \in \mathbb{I}} |\langle u_i, v \rangle_H|^2 \leq C_2 \|v\|_H^2, \quad \forall v \in H.$$

Intuitively, a frame is a generating set for H (in the sense of the lower bound in Definition 3.1) that is not “too overcomplete” (in the sense of the upper bound). In particular, the frame condition (3.2) implies:

- The operator $T: H \rightarrow \ell^2(\mathbb{I})$ where $Tv = (\langle u_i, v \rangle_H)_{i \in \mathbb{I}}$, called *analysis operator* is bounded, injective, and has closed range.
- The adjoint $T^*: \ell^2(\mathbb{I}) \rightarrow H$, called *synthesis operator*, is surjective (and bounded).
- $S := T^*T$, called *frame operator*, is a bounded operator on H with bounded inverse.
- $G := TT^*$, called *Gram operator*, is a bounded operator on $\ell^2(\mathbb{I})$ with closed range, and thus with bounded pseudoinverse, $G^+ = TS^{-2}T^*$.

Using these operators, we can express any vector $v \in H$ as the following linear combination of frame elements,

$$(3.3) \quad v = T^*G^+Tv = \sum_{i \in \mathbb{I}} v_i u_i, \quad v_i = \langle S^{-1}u_i, v \rangle_H.$$

While, in general, this expansion is not unique, the sequence of coefficients v_i in (3.3) has the minimal ℓ^2 norm among all sequences $(\tilde{v}_i)_i \in \ell^2(\mathbb{I})$ such that $v = \sum_{i \in \mathbb{I}} \tilde{v}_i u_i$. Note that if $\{u_i\}_{i \in \mathbb{I}}$ is an orthonormal basis of H then S and G are equal to identity operators and (3.3) reduces to a standard orthonormal basis expansion.

Returning to our SEC example, we build frames for \mathbb{H} that consist of vector fields of the form

$$(3.4) \quad B_{ij} := \phi_i \nabla \phi_j \in C(\mathcal{M}, T\mathcal{M}), \quad i \in \mathbb{N}_0, \quad j \in \mathbb{N}.$$

The following result [7, Theorem 4.3] guarantees that such fields form a frame so long as a sufficiently large, but finite, number of gradient fields $\nabla \phi_j$ are employed.

Theorem 3.2. Under Assumptions 1–3, there exists an integer $J \geq \dim(\mathcal{M})$ such that for every $J' \geq J$, the collection $\{B_{ij} : i \in \mathbb{N}_0, 1 \leq j \leq J'\}$ is a frame for \mathbb{H} .

Remark 3.3. A motivation of the frame element construction in (3.4) is that the space $C(\mathcal{M}, T\mathcal{M})$ of continuous vector fields on a closed smooth manifold is a finitely generated $C(\mathcal{M})$ -module. That is, there exists a finite collection of vector fields $W_1, \dots, W_J \in C(\mathcal{M}, T\mathcal{M})$ such that $V = \sum_{j=1}^J f_j W_j$ for every vector field $V \in C(\mathcal{M}, T\mathcal{M})$ for a collection $f_1, \dots, f_J \in C(\mathcal{M})$ of continuous functions. By results from differential geometry, unless \mathcal{M} is a parallelizable manifold (i.e., $T\mathcal{M}$ is diffeomorphic to the product manifold $\mathcal{M} \times \mathbb{R}^{\dim(\mathcal{M})}$), $C(\mathcal{M}, T\mathcal{M})$ does not have a basis as a $C(\mathcal{M})$ -module. That is, if \mathcal{M} is not parallelizable there exists no collection $\{W_j\}$ such that the functions f_j are unique for every $V \in C(\mathcal{M}, T\mathcal{M})$. For smooth closed manifolds it so happens that finitely many gradient fields $\nabla \phi_j$ obtained from Laplace–Beltrami eigenfunctions are a generating set for $C(\mathcal{M}, T\mathcal{M})$, but unless \mathcal{M} is parallelizable the frames of \mathbb{H} in Theorem 3.2 are necessarily overcomplete.

For simplicity, we shall henceforth denote the number J' in Theorem 3.2 by J . We denote the associated index set for the frame by $\mathbb{I} := \mathbb{N}_0 \times \{1, \dots, J\}$.

3.2. Operator representation. Being continuous vector fields, the frame elements B_{ij} are characterized by their action as linear operators on $C^1(\mathcal{M})$ functions. To evaluate this action, the SEC employs the carré du champ identity (1.4) for the Laplace–Beltrami operator:

$$(3.5) \quad \begin{aligned} B_{ij} \triangleright \phi_k &= \phi_i \nabla \phi_j \triangleright \phi_k = \phi_i (\nabla \phi_j \cdot \nabla \phi_k) \\ &= \frac{\phi_i}{2} ((\Delta \phi_j) \phi_k + \phi_j (\Delta \phi_k) - \Delta(\phi_j \phi_k)) = \frac{\phi_i}{2} ((\lambda_j + \lambda_k) \phi_j \phi_k - \Delta(\phi_j \phi_k)). \end{aligned}$$

From the above we see that the action of B_{ij} on the Laplace–Beltrami eigenfunctions ϕ_k (and, by linear extension, on linear combinations of eigenfunctions) is completely determined by the spectrum (eigenvalues) of Δ and the pointwise products of the corresponding eigenfunctions.

The operator representation of continuous vector fields also allows to compute the action of the analysis operator $T: \mathbb{H} \rightarrow \ell^2(\mathbb{I})$ on continuous vector fields $V \in C(\mathcal{M}, TM)$; that is, $TV = (V_{ij})_{(i,j) \in \mathbb{I}}$, where

$$(3.6) \quad V_{ij} = \langle \phi_i, \nabla \phi_j \cdot V \rangle_{L^2(\nu)} = \langle \phi_i, \nabla \phi_j \cdot V \rangle_{L^2(\nu)}.$$

Note that the second equality above reinterprets the coefficients V_{ij} as the the matrix elements $\langle \phi_i, V \triangleright \phi_j \rangle_{L^2(\nu)}$ in the Laplace–Beltrami eigenfunction basis of $L^2(\nu)$.

Equations (3.5) and (3.6) provide a route for using any consistent spectral approximation technique for the Laplacian to perform vector field computations. Note, in particular, that these expressions do not make use of extrinsic information to the Riemannian manifold $(\mathcal{M}, \mathbf{g})$ such as local tangent plane approximation in the (potentially high-dimensional) embedding space \mathbb{R}^d . Moreover, (3.6) links the representation of V through its coefficients under the analysis operator to its matrix representation as an operator on functions.

We now show how the coefficients V_{ij} are computable using information from the pushforward vector field $\vec{V}: Y \rightarrow \mathbb{R}^d$, $\vec{V} = F_* V$, underpinning our training data.

First, since $Y \subset \mathbb{R}^d$ is a submanifold of Euclidean space, the action of the pushforward map F_* can be expressed as linear operator application,

$$(3.7) \quad \vec{V}(y) = F_{*x}(V(x)) = (V \triangleright F)(x), \quad y = F(x),$$

where V acts on F componentwise. Let $F = \sum_{k=0}^{\infty} F_k \phi_k$, F_k by the expansion of the embedding in the Laplace–Beltrami eigenfunction basis, where the expansion coefficients $F_k \in \mathbb{R}^d$ are given by the componentwise integrals $F_k = \int_{\mathcal{M}} \phi_k(x) F(x) d\nu(x)$, and the truncated sum $\sum_{k=0}^{K-1} F_k \phi_k$ converges in $C^r(\mathcal{M}; \mathbb{R}^d)$ for every $r \in \mathbb{N}_0$ by C^∞ smoothness of F . Defining

$$\hat{V}_{kl} = \langle (F_k \cdot \vec{V}) \circ F, \phi_l \rangle_{L^2(\nu)}$$

and noticing that

$$V \phi_j = V \cdot \nabla \phi_j = (F_* V \cdot F_* \nabla \phi_j) \circ F = (\vec{V} \circ F) \cdot (\nabla \phi_j \triangleright F),$$

we get

$$(3.8) \quad V_{ij} = \langle \phi_i, (\vec{V} \circ F) \cdot (\nabla \phi_j \triangleright F) \rangle_{L^2(\nu)} = \sum_{k=0}^{\infty} \langle (F_k \cdot \vec{V}) \circ F, B_{ij} \triangleright \phi_k \rangle_{L^2(\nu)} = \sum_{k,l=0}^{\infty} \hat{V}_{kl} \langle \phi_l, B_{ij} \triangleright \phi_k \rangle_{L^2(\nu)}.$$

Next, let $\{\varepsilon_{ij}\}_{(i,j) \in \mathbb{I}}$ be the standard orthonormal basis of $\ell^2(\mathbb{I})$ for an enumeration $\varepsilon_{i_1, j_1}, \varepsilon_{i_2, j_2}, \dots$ of the indices $(i, j) \in \mathbb{I} = \mathbb{N}_0 \times \{1, \dots, J\}$, and similarly let $\{\hat{\varepsilon}_{ij}\}_{(i,j) \in \hat{\mathbb{I}}}$ be the standard orthonormal basis of $\ell^2(\hat{\mathbb{I}})$ for the index set $\hat{\mathbb{I}} = \mathbb{N}_0^2$. Define the linear map $\mathcal{F}: L^2(\nu_*; \mathbb{R}^d) \rightarrow \ell^2(\hat{\mathbb{I}})$ as

$$\mathcal{F}\vec{W} = \sum_{(k,l) \in \hat{\mathbb{I}}} \langle (F_k \cdot \vec{W}) \circ F, \phi_l \rangle_{L^2(\nu)} \hat{\varepsilon}_{kl}.$$

Define also $\mathcal{D}: \text{dom}(\mathcal{D}) \rightarrow \ell^2(\mathbb{I})$ on a dense domain $\text{dom}(\mathcal{D}) \subset \ell^2(\hat{\mathbb{I}})$ through its matrix representation,

$$\langle \varepsilon_{ij}, \mathcal{D}\hat{\varepsilon}_{kl} \rangle_{\ell^2(\mathbb{I})} = \langle \phi_l, B_{ij} \triangleright \phi_k \rangle_{L^2(\nu)}.$$

We will verify in [section 5](#) that \mathcal{F} is bounded and its range is a subspace of $\text{dom}(\mathcal{D})$.

With these definitions, (3.8) becomes

$$V_{ij} = (\mathcal{D}\mathcal{F}\vec{V})_{ij}, \quad \forall (i, j) \in \mathbb{I}.$$

Inserting the last expression in (3.6), we deduce that every sequence $b \in \ell^2(\mathbb{I})$ that reconstructs V in the frame (i.e., $V = T^*b$) satisfies

$$(3.9) \quad Gb = \mathcal{D}\mathcal{F}\vec{V}.$$

[Equation \(3.9\)](#) provides the theoretical basis for setting up the vector field learning problem as a problem in linear regression. As we will see below, the Gram operator G in the left-hand side is computed using the SEC framework, whereas the vector of coefficients $\mathcal{D}\mathcal{F}\vec{V}$ in the right-hand side can be empirically approximated from the vector field samples $\vec{v}_n \in \mathbb{R}^d$.

3.3. Building the frame elements. We now describe various SEC constructions that will be used in our vector field approximation scheme. These constructions involve scalar coefficients derived from the spectral data (λ_j, ϕ_j) of the Laplacian, which we name c -, g -, and d - coefficients, respectively.

We first consider the c -coefficients,

$$(3.10) \quad c_{ijp} := \langle \phi_p, \phi_i \phi_j \rangle_{L^2(\nu)}, \quad \forall i, j, p \in \mathbb{N}_0.$$

These coefficients provide a representation of products of Laplace–Beltrami eigenfunctions,

$$\phi_i \phi_j = \sum_{p=0}^{\infty} c_{ijp} \phi_p, \quad \forall i, j, p \in \mathbb{N}_0,$$

where the summation over p giving $\phi_i(x)\phi_j(x)$ converges uniformly over $x \in \mathcal{M}$ by compactness of \mathcal{M} and C^∞ smoothness of the ϕ_j .

Next, we consider the g -coefficients,

$$(3.11) \quad g_{pjk} := \langle \phi_p, \nabla \phi_j \cdot \nabla \phi_k \rangle_{L^2(\nu)}, \quad \forall p, j, k \in \mathbb{N}_0.$$

These coefficients enable a representation of Riemannian inner products of gradient vector fields,

$$\nabla \phi_j \cdot \nabla \phi_k = \sum_{p=0}^{\infty} g_{pjk} \phi_p,$$

where the sums over p converge again uniformly on \mathcal{M} . Using (3.5), we get

$$(3.12) \quad g_{pjk} = \frac{1}{2} (\lambda_j + \lambda_k - \lambda_p) c_{jpk}.$$

Finally, the d -coefficients are defined as

$$(3.13) \quad d_{ijkl} = \langle \phi_l, B_{ij} \triangleright \phi_k \rangle_{L^2(\nu)}, \quad \forall i, k, l \in \mathbb{N}_0, \quad \forall j \in \{1, \dots, J\}.$$

These coefficients represent the action $B_{ij} \triangleright \phi_k$ of the frame elements on the basis functions, or, equivalently, the matrix elements of the \mathcal{D} operator from subsection 3.1, $\langle \varepsilon_{ij}, \mathcal{D} \hat{\varepsilon}_{kl} \rangle_{\ell^2(\hat{\mathbb{I}})} = d_{ijkl}$. Using again (3.5), we obtain

$$(3.14) \quad d_{ijkl} = \langle \phi_l, B_{ij} \triangleright \phi_k \rangle_{L^2(\nu)} = \langle \phi_i \phi_l, \nabla \phi_j \cdot \nabla \phi_k \rangle_{L^2(\nu)} = \sum_{p=0}^{\infty} c_{ilp} g_{pjk}.$$

A restriction of the d -coefficients lead to the matrix elements of the Gram operator $G: \ell^2(\mathbb{I}) \rightarrow \ell^2(\mathbb{I})$,

$$(3.15) \quad G_{ijkl} := \langle \varepsilon_{ij}, G \varepsilon_{kl} \rangle_{\ell^2} = \langle T^* \varepsilon_{ij}, T^* \varepsilon_{kl} \rangle_{\mathbb{H}} = d_{ijlk}, \quad (i, j), (k, l) \in \mathbb{I}.$$

Using the d -coefficients, we can also represent the pushforward $\vec{V} = F_* V$ of V in data space from (3.7) :

$$(3.16) \quad \vec{V} = \sum_{i,k,l=0}^{\infty} \sum_{j=1}^J b_{ij} F_k \langle \phi_l, B_{ij} \triangleright \phi_k \rangle_{L^2(\nu)} \varphi_l = \sum_{i,k,l=0}^{\infty} \sum_{j=1}^J d_{ijkl} b_{ij} F_k \varphi_l,$$

where the summation over i, j, k, l converges in $L^2(\nu_*; \mathbb{R}^d)$ norm.

Equation (3.16) establishes the pushforward of the vector field as an infinite sum of various coefficients. Among these, the d -coefficients depend only on the geometry of \mathcal{M} , the F -coefficients depend only on the embedding, and the b -coefficients are the only coefficients that depend on the vector field V . We next describe how the various infinite sums determining these coefficients may be truncated to finite sums. This is essential for our eventual data-driven realization of (3.16).

3.4. The truncated problem and vector field projection. We fix a positive integer $L \in \mathbb{N}$ which we call the *spectral resolution parameter*. It determines the hypothesis space $\mathbb{H}_L \subset \mathbb{H}$ of dimension $D_L = LJ$ which we shall use for the approximation of the vector field V :

$$\mathbb{H}_L = \text{span}\{B_{ij} : 0 \leq i \leq L-1, 1 \leq j \leq J\}.$$

Let similarly $\ell^2(\mathbb{I}_L) = \text{span}\{\varepsilon_{ij} : 0 \leq i \leq L, 1 \leq j \leq J\} \subset \ell^2(\mathbb{I})$, $\dim \ell^2(\mathbb{I}_L) = D_L$, and define $\pi_L: \ell^2 \rightarrow \ell^2$ as the orthogonal projection with $\text{ran } \pi_L = \ell^2(\mathbb{I}_L)$.

Associated with π_L is a projected analysis operator $T_L := \pi_L T$, whose adjoint $T_L^* = T^* \pi_L$ also maps into \mathbb{H}_L . Moreover, since π_L is an orthogonal projection, we have $\pi_L^+ = \pi_L$ and the pseudoinverse of T_L is given by $T_L^+ = \pi_L T^+$. Defining the projected Gram operator

$$G_L := \pi_L G \pi_L = T_L T_L^*,$$

we have $G_L^+ = \pi_L G^+ \pi_L$. Moreover, G_L is represented by a $D_L \times D_L$ symmetric matrix \mathbf{G}_L with elements

$$G_{mn} = \langle \varepsilon_{i_m j_m}, G \varepsilon_{i_n j_n} \rangle_{\ell^2(\mathbb{I})} = \langle B_{i_m, j_m}, B_{i_n, j_n} \rangle_{\mathbb{H}} = d_{i_m j_m j_n i_n},$$

whose pseudoinverse $\mathbf{G}_L^+ = [G_{mn}^+]$ gives the matrix elements $\langle \varepsilon_{i_m, j_m}, G_L^+ \varepsilon_{i_n, j_n} \rangle_{\ell^2(\mathbb{I})} = G_{mn}^+$ of G_L^+ .

Let $\mathcal{D}_L = \pi_L \mathcal{D}$. We define a spectrally truncated approximation $V^{(L)} \in C(\mathcal{M}, T\mathcal{M})$ of the dynamical vector field V as

$$(3.17) \quad V^{(L)} := T^* b^{(L)}, \quad b^{(L)} = G_L^+ \mathcal{D}_L \mathcal{F} V,$$

where $b^{(L)} \in \ell^2(\mathbb{I}_L)$ is a minimum-norm solution of the regression problem (cf. (3.9))

$$(3.18) \quad G_L b^{(L)} = \mathcal{D}_L \mathcal{F} V.$$

Equivalently, $V^{(L)}$ is the best-fit approximation of V among vector fields in the hypothesis space \mathbb{H}_L , in the sense of minimizing the square error functional $\mathcal{E}_L: \mathbb{H}_L \rightarrow \mathbb{R}$, $\mathcal{E}_L(W) = \|V - W\|_{\mathbb{H}_L}^2$.

As $L \rightarrow \infty$, π_L converges strongly to the identity on $\ell^2(\mathbb{I})$, and thus T_L and G_L^+ converge to T and G , respectively. We deduce that $b^{(L)}$ converges to a (minimum-norm) solution b of (3.9) and $V^{(L)}$ converges to $V = T^* b$ in the Hodge norm,

$$(3.19) \quad \lim_{L \rightarrow \infty} \|V^{(L)} - V\|_{\mathbb{H}} = 0.$$

Next, we can write down an expansion for $V^{(L)}$ in the frame,

$$V^{(L)} = \sum_{i=0}^{L-1} \sum_{j=1}^J b_{ij}^{(L)} B_{ij}, \quad b_{ij}^{(L)} = \langle \varepsilon_{ij}, b^{(L)} \rangle_{\ell^2(\mathbb{I})}.$$

This expansion leads to an approximation $\vec{V}^{(L)} = F_* V^{(L)}$ of the pushforward vector field \vec{V} , given by (cf. (3.16))

$$(3.20) \quad \vec{V}^{(L)} = \sum_{i=0}^{L-1} \sum_{j=1}^J \sum_{k,l=0}^{\infty} d_{ijkl} b_{ij}^{(L)} F_k \varphi_l.$$

Note that (3.20) gives $V^{(L)}$ as a finite linear combination of infinite-rank (unbounded) operators B_{ij} . Similarly, the regression problem yielding the coefficients $b^{(L)}$ involves an infinite-rank (bounded) operator, \mathcal{F} , in the right-hand side. To arrive at an approximation that is more amenable to numerical implementation, we make a number of modifications to (3.20). In what follows, $H_M = \text{span}\{\phi_0, \dots, \phi_{M-1}\}$ will be the M -dimensional subspace of $L^2(\nu)$ spanned by the leading M Laplace–Beltrami eigenfunctions, and $\Pi_M: L^2(\nu) \rightarrow L^2(\nu)$ the orthogonal projection mapping into this subspace. Given $M_1, M_2 \in \mathbb{N}$, we let $\ell^2(\hat{\mathbb{I}}_{M_1, M_2})$ be the $(M_1 M_2)$ -dimensional subspace of $\ell^2(\hat{\mathbb{I}})$ corresponding to the index set $\hat{\mathbb{I}}_{M_1, M_2} = \{(i, j) \in \mathbb{N}_0 \times \mathbb{N}_0 : i < M_1, j < M_2\}$, and define $\hat{\pi}_{M_1, M_2}: \ell^2(\hat{\mathbb{I}}) \rightarrow \ell^2(\hat{\mathbb{I}})$ as the orthogonal projection mapping onto this subspace.

- We approximate the d -coefficients by truncating the summation in (3.14) to $L_D \in \mathbb{N}$ terms,

$$d_{ijkl}^{(L_D)} := \sum_{p=0}^{L_D-1} c_{ilp} g_{pjk}.$$

This truncation induces densely-defined operators $B_{ij}^{(L_D)}$ on $L^2(\nu)$ defined via the matrix representation

$$\langle \phi_l, B_{ij}^{(L_D)} \phi_k \rangle_{L^2(\nu)} = d_{ijkl}^{(L_D)}.$$

We view $B_{ij}^{(L_D)}$ as approximations of the vector fields B_{ij} .

- Using the $B_{ij}^{(L_D)}$, we introduce $G^{(L, L_D)}: \ell^2(\mathbb{I}) \rightarrow \ell^2(\mathbb{I})$ as an approximation of the Gram operator G_L with matrix elements (cf. (3.15))

$$G_{ijkl}^{(L, L_D)} := \langle \varepsilon_{ij}, G^{(L, L_D)} \varepsilon_{kl} \rangle_{\ell^2(\mathbb{I})} = d_{ijlk}^{(L_D)}.$$

We also define $\mathcal{D}^{(L, L_D)}: \text{dom}(\mathcal{D}^{(L, L_D)}) \rightarrow \ell^2(\mathbb{I})$ as an approximation of \mathcal{D}_L on a dense domain $\text{dom}(\mathcal{D}^{(L, L_D)}) \subset \ell^2(\hat{I})$, where

$$\langle \varepsilon_{ij}, \mathcal{D}^{(L, L_D)} \hat{\varepsilon}_{kl} \rangle_{\ell^2(\mathbb{I})} = d_{ijkl}^{(L_D)}.$$

- Let $G^{(L, L_D)} u_j = \gamma_j u_j$ be an eigendecomposition of $G^{(L, L_D)}$ where the eigenvalues γ_j are non-negative and the eigenvectors u_j are orthonormal in $\ell^2(\mathbb{I})$. For $\eta > 0$, we approximate $G^{(L, L_D)}$ by the spectrally truncated Gram operator $G^{(L, \eta, L_D)}: \ell^2(\mathbb{I}) \rightarrow \ell^2(\mathbb{I})$, where

$$(3.21) \quad G^{(L, \eta, L_D)} = \sum_{j: \gamma_j > \eta} \gamma_j u_j \langle u_j, \cdot \rangle_{\ell^2(\mathbb{I})}.$$

Note that the pseudoinverse of this operator is given by

$$G^{(L, \eta, L_D), +} = \sum_{j: \gamma_j > \eta} \gamma_j^{-1} u_j \langle u_j, \cdot \rangle_{\ell^2(\mathbb{I})},$$

and has norm at most $1/\eta$ by construction.

- With $L_1, L_2 \in \mathbb{N}$, we replace the regression problem for $b^{(L)}$ in (3.18) by

$$(3.22) \quad G^{(L, \eta, L_D)} b^{(L, L_1, L_2, \eta, L_D)} = \mathcal{D}^{(L, L_D)} \mathcal{F}^{(L_1, L_2)} \vec{V},$$

where $\mathcal{F}^{(L_1, L_2)} = \hat{\pi}_{L_1, L_2} \mathcal{F}$. We use the solution

$$(3.23) \quad b^{(L, L_1, L_2, \eta, L_D)} = G^{(L, \eta, L_D), +} \mathcal{D}^{(L, L_D)} \mathcal{F}^{(L_1, L_2)} \vec{V} \in \ell^2(\mathbb{I}_L)$$

of this problem as an approximation of $b^{(L)}$ from (3.17).

- We approximate $B_{ij}^{(L_D)}$ by finite-rank operators $B_{ij}^{(L_1, L_2, L_D)}: L^2(\nu) \rightarrow L^2(\nu)$ where

$$B_{ij}^{(L_1, L_2, L_D)} = \Pi_{L_2} B_{ij}^{(L_D)} \Pi_{L_1}.$$

Letting $\ell = (L, L_1, L_2, \eta, L_D)$ collectively represent the truncation parameters introduced in this subsection, we arrive at the following modified version of (3.20),

$$(3.24) \quad \vec{V}^{(\ell)} := \sum_{i=0}^{L-1} \sum_{j=1}^J \sum_{k=0}^{L_1-1} \sum_{l=0}^{L_2-1} d_{ijkl}^{(L_D)} b_{ij}^{(\ell)} F_k \varphi_l,$$

where $b_{ij}^{(\ell)} = \langle \varepsilon_{ij}, b^{(\ell)} \rangle_{\ell^2(\mathbb{I})}$.

The vector field $\vec{V}^{(\ell)}$ in (3.24) can be employed as a practical approximation to \vec{V} assuming that (finitely many) coefficients $d_{ijkl}^{(L_D)}$, $b_{ij}^{(L, L_D)}$, F_k and basis functions φ_l are known. In section 4 we will relax that assumption and take up the problem of data-driven approximation of these objects using kernel methods.

Remark 3.4. As we will discuss in subsection 5.1, the spectral truncation of the Gram operators by η from below is introduced in order to ensure convergence of the pseudoinverse of $G^{(L, \eta, L_D)}$ as $L_D \rightarrow \infty$. In particular, the approximation of G_L by $G^{(L, L_D)}$ is not an approximation by projection (in contrast to the approximation of G by G_L), and we cannot deduce convergence of $G^{(L, L_D), +}$ as $L_D \rightarrow \infty$.

4. Data-driven formulation. The data-driven formulation of our scheme is built on kernel-based approximations of the Laplace–Beltrami eigenpairs, (λ_j, ϕ_j) , as well as the corresponding extensions φ_j of the eigenfunctions on data space. With such approximations, the data-driven scheme is essentially structurally identical to the approach presented in subsection 3.4 leading to the projected vector field $\vec{V}^{(\ell)}$ in (3.24).

We begin in subsection 4.1 by stating abstract conditions on kernel methods that are sufficient to guarantee asymptotic consistency of our scheme in the limit of large data. A concrete example satisfying these conditions is the diffusion maps algorithm [16], which is used in the numerical experiments of section 6 and summarized in Appendix A.1. In subsection 4.2, we present the associated SEC-based scheme for learning vector fields.

In what follows, $L^2(\mu_N)$ will be the N -dimensional L^2 space on \mathcal{M} associated with the sampling measure μ_N . Note that elements of $L^2(\mu_N)$ are equivalence classes of functions f with values $f(x_n)$ known at the sampling points $x_n \in \mathcal{M}$. As such, every such element can be represented by a column vector $\mathbf{f} = (f(x_1), \dots, f(x_N))^T \in \mathbb{R}^N$ and every linear operator $A: L^2(\mu_N) \rightarrow L^2(\mu_N)$ is representable by an $N \times N$ matrix \mathbf{A} such that $\mathbf{A}\mathbf{f}$ is the column vector representation of Af .

4.1. Kernel-based approximation of the Laplacian. The approximations of the Laplacian used in this paper are based on normalized kernels $\rho_{\epsilon, N}: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_{>0}$, parameterized by a bandwidth (lengthscale) parameter $\epsilon > 0$. The kernels are C^∞ smooth, and are constructed using the training dataset $\{y_n\}_{n=0}^{N-1} \subset \mathbb{R}^d$ so that they are Markovian with respect to the sampling measure μ_N ; that is,

$$(4.1) \quad \int_{\mathcal{M}} \rho_{\epsilon, N}(y, F(x)) d\mu_N(x) = 1, \quad \forall y \in \mathbb{R}^d.$$

Every such kernel pulls back to a C^∞ kernel $p_{\epsilon, N}: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}_{>0}$ on the manifold,

$p_{\epsilon,N}(x, x') = \rho_{\epsilon,N}(F(x), F(x'))$, with an associated integral operator $P_{\epsilon,N}: L^2(\mu_N) \rightarrow L^2(\mu_N)$,

$$P_{\epsilon,N}f = \int_{\mathcal{M}} p_{\epsilon,N}(\cdot, x)f(x) d\mu_N(x) \equiv \frac{1}{N} \sum_{n=1}^N \rho_{\epsilon,N}(F(\cdot), x_n)f(x_n).$$

Defining

$$(4.2) \quad \Delta_{\epsilon,N} := -\frac{\log P_{\epsilon,N}}{c\epsilon^2}$$

for an ϵ -independent constant c that depends on the parametric family of the kernel $\rho_{\epsilon,N}$ (but not on ϵ or N), our data-driven approach employs kernels such that $\Delta_{\epsilon,N}$ is an approximation of the Laplacian Δ and $P_{\epsilon,N}$ is an approximation of the heat operator $e^{-c\epsilon^2\Delta}$. In addition, $\rho_{\epsilon,N}$ is constructed such that $P_{\epsilon,N}$ is ergodic (i.e., its eigenvalue equal to 1 is simple) and reversible (i.e., it satisfies detailed balance),

$$(4.3) \quad w_{\epsilon,N}(x)p_{\epsilon,N}(x, x') = w_{\epsilon,N}(x')p_{\epsilon,N}(x', x), \quad \mu_N\text{-a.e.}, x, x' \in \mathcal{M}.$$

Recall the sampling density σ introduced in [Assumption 3](#). In (4.3), $w_{\epsilon,N} \in L^1(\mu_N)$ is a vector with strictly positive elements that provides an estimate of the reciprocal $\frac{1}{\sigma}$, i.e., the density of the volume measure relative to the sampling measure of the training data. To build this estimate, we normalize $w_{\epsilon,N}$ such that $\int_{\mathcal{M}} w_{\epsilon,N} d\mu_N$ is an approximation of the volume $\nu(\mathcal{M})$ of the manifold (obtained using heat kernel asymptotics; see [Appendix A.1](#)). With this normalization, empirical integrals $\int_{\mathcal{M}} f d\nu_{\epsilon,N}$ of continuous functions $f \in C(\mathcal{M})$ with respect to the weighted measure

$$(4.4) \quad \nu_{\epsilon,N} := \sum_{n=1}^N w_{\epsilon,N}(x_n)\delta_{x_n},$$

approximate integrals $\int_{\mathcal{M}} f d\nu$ with respect to the volume measure ν of the manifold. Reversibility also implies that $P_{\epsilon,N}$ is diagonalizable and has real eigenvalues. When convenient, we will use the symbol α to collectively denote data-driven approximation parameters; e.g., $\alpha = (\epsilon, N)$ in the present context.

Next, let us consider an eigendecomposition

$$(4.5) \quad P_{\alpha}\phi_j^{(\alpha)} = \Lambda_j^{(\alpha)}\phi_j^{(\alpha)},$$

where the eigenvalues $\Lambda_j^{(\alpha)} \in \mathbb{R}$ are bounded in modulus by 1 since P_{α} is a Markov operator and the eigenvectors $\phi_j^{(\alpha)}$ are normalized to unit norm on $L^2(\nu_{\alpha})$. Moreover, w_{α} is an eigenvector of P_{α}^* at eigenvalue 1, $P_{\alpha}^*w_{\alpha} = w_{\alpha}$. Observe that (4.5) is equivalent to the $N \times N$ matrix eigenvalue problem

$$\mathbf{P}_{\alpha}\phi_j^{(\alpha)} = \Lambda_j^{(\alpha)}\phi_j^{(\alpha)},$$

where $\mathbf{P}_{\alpha} = [P_{ij}^{(\alpha)}]_{i,j=1}^N$ is a stochastic matrix with entries $P_{ij}^{(\alpha)} = p_{\alpha}(x_i, x_j)/N$ and the eigenvectors $\phi_j^{(\alpha)} \in \mathbb{R}^N$ are normalized such that $\phi_j^{(\alpha),\top} \mathbf{W}_{\alpha}\phi_j^{(\alpha)} = 1$ for the diagonal matrix

$W_\alpha = \text{diag}(w_\alpha(x_1), \dots, w_\alpha(x_N))$. Note also that every eigenvector $\phi_j^{(\alpha)}$ with nonzero corresponding eigenvalue $\Lambda_j^{(\alpha)}$ has a smooth extension $\varphi_j^{(\alpha)} \in C^\infty(\mathbb{R}^d)$ on data space,

$$(4.6) \quad \varphi_j^{(\alpha)}(y) = \frac{1}{\Lambda_j^{(\alpha)}} \int_{\mathcal{M}} \rho_\alpha(y, x) \phi_j(x) d\mu_N(x),$$

satisfying $\varphi_j^{(\alpha)}(y_n) = \phi_j^{(\alpha)}(x_n)$ for every $n \in \{1, \dots, N\}$.

The operator Δ_α has an eigendecomposition

$$\Delta_\alpha \phi_j^{(\alpha)} = \lambda_j^{(\alpha)} \phi_j^{(\alpha)}$$

with eigenvalues $\lambda_j^{(\alpha)} = -\frac{1}{c\epsilon^2} \log \Lambda_j^{(\alpha)}$. For simplicity of exposition, we will assume that all eigenvalues $\Lambda_j^{(\alpha)}$ are nonnegative, and can thus be ordered as $1 = \Lambda_0^{(\alpha)} \geq \Lambda_1^{(\alpha)} \geq \dots \geq \Lambda_{N-1}^{(\alpha)} \geq 0$ by Markovianity of P_α .

As we will see in [subsection 4.2](#) below, the following assumption encapsulates sufficient conditions for convergence of our data-driven scheme.

Assumption 4. *The operators Δ_α converge spectrally to the Laplacian Δ as $\alpha = (\epsilon, N)$ tends to $(0^+, \infty)$ along a sequence $\alpha_N = (\epsilon_N, N)$, in the following sense:*

- (i) *For every $j \in \mathbb{N}$, the eigenvalues $\lambda_j^{(\alpha_N)}$ converge to the Laplace–Beltrami eigenvalue λ_j , including multiplicities.*
- (ii) *For every Laplace–Beltrami eigenfunction $\phi_j \in C^\infty(\mathcal{M})$ corresponding to λ_j there exists a sequence of eigenvectors $\phi_j^{(\alpha_N)} \in L^2(\mu_N)$ of Δ_α corresponding to $\lambda_j^{(\alpha)}$ whose representatives $\varphi_j^{(\alpha_N)}$ converge uniformly to ϕ_j on \mathcal{M} :*

$$\lim_{N \rightarrow \infty} \left\| \varphi_j^{(\alpha_N)} \circ F - \phi_j \right\|_{C(\mathcal{M})} = 0.$$

- (iii) *The convergence of the eigenvalues in (i) takes place at a sufficiently rapid rate such that*

$$\lim_{N \rightarrow \infty} \epsilon_N^{\dim(\mathcal{M})} \sum_{j=0}^{N-1} \left(e^{-\epsilon_N^2 \lambda_j^{(\alpha_N)}} - e^{-\epsilon_N^2 \lambda_j} \right) = 0.$$

A number of techniques from the field of manifold learning satisfy [Assumption 4](#); e.g., [\[73, 71, 76, 13, 63\]](#). Convergence results from these methods are summarized in [Table 1](#). As mentioned earlier, the numerical examples presented in this paper utilize the diffusion maps algorithm [\[16\]](#), which is a popular kernel method for approximation of Laplacians and weighted Laplacians under potentially nonuniform sampling distributions μ with respect to the volume measure ν of the manifold. A summary of diffusion maps as used in this paper is included in [Appendix A.1](#). The paper [\[76\]](#) proves spectral convergence of this techniques for data sampled on flat tori. This analysis thus covers the numerical examples in [subsection 6.1](#), which take place on circles. In [subsection 6.2](#), we consider examples on tori with curvature, which lie outside the theoretical scope of [\[76\]](#). Nevertheless, there is strong numerical evidence that diffusion maps provides approximations of the Laplacian of sufficient spectral accuracy even in those cases.

Table 1

Representative data-driven approximation techniques with spectral convergence for the Laplacian. *Assumption 4* addresses spectral convergence as the data parameters $\alpha := (\epsilon, N)$ approach $(0^+, \infty)$. Convergence results have been established for various approaches to this limit, and under various assumptions on the manifold \mathcal{M} . The table lists a few such results in chronological order of the corresponding reference in the literature. All paths of approach of α to $(0^+, \infty)$ are along a sequence $\alpha_N = (\epsilon_N, N)$. The variable m represents the dimension of the manifold \mathcal{M} , p_m is a constant determined by m , and k is the maximum multiplicity among a fixed number of top eigenvalues.

Reference	Assumptions on the manifold	Rate law ϵ_N
[73]	Closed	$O(\log(N)^{p_m} N^{1/2m})$
[71]	Closed	$\epsilon_N^{1/2} + \frac{\ln N - \ln \epsilon_N + 1}{\epsilon_N^{k+3} \sqrt{N}} = o(1)$
[76]	Torus	N/A
[13]	Closed	$O\left((\ln N/N)^{1/O(m)}\right)$
[63]	Compact with boundary	$O\left((\ln N/N)^{1/O(m)}\right)$

Remark 4.1. In supervised and unsupervised learning applications utilizing the $\phi_j^{(\alpha)}$ as basis vectors (e.g., feature extraction [16] and kernel regression [1]) the normalization of the weight vector w_α can oftentimes be chosen arbitrarily without affecting the results. On the other hand, the SEC approach presented in this paper uses the product structure of the $\phi_j^{(\alpha)}$ in conjunction with the eigenvalues of Δ_α to implement the carré du champ identity of the Laplacian (1.4) (through data-driven analogs of the c - and g -coefficients; see subsection 4.2 below). Consistency of this implementation requires that the $\phi_j^{(\alpha)}$ are normalized consistently with the volume form of the manifold. Geometrically, this stems from the fact that the volume $\nu(\mathcal{M})$ is related to the trace of the heat operator (i.e., the sum of its eigenvalues). *Assumption 4(iii)* is essentially a trace condition on P_α that ensures that $\nu(\mathcal{M})$ can be consistently approximated using the trace of the matrices \mathbf{P}_α ; see Appendix A.1.

4.2. Data-driven SEC framework. Let $\ell = (L, L_1, L_2, \eta, L_D)$ be the vector of discretization parameters from subsection 3.4 and set $M = \max\{L, L_1, L_2, L_D\}$. By Assumption 4, for a sufficiently large number of samples N , the leading M eigenvectors $\phi_0^{(\alpha)}, \dots, \phi_{M-1}^{(\alpha)}$ are linearly independent and form a basis of an M -dimensional subspace $H_{M,N} \subseteq L^2(\nu_\alpha)$ for the empirical volume measure ν_α from (4.4). By convention, we always choose the basis vectors so that they are orthonormal in this space, $\langle \phi_i^{(\alpha)}, \phi_j^{(\alpha)} \rangle_{L^2(\nu_\alpha)} = \delta_{ij}$. We also define $\Pi_{M,N}: L^2(\nu_\alpha) \rightarrow L^2(\nu_\alpha)$ as the orthogonal projection with $\text{ran } \Pi_{M,N} = H_{M,N}$. Similarly to the expansion of the embedding $F: \mathcal{M} \rightarrow \mathbb{R}^d$ in the Laplace-Beltrami eigenfunction basis, we can write down an expansion $F^{(\alpha)} = \sum_{k=0}^{N-1} F_k^{(\alpha)} \phi_k^{(\alpha)} \in L^2(\nu_\alpha)$ with coefficients $F_k^{(\alpha)} \in \mathbb{R}^d$ given by componentwise discrete integrals

$$F_k^{(\alpha)} = \int_{\mathcal{M}} \phi_k^{(\alpha)} F d\nu_\alpha = \sum_{n=1}^N \phi_k^{(\alpha)}(x_n) y_n w_\alpha(x_n).$$

This expansion recovers F on the training dataset (i.e., μ_N -a.e.), $F^{(\alpha)}(x_n) = F(x_n)$ for all $n \in \{1, \dots, N\}$.

The data-driven implementation of our framework is based on the following approximations of the c -, g -, and d -coefficients from [section 3](#):

$$(4.7) \quad \begin{aligned} c_{ijp}^{(\alpha)} &:= \left\langle \phi_p^{(\alpha)}, \phi_i^{(\alpha)} \phi_j^{(\alpha)} \right\rangle_{L^2(\nu_\alpha)}, \quad g_{pjk}^{(\alpha)} := \frac{1}{2} \left(\lambda_j^{(\alpha)} + \lambda_k^{(\alpha)} - \lambda_p^{(\alpha)} \right) c_{jkp}^{(\alpha)}, \\ d_{ijkl}^{(L_D, \alpha)} &:= \sum_{p=0}^{L_D-1} c_{ilp}^{(\alpha)} g_{pjk}^{(\alpha)}. \end{aligned}$$

With these approximations, we define operators $G^{(L_D, \alpha)}: \ell^2(\mathbb{I}_N) \rightarrow \ell^2(\mathbb{I}_N)$ and $\mathcal{D}^{(L_D, \alpha)}: \ell^2(\hat{\mathbb{I}}_{N, N}) \rightarrow \ell^2(\mathbb{I}_{N, N})$ of the $G^{(L_D)}$ and $\mathcal{D}^{(L_D)}$ operators from [subsection 3.4](#) to have matrix elements

$$(4.8) \quad \langle \varepsilon_{ij}, G^{(L_D, \alpha)} \varepsilon_{kl} \rangle_{\ell^2(\mathbb{I})} = d_{ijlk}^{(L_D, \alpha)}, \quad \langle \varepsilon_{ij}, \mathcal{D}^{(L_D, \alpha)} \hat{\varepsilon}_{kl} \rangle_{\ell^2(\mathbb{I})} = d_{ijkl}^{(L_D, \alpha)}.$$

For $L \in \mathbb{N}$ and $\eta > 0$, we also define projected Gram operators $G^{(L, L_D, \alpha)} = \pi_L G^{(L_D, \alpha)} \pi_L$ and their spectrally truncated approximations $G^{(L, \eta, L_D, \alpha)}$ on $\ell^2(\mathbb{I}_N)$ analogously to $G^{(L, L_D)}$ and $G^{(L, \eta, L_D)}$ from [subsection 3.4](#), respectively.

Next, as a data-driven approximation of $\mathcal{F}: L^2(\nu_*; \mathbb{R}^d) \rightarrow \ell^2(\hat{\mathbb{I}})$, we introduce the operator $\mathcal{F}^{(\alpha)}: L^2(\nu_{\alpha*}; \mathbb{R}^d) \rightarrow \ell^2(\hat{\mathbb{I}}_{N, N})$, where

$$\mathcal{F}^{(\alpha)} \vec{W} = \sum_{(k, l) \in \hat{\mathbb{I}}_{N, N}} \langle (F_k^{(\alpha)} \cdot \vec{W}) \circ F, \phi_l^{(\alpha)} \rangle_{L^2(\nu_\alpha)} \hat{\varepsilon}_{kl}.$$

Note that the action of this map on the pushforward vector field can be evaluated using the vector field samples (“arrows”) $\vec{v}_n \in \mathbb{R}^d$ in our training dataset, viz.

$$(4.9) \quad \mathcal{F}^{(\alpha)} \vec{V} = \sum_{(k, l) \in \hat{\mathbb{I}}_{N, N}} \sum_{n=1}^N F_k^{(\alpha)} \cdot \vec{v}_n \phi_l^{(\alpha)}(x_n) w_\alpha(x_n) \hat{\varepsilon}_{kl}.$$

Combining [\(4.8\)](#) and [\(4.9\)](#), and using projections $\pi_L: \ell^2(\mathbb{I}_N) \rightarrow \ell^2(\mathbb{I}_N)$ and $\hat{\pi}_{L_1, L_2}: \ell^2(\hat{\mathbb{I}}_{N, N}) \rightarrow \ell^2(\hat{\mathbb{I}}_{N, N})$ and spectral resolution parameters $\ell = (L, L_1, L_2, \eta, L_D)$ as in [subsection 3.4](#), we arrive at the regression problem

$$(4.10) \quad G^{(L, \eta, L_D, \alpha)} b^{(\ell, \alpha)} = \mathcal{D}^{(L, L_D, \alpha)} \mathcal{F}^{(L_1, L_2, \alpha)} \vec{V},$$

where $\mathcal{F}^{(L_1, L_2, \alpha)} = \hat{\pi}_{L_1, L_2} \mathcal{F}^{(\alpha)}$. This regression problem is a data-driven analog of [\(3.22\)](#). The solution of [\(4.10\)](#),

$$(4.11) \quad b^{(\ell, \alpha)} = G^{(L, L_D, \eta, \alpha), +} \mathcal{D}^{(L, L_D, \alpha)} \mathcal{F}^{(L_1, L_2, \alpha)} \vec{V},$$

is our data-driven approximation of $b^{(\ell)}$ from [\(3.23\)](#) with coefficients $b_{ij}^{(\ell, \alpha)} = \langle \varepsilon_{ij}, b^{(\ell, \alpha)} \rangle_{\ell^2(\mathbb{I})}$.

Finally, defining $B_{ij}^{(L_D, \alpha)}: L^2(\nu_\alpha) \rightarrow L^2(\nu_\alpha)$ to have matrix elements

$$\langle \phi_l^{(\alpha)}, B_{ij}^{(L_D, \alpha)} \phi_k \rangle_{L^2(\nu_\alpha)} = d_{ijkl}^{(L_D, \alpha)},$$

and setting $B_{ij}^{(L_1, L_2, L_D, \alpha)} = \Pi_{L_1, N} B_{ij}^{(L_D, \alpha)} \Pi_{L_2, N}$, leads to our data-driven approximation $\vec{V}^{(\ell, \alpha)} \in C^\infty(\mathbb{R}^d; \mathbb{R}^d)$ of $\vec{V}^{(\ell)}$ from (3.24),

$$(4.12) \quad \vec{V}^{(\ell, \alpha)} = \sum_{i=0}^{L-1} \sum_{j=1}^J \sum_{k=0}^{L_1-1} \sum_{l=0}^{L_2-1} d_{ijkl}^{(L_D, \alpha)} b_{ij}^{(\ell, \alpha)} F_k^{(\alpha)} \varphi_l^{(\alpha)}.$$

Each of the terms in (4.12) is data-driven and computable. The schematic in Figure 1 includes information about the time complexity of training (i.e., computation of $d_{ijkl}^{(L_D, \alpha)}$, $b_{ij}^{(\ell, \alpha)}$, and $F_k^{(\alpha)}$) and prediction (i.e., evaluation of $\vec{V}^{(\ell, \alpha)}(y)$) with our scheme. In the next section, we study the convergence of the reconstructed vector field (4.12).

5. Convergence analysis. We study the asymptotic consistency of the data-driven, SEC-based approximation of the dynamical vector field (subsection 5.1) and the accuracy of the associated dynamical orbits (subsection 5.2).

5.1. Consistency of vector field approximation. Our main approximation result establishes convergence of the data-driven vector field approximation scheme from subsection 4.2 to the true vector field \vec{V} in L^2 sense in an asymptotic limit of large data and infinite spectral resolution.

Theorem 5.1. *Let $\vec{V}^{(\ell, \alpha)}$ be the data-driven vector field (4.12), where $\alpha = (\epsilon, N)$, N is the training dataset size, ϵ is the kernel bandwidth parameter and $\ell = (L, L_1, L_2, \eta, L_D)$ is the collection of SEC spectral resolution parameters. Let Assumptions 1–4 hold, where the large-data limit in Assumption 4 is taken along a sequence $\alpha_N = (\epsilon_N, N)$ tending to $(0^+, \infty)$. Then,*

$$\lim_{\ell \rightarrow \infty} \lim_{N \rightarrow \infty} \left\| \vec{V}^{(\ell, \alpha_N)} - \vec{V} \right\|_{L^2(\nu_*, \mathbb{R}^d)} = 0,$$

where $\lim_{\ell \rightarrow \infty}$ denotes the iterated limit $\lim_{L \rightarrow \infty} \lim_{L_1 \rightarrow \infty} \lim_{L_2 \rightarrow \infty} \lim_{\eta \rightarrow 0^+} \lim_{L_D \rightarrow \infty}$.

Proof of Theorem 5.1. The convergence claim in Theorem 5.1 involves two types of limits: (i) the large-data limit $\alpha_N \rightarrow (\infty, 0^+)$; and (ii) infinite spectral resolution, $\ell \rightarrow \infty$. We study these limits separately in the ensuing subsections.

5.1.1. Large-data limit. We begin by showing that the weighted empirical measures ν_α from (4.4) consistently approximate the Riemannian volume measure ν .

Lemma 5.2. *Under Assumptions 3 and 4 and for the normalization of the weights w_α in (A.6) and (A.7), the measures ν_{α_N} converge weakly to ν ,*

$$\lim_{N \rightarrow \infty} \int_{\mathcal{M}} f d\nu_{\alpha_N} = \int_{\mathcal{M}} f d\nu, \quad \forall f \in C(\mathcal{M}).$$

Proof. By construction, the operator P_α is self-adjoint on $L^2(\nu_\alpha)$ and admits the decomposition $P_\alpha = \sum_{j=0}^{N-1} \Lambda_j^{(\alpha)} \phi_j^{(\alpha)} \langle \phi_j^{(\alpha)}, \cdot \rangle_{L^2(\nu_\alpha)}$. This implies that the kernel p_α admits the Mercer-type expansion $p_\alpha(x_m, x_n) = \sum_{j=0}^{N-1} \Lambda_j^{(\alpha)} \phi_j^{(\alpha)}(x_m) \phi_j^{(\alpha)}(x_n) w_\alpha(x_n)$, giving

$$\text{tr}_{L^2(\nu_\alpha)} P_\alpha = \sum_{j=0}^{N-1} \langle \phi_j^{(\alpha)}, P_\alpha \phi_j^{(\alpha)} \rangle_{L^2(\nu_\alpha)} = \int_{\mathcal{M}} p_\alpha(x, x) d\mu_N(x) = \text{tr } P_\alpha.$$

By [Assumption 4](#)(i, iii), we have $\lim_{N \rightarrow \infty} \epsilon_N^{\dim(\mathcal{M})} \left(\text{tr}_{L^2(\nu_{\alpha_N})} P_{\alpha_N} - \text{tr}_{L^2(\nu)} e^{-\epsilon_N^2 \Delta} \right) = 0$, where $e^{-\epsilon_N^2 \Delta}$ is the time- ϵ_N^2 heat operator on $L^2(\nu)$. Meanwhile, by the Minakshisundaram–Pleijel heat-trace formula (see [Appendix A.1](#)) we have $\lim_{N \rightarrow \infty} (4\pi\epsilon_N^2)^{\dim(\mathcal{M})/2} \text{tr}_{L^2(\nu)} e^{-\epsilon_N^2 \Delta} = \nu(\mathcal{M})$. As a result, under the normalization from [\(A.6\)](#) and [\(A.7\)](#), the weights w_α satisfy

$$(5.1) \quad \lim_{N \rightarrow \infty} \int_{\mathcal{M}} w_{\alpha_N} d\mu_N = \nu(\mathcal{M}).$$

Now, by [Assumption 3](#) and [Assumption 4](#)(ii) (and noting [\(A.4\)](#) and [\(A.4\)](#)), we have

$$\lim_{N \rightarrow \infty} \int_{\mathcal{M}} f d\nu_\alpha = C \int_{\mathcal{M}} f \frac{1}{\sigma} d\mu = C \int_{\mathcal{M}} f d\nu, \quad \forall f \in C(\mathcal{M}),$$

where C is a normalization constant. By [\(5.1\)](#), we have $C = 1$ and the claim of the lemma follows. ■

[Lemma 5.2](#), in conjunction with the uniform convergence of the eigenvectors in [Assumption 4](#)(ii), implies that the data-driven c -coefficients from [subsection 4.2](#) converge. Indeed, for every $i, j, p \in \mathbb{N}$ and sufficiently large N we have

$$\begin{aligned} |c_{ijp}^{(\alpha)} - c_{ijp}| &= |\langle \phi_p^{(\alpha)}, \phi_i^{(\alpha)} \phi_j^{(\alpha)} \rangle_{L^2(\nu_\alpha)} - \langle \phi_p, \phi_i \phi_j \rangle_{L^2(\nu)}| \\ &= |\nu_\alpha(\phi_p^{(\alpha)} \phi_i^{(\alpha)} \phi_j^{(\alpha)}) - \nu(\phi_p \phi_i \phi_j)| \\ &\leq |\nu_\alpha(\phi_p^{(\alpha)} \phi_i^{(\alpha)} \phi_j^{(\alpha)}) - \nu(\phi_p \phi_i \phi_j)| + |(\nu_\alpha - \nu)(\phi_p \phi_i \phi_j)|, \end{aligned}$$

and both terms in the last line converge to 0 on the sequence α_N as $N \rightarrow \infty$. Using [Assumption 4](#)(i), we can therefore deduce convergence of the g - and d - coefficients (the latter, at fixed L_D),

$$(5.2) \quad \lim_{N \rightarrow \infty} g_{pjk}^{(\alpha_N)} = g_{pjk}, \quad \lim_{\alpha_N \rightarrow \infty} d_{ijkl}^{(L_D, \alpha)} = d_{ijkl}^{(L_D)}, \quad \forall i, j, k, l, p \in \mathbb{N}.$$

[Lemma 5.2](#) and continuity of F ([Assumption 2](#)) also imply convergence of the expansion coefficients of the embedding, $\lim_{N \rightarrow \infty} F_k^{(\alpha_N)} = F_k$ for every $k \in \mathbb{N}$, as well as convergence of each of the terms of the sequence $\mathcal{F}^{(\alpha_N)} \vec{V}$, i.e., $\lim_{N \rightarrow \infty} (\mathcal{F}^{(\alpha_N)} \vec{V})_{kl} = (\mathcal{F} \vec{V})_{kl}$ for all $k, l \in \mathbb{N}_0$. The terms of that sequence and the $d_{ijkl}^{(L_D, \alpha)}$ coefficients completely determine the right-hand side of the regression problem in [\(4.11\)](#), so that

$$(5.3) \quad \lim_{N \rightarrow \infty} (\mathcal{D}^{(L, L_D, \alpha_N)} \mathcal{F}^{(L_1, L_2, \alpha_N)} \vec{V})_{ij} = (\mathcal{D}^{(L, L_D)} \mathcal{F}^{(L_1, L_2)} \vec{V})_{ij}, \quad \forall (i, j) \in \mathbb{I}.$$

Similarly, the $d_{ijkl}^{(\alpha_N)}$ coefficients determine the spectrally truncated Gram operators $G^{(L, \eta, L_D, \alpha_N)}$, giving

$$\lim_{N \rightarrow \infty} \langle \varepsilon_{ij}, G^{(L, \eta, L_D, \alpha_N)} \varepsilon_{kl} \rangle_{\ell^2(\mathbb{I})} = \lim_{N \rightarrow \infty} d_{ijkl}^{(L_D, \alpha_N)} = d_{ijkl}^{(L_D)} = \langle \varepsilon_{ij}, G^{(L, \eta, L_D)} \varepsilon_{kl} \rangle_{\ell^2(\mathbb{I})},$$

for all $(i, j), (k, l) \in \ell^2(\mathbb{I})$, where $G^{(L, \eta, L_D)}$ was defined in [\(3.21\)](#).

Next, we recall the following result on convergence of pseudoinverses of finite-rank operators (e.g., [\[66, Theorem 4.2\]](#)).

Lemma 5.3. *Suppose that A_τ is a family of operators on a finite-dimensional Hilbert space converging in norm as $\tau \rightarrow \infty$ to an operator A . Then*

$$\lim_{\tau \rightarrow \infty} \|A_\tau^+ - A^+\| = 0 \quad \text{iff} \quad \lim_{\tau \rightarrow \infty} \text{rank } A_\tau = \text{rank } A.$$

Since $\eta > 0$, the norm convergence of $G^{(L, \eta, L_D, \alpha)}$ to $G^{(L, \eta, L_D)}$ implies rank convergence, so by [Lemma 5.3](#) $G^{(L, \eta, L_D, \alpha), +}$ converges to $G^{(L, \eta, L_D), +}$ in the operator norm induced from $\ell^2(I_L)$. It follows that as $N \rightarrow \infty$ the solution vector $b^{(\ell, \alpha_N)}$ from (4.11) converges to $b^{(\ell)}$ from (3.23). As a result, the data-driven vector field $\vec{V}^{(\ell, \alpha)}$ converges to $\vec{V}^{(\ell)}$ from (3.24); that is, $\lim_{N \rightarrow \infty} \vec{V}^{(\ell, \alpha_N)}(y) = \vec{V}^{(\ell)}(y)$, where the convergence is uniform for y in compact subsets of \mathbb{R}^d .

5.1.2. Limit of infinite spectral resolution. We now show that $\vec{V}^{(\ell)} \equiv \vec{V}^{(L, L_1, L_2, \eta, L_D)}$ converges to \vec{V} in $L^2(\nu_*; \mathbb{R}^d)$ in the iterated limit $L \rightarrow \infty$ after $L_1 \rightarrow \infty$ after $L_2 \rightarrow \infty$ after $\eta \rightarrow 0^+$ after $L_D \rightarrow \infty$. Since in [subsection 3.4](#) we have already shown that $V^{(L)}$ converges to V as $L \rightarrow \infty$ in Hodge norm (see, in particular, (3.19)), it is enough to show that $\vec{V}^{(\ell)}$ converges to $\vec{V}^{(L)}$ from (3.20).

First, by definition of the $d_{ijkl}^{(L_D)}$ coefficients we have $\lim_{L_D \rightarrow \infty} d_{ijkl}^{(L_D)} = d_{ijkl}$ for each $(i, j) \in \mathbb{I}$ and $(k, l) \in \hat{\mathbb{I}}$. This implies that

$$(5.4) \quad \lim_{L_D \rightarrow \infty} \mathcal{D}^{(L, L_D)} \mathcal{F}^{(L_1, L_2)} \vec{V} = \mathcal{D}_L \mathcal{F}^{(L_1, L_2)} \vec{V}$$

in $\ell^2(\mathbb{I})$ norm. Moreover, letting $G^{(L, \eta)}$ be the spectrally truncated Gram operator defined analogously to $G^{(L, \eta, L_D)}$ in (3.21) using the d_{ijkl} coefficients instead of $d_{ijkl}^{(L_D)}$, we have that $G^{(L, \eta, L_D)}$ converges to $G^{(L, \eta)}$ in norm and rank (the latter since $\eta > 0$) as $L \rightarrow \infty$. Thus, by [Lemma 5.3](#) and (5.4), it follows that the regression solution $b^{(\ell)}$ in (3.23) converges to $b^{(L, L_1, L_2, \eta)} = G^{(L, \eta), +} \mathcal{D}_L \mathcal{F}^{(L_1, L_2)}$.

Next, we consider the $\eta \rightarrow 0^+$ limit. Given a self-adjoint operator A on a finite-dimensional Hilbert space H with orthonormal eigendecomposition $Au_j = \alpha_j u_j$ and a function $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(A)$ will denote the corresponding operator-valued function from the functional calculus, $f(A) = \sum_j f(\alpha_j) u_j \langle u_j, \cdot \rangle_H$. Note that $\chi_{(\eta, \infty)}(A)$ for the characteristic function $\chi_{(0, \infty)}$ of $(0, \infty)$ gives the spectral truncation of A analogously to (3.21). We will use the shorthand notation $A_\eta \equiv \chi_{(\eta, \infty)}(A)$.

Lemma 5.4. *Let A be a positive operator on a finite-dimensional Hilbert space and A_η the corresponding spectrally truncated operators for $\eta > 0$. Then, we have convergence of the respective pseudoinverses, $\lim_{\eta \rightarrow 0^+} \|A_\eta^+ - A^+\| = 0$.*

Proof. Let $\alpha_* > 0$ be strictly smaller than the smallest strictly positive eigenvalue of A . Fix a continuous function $f: \mathbb{R} \rightarrow \mathbb{R}$ such that $f(\eta) = 1/\eta$ for $\eta \in (\alpha_*, \infty)$ and $f(\eta) = 0$ for $\eta \leq 0$. Then, we have $f(A) = A^+$ and there exists $\eta_* > 0$ such that $f(A_\eta) = A_\eta^+$ for all $\eta \in (0, \eta_*)$. It then follows from the continuous functional calculus that

$$\lim_{\eta \rightarrow 0^+} A_\eta^+ = \lim_{\eta \rightarrow 0^+} f(A_\eta) = f\left(\lim_{\eta \rightarrow 0^+} A_\eta\right) = f(A) = A^+,$$

where the limits are taken in operator norm. ■

Applying [Lemma 5.4](#) for the spectrally truncated approximations $G^{(L,\eta)}$ of the Gram operator G_L , we deduce that as $\eta \rightarrow 0^+$, $G^{(L,\eta),+}$ converges to G_L^+ and thus that $b^{(L,L_1,L_2,\eta)}$ converges to $b^{(L,L_1,L_2)} := G_L^+ \mathcal{D}_L \mathcal{F}^{(L_1,L_2)} \vec{V}$. In this limit, our approximate vector field, $\vec{V}^{(L,L_1,L_2)} \in C^\infty(\mathbb{R}^d; \mathbb{R}^d)$ is given by

$$(5.5) \quad \vec{V}^{(L,L_1,L_2)} = \sum_{i=0}^{L-1} \sum_{j=1}^J \sum_{k=0}^{L_1-1} \sum_{l=0}^{L_2-1} d_{ijkl} b_{ij}^{(L,L_1,L_2)} F_k \varphi_l$$

Note that this reconstruction is close to $\vec{V}^{(\ell)}$ in [\(3.24\)](#), but without the use of the L_D -truncation for the d -coefficients and η -regularization.

Next, observe that $\vec{V}^{(L,L_1,L_2)}(F(x)) = \sum_{i=0}^{L-1} \sum_{j=1}^J B_{ij}^{(L,L_1,L_2)} F(x)$ for all $x \in \mathcal{M}$. We use regularity of the SEC frame elements and the embedding F to show that $B_{ij}^{(L,L_1,L_2)} F$ converges in $L^2(\nu; \mathbb{R}^d)$ as $L_1, L_2 \rightarrow \infty$.

Lemma 5.5. *The frame elements $\{B_{ij}\}_{(i,j) \in \mathbb{I}}$ are uniformly norm-bounded as operators from $C^1(\mathcal{M})$ to $L^2(\nu)$.*

Proof. The claim follows by direct verification. With $f \in C^1(\mathcal{M})$, we have

$$\begin{aligned} \|B_{ij} \triangleright f\|_{L^2(\nu)}^2 &= \langle \phi_i^2, (\nabla \phi_j \cdot \nabla f)^2 \rangle_{L^2(\nu)} \leq \|\phi_i^2\|_{L^1(\nu)} \|\nabla \phi_j \cdot \nabla f\|_{L^\infty(\nu)}^2 = \|\nabla \phi_j \cdot \nabla f\|_{L^\infty(\nu)}^2 \\ &\leq \|\phi_j\|_{C^1(\mathcal{M})} \|f\|_{C^1(\mathcal{M})}^2 \leq M \|f\|_{C^1(\mathcal{M})}^2, \end{aligned}$$

where $M = \max\{\|\phi_j\|_{C^1(\mathcal{M})}\}_{j=1}^J$. ■

Since F is C^∞ , $\Pi_{L_2} F$ converges to F componentwise in $C^1(\mathcal{M})$ norm. Thus, by [Lemma 5.5](#), ■

$$(5.6) \quad \lim_{L_1 \rightarrow \infty} \lim_{L_2 \rightarrow \infty} B_{ij}^{(L_1,L_2)} F = \lim_{L_1 \rightarrow \infty} \lim_{L_2 \rightarrow \infty} \Pi_{L_1} B_{ij} \Pi_{L_2} F = \lim_{L_1 \rightarrow \infty} \Pi_{L_1} B_{ij} F = B_{ij} F,$$

and the convergence is uniform with respect to $(i, j) \in \mathbb{I}$.

Using similar arguments, one can verify that $\mathcal{D}_L \mathcal{F}$ is bounded as an operator from $C^1(Y; \mathbb{R}^d)$ to $\ell^2(\mathbb{I})$, giving $\lim_{L_1 \rightarrow \infty} \lim_{L_2 \rightarrow \infty} \mathcal{D}_L \mathcal{F}^{(L_1,L_2)} \vec{V} = \mathcal{D}_L \mathcal{F} \vec{V}$, and thus

$$(5.7) \quad \lim_{L_1 \rightarrow \infty} \lim_{L_2 \rightarrow \infty} b^{(L,L_1,L_2)} = b^{(L)}.$$

Using [\(5.6\)](#) and [\(5.7\)](#) in [\(5.5\)](#), we conclude that

$$\lim_{L_1 \rightarrow \infty} \lim_{L_2 \rightarrow \infty} \vec{V}^{(L,L_1,L_2)} = \vec{V}^{(L)}$$

in $L^2(\nu^*; \mathbb{R}^d)$. This completes our proof of [Theorem 5.1](#). ■

5.2. Orbit reconstruction. As discussed in [section 3](#), for every $x \in \mathcal{M}$, the SEC-approximated vector field $\vec{V}^{(L)}$ from [\(3.7\)](#) induces a dynamical trajectory $t \mapsto y^{(L)}(x, t) \in \mathbb{R}^d$ in data space, where $y^{(L)}(x, \cdot) \in C^1(\mathbb{R}; \mathbb{R}^d)$ is the solution of the initial-value problem

$$(5.8) \quad \dot{y}^{(L)}(x, t) = \vec{V}^{(L)}(y^{(L)}(x, t)), \quad y^{(L)}(x, 0) = F(x).$$

In this subsection, we assess the error of such orbits relative to orbits $t \mapsto y(x, t)$ generated by the true vector field V on the manifold, i.e., $y(x, t) = F(\Phi^t(x))$. Orbits generated by the vector fields $\vec{V}^{(\ell)}$ and $\vec{V}^{(\ell, \alpha)}$ from (3.24) and (4.12), respectively, which involve additional approximations compared to $V^{(L)}$, can be treated similarly and obey analogous bounds as orbits under $V^{(L)}$ in the appropriate asymptotic limits.

As a measure of expected error over initial conditions, we will use the norm

$$\|f\|_{L^1(\nu; \mathbb{R}^d)} := \int_{\mathcal{M}} \|f(x)\|_1 d\nu(x) \equiv \sum_{j=1}^d \|f_j\|_{L^1(\nu)},$$

where $f = (f_1, \dots, f_d): \mathcal{M} \rightarrow \mathbb{R}^d$ has components $f_j \in L^1(\nu)$. We will also make use of the following seminorms on $C(\mathcal{M}; \mathbb{R}^d)$,

$$|f|_{1,2} = \sum_{j=1}^d \|\nabla f_j\|_{\mathbb{H}} \equiv - \sum_{j=1}^d \langle f_j, \Delta f_j \rangle_{L^2(\nu)}, \quad |f|_{2,\infty} = \sum_{j=1}^d \|\Delta f_j\|_{L^\infty(\nu)}.$$

With these definitions, we have the following bound on the expected orbit distance over initial conditions based on the distance between two vector fields in the Hodge norm.

Proposition 5.6. *Assume \mathcal{M} , V as in Assumption 1. Let $\hat{V} \in C(\mathcal{M}, TM)$ be any other continuous vector field on \mathcal{M} generating a flow $\hat{\Phi}^t: \mathcal{M} \rightarrow \mathcal{M}$ and associated data space trajectory map $\hat{y}(x, t) = F(\hat{\Phi}^t(x))$. Fix $T > 0$ and set*

$$\alpha := \left\| V - \hat{V} \right\|_{\mathbb{H}} \max_{t \in [0, T]} |F \circ \Phi^t|_{1,2}, \quad \beta = \|\hat{V}\|_{\mathbb{H}} \max_{t \in [0, T]} \left| F \circ \Phi^t - F \circ \hat{\Phi}^t \right|_{2,\infty}.$$

Then, for every $t \in [0, T]$, the error norm $\varepsilon(t) = \|y(t, \cdot) - \hat{y}(t, \cdot)\|_{L^1(\nu; \mathbb{R}^d)}$ satisfies the following bounds on growth rate and value:

$$\dot{\varepsilon}(t) \leq \alpha + \beta \sqrt{\varepsilon(t)}, \quad \varepsilon(t) \leq t \left(\sqrt{\alpha} + \frac{\beta \sqrt{t}}{2} \right)^2.$$

Proof. See Appendix B. ■

In the case of the SEC-approximated vector field, $\hat{V} \equiv V^{(L)}$, the Hodge norm convergence (3.19) implies that the corresponding error term $\alpha \equiv \alpha_L$ in Proposition 5.6 can be made arbitrarily small by increasing the spectral resolution parameter L . On the other hand, convergence of $V^{(L)}$ to V in Hodge norm does not, in general, provide control of $\beta \equiv \beta_L$ as L increases. Thus, we cannot conclude from (3.19) and Proposition 5.6 alone that the expected error $\varepsilon(t)$ can be made arbitrarily small by increasing L on a fixed time interval $[0, T]$. Nevertheless, we can still conclude that the short-term error and error growth rate can be made arbitrarily small. In particular, for $T_L = \min\{T, 2\sqrt{\alpha_L}/\beta_L\}$ and $t \in [0, T_L]$ we have $\varepsilon(t) \leq 2\sqrt{\alpha_L}t$. Since $\alpha_L \rightarrow 0$ as $L \rightarrow \infty$, the expected error $\varepsilon(t)$ can be made arbitrarily small on an L -independent time interval if $\liminf_{L \rightarrow \infty} T_L > 0$.

Table 2

Dataset attributes and numerical approximation parameters used in our experiments.

	Circle examples	2-Torus examples
Manifold \mathcal{M}	S^1	\mathbb{T}^2
Data space	\mathbb{R}^2	\mathbb{R}^3
Embedding map F	(6.4)	(6.5)
Training dataset size N	800	$150 \times 150 = 22,500$
Kernel bandwidth ϵ	0.2	0.0966
Number of gradient fields J	10	20
SEC discretization parameters L, L_1	20	120
SEC discretization parameters L_D, L_2	40	280
Gram operator spectral truncation parameter η	0.01	0.05

6. Numerical examples. We apply the SEC framework described in sections 2–4 to various dynamical systems on the unit circle S^1 and 2-torus \mathbb{T}^2 , viewed as embedded submanifolds of \mathbb{R}^2 and \mathbb{R}^3 under standard (smooth) embeddings, respectively.

In each example, we use training data $\{y_n \in \mathbb{R}^d\}_{n=1}^N$ and $\{\vec{v}_n \in \mathbb{R}^d\}$ for the embedded manifold and dynamical vector field, respectively, generated from a set $\{x_n \in \mathcal{M}\}_{n=1}^N$ of points on the manifold lying in a uniform grid. Using the y_n and \vec{v}_n data (but assuming no knowledge of the underlying manifold points x_n), we apply the diffusion maps algorithm (Algorithm A.1) to compute the eigenpairs $(\lambda_j^{(\alpha)}, \phi_j^{(\alpha)})$ of the Laplacian, and build the SEC approximation $V^{(\ell, \alpha)}$ of the vector field using Algorithm A.3. We then reconstruct the pushforward vector field $\vec{V}^{(\ell, \alpha)}$ in data space using (4.12) and solve the associated initial-value problem (5.8).

We assess the performance of our scheme quantitatively by computing normalized mean square errors (R^2 coefficients) of the reconstructed vector fields,

$$R^2 = \frac{\sum_{n=1}^N \|\vec{V}(x_n) - \vec{V}^{(\ell, \alpha)}(x_n)\|_2^2}{\sum_{n=1}^N \|\vec{V}(x_n)\|_2^2},$$

which provide estimates of the SEC approximation accuracy with respect to the Hodge norm on vector fields. We also assess performance qualitatively by comparing dynamical trajectories in embedding space \mathbb{R}^d under the true vector field \vec{V} and the SEC approximation $\vec{V}^{(\ell, \alpha)}$. The latter are given by solutions of the initial-value problem (cf. (5.8))

$$\dot{y}^{(\ell, \alpha)}(t) = \vec{V}^{(\ell, \alpha)}(y^{(\ell, \alpha)}(t)), \quad y^{(\ell, \alpha)}(0) = y_0 \in \mathbb{R}^d.$$

In our experiments, we will take the initial condition y_0 to lie in the embedded manifold $F(\mathcal{M})$, i.e., $y_0 = F(x_0)$ for some $x_0 \in \mathcal{M}$. The true dynamical trajectory that we compare against is thus given by $y(t) = F(x(t))$, where $x(t) = \Phi^t(x_0)$ solves the initial-value problem

$$(6.1) \quad \dot{x}(t) = V(x(t)), \quad x(0) = x_0$$

on the manifold \mathcal{M} (see subsection 5.2).

The dataset attributes and approximation parameters used in our experiments are summarized in Table 2.

6.1. Dynamical systems on the circle. Every continuous vector field V on S^1 is of the form

$$(6.2) \quad V(\theta) = h(\theta) \frac{\partial}{\partial \theta},$$

where h is a continuous function, $\theta \in [0, 2\pi)$ is an angle coordinate, and $\frac{\partial}{\partial \theta} \in T_\theta(S^1)$ the associated coordinate tangent vector. Here, we consider three vector fields with distinct dynamical behavior:

- A uniform rotation, $h(\theta) = 1$.
- A dynamical flow with fixed points, $h(\theta) = 1 + c \cos \theta$ for a constant $c > 1$, set here to $c = 1.5$.
- A variable-speed rotation, with $h(\theta) = e^{c \cos \theta}$ for a constant $c \neq 0$, set here to $c = 0.5$.

We denote the corresponding vector fields as V_1 , V_2 , and V_3 , respectively. Note that V_2 has two unstable equilibria given by the solutions of $h(\theta) = 0$. For our chosen value of c these turn out to be

$$(6.3) \quad \theta \approx 2.30 \text{ rad}, \quad \theta \approx 3.98 \text{ rad}.$$

The resulting dynamical flow is monotonic and takes place along a circular arc connecting these equilibria. Meanwhile, the variable-speed rotation under V_3 attains its maximum speed $e^c \approx 1.65$ at $\theta = 0$ and its minimum speed $e^{-c} \approx 0.61$ at $\theta = \pi$.

In all three cases, we use the standard embedding $F: S^1 \rightarrow \mathbb{R}^2$, given by

$$(6.4) \quad F(\theta) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

The associated pushforward of the vector field from (6.2) is

$$\vec{V}(y) = F_* V(y) = V F(\theta) = h(\theta) \frac{\partial}{\partial \theta} \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} = h(\theta) \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix},$$

where $y = F(\theta)$. For this embedding of the circle, the eigenvalue problem (2.1) for the Laplacian can be analytically solved, giving eigenvalues $\lambda_j = -j^2$ and

$$\phi_j(\theta) = \begin{cases} 1/\sqrt{2\pi}, & j = 0, \\ \sin\left(\frac{(j+1)\theta}{2}\right)/\sqrt{\pi}, & j \text{ is odd}, \\ \cos\left(\frac{j\theta}{2}\right)/\sqrt{\pi}, & j \text{ is even}. \end{cases}$$

for corresponding eigenfunctions normalized to unit norm on $L^2(\nu)$. Consequently, F can be expressed componentwise as a finite linear combination of Laplace–Beltrami eigenfunctions,

$$F(\theta) = \sqrt{\pi} \begin{pmatrix} \phi_2(\theta) \\ \phi_1(\theta) \end{pmatrix}.$$

Our training datasets are based on $N = 800$ equidistant samples obtained from the embedding (6.4) and the vector fields V_1 , V_2 , and V_3 . We perform diffusion maps using the

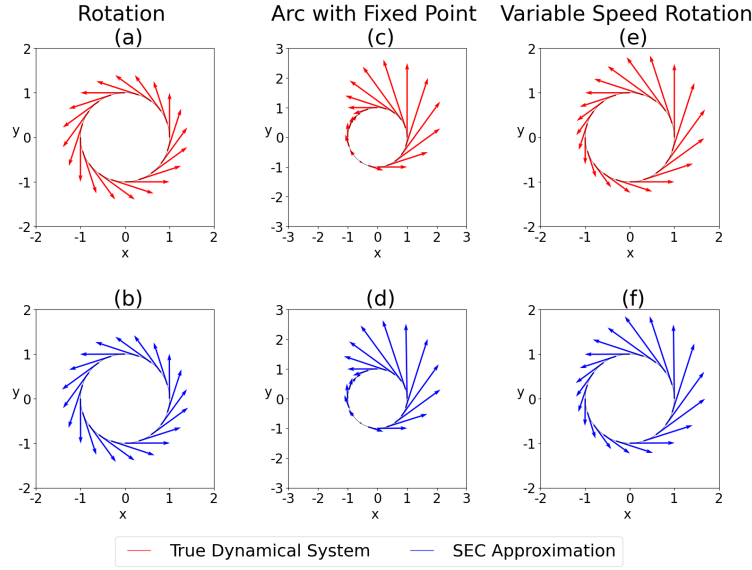


Figure 2. Quiver plots of the true (\vec{V} ; a, c, e) and SEC-approximated vector fields ($\vec{V}^{(\ell, \alpha)}$; b, d, f) on the unit circle embedded in \mathbb{R}^2 . (a, b) Circle rotation, V_1 ; (c, d) connecting arc, V_2 ; and (e, f) variable-speed rotation, V_3 . The panels show tangent vectors on the unit circle in \mathbb{R}^2 for a subset of points in the training dataset.

kernel bandwidth parameter $\epsilon = 0.2$. In all three cases, we use $J = 10$ gradient fields $\nabla\phi_j$ to build SEC frames and set the remaining SEC spectral resolution parameters to $L = L_1 = 20$ and $L_D = L_2 = 40$. These parameter values were selected from a suite of cross-validation experiments. Even though L , L_D , L_1 , and L_2 are independent discretization parameters, experimentally we found that fixing $L = L_1$ and $L_D = L_2$ yields accurate reconstruction results while simplifying the parameter tuning process. The R^2 coefficients for our chosen parameter values are 0.999731, 0.999220, and 0.999632, respectively for V_1 , V_2 , and V_3 .

Figure 2 shows a comparison of the true vs. SEC-approximated vector fields for the three examples, visualized as quiver plots on representative training points on the unit circle in \mathbb{R}^2 . As expected from the large R^2 coefficients, the SEC approximations are seen to accurately reconstruct the true vector fields, exhibiting a high degree of tangency on the unit circle.

Recall that, unlike the true pushforward vector field \vec{V} , the SEC approximation $\vec{V}^{(\ell, \alpha)}$ can be evaluated at arbitrary points in \mathbb{R}^2 , including points which do not lie in the embedded manifold $Y = F(\mathcal{M})$. In Figure 3 we plot these out-of-sample extensions on the rectangular domain $[-5, 5] \times [-5, 5]$ that contains Y . Intriguingly, the vector fields $\vec{V}^{(\ell, \alpha)}$ appear to consistently extend the angular structure of the true vector fields; that is, we have $\vec{V}^{(\ell, \alpha)}(y) \approx a(r)\vec{V}(F(\theta))$ for $y = r(\cos \theta, \sin \theta)$ and a near-constant amplitude function $a(r)$. This behavior suggests favorable generalization behavior of the SEC approximations to points outside the support of the sampling distribution in data space. This could aid, e.g., orbit stability of (5.8) under perturbations of the initial condition and/or error due to numerical timestepping.

Next in, Figures 4 and 5, we examine representative dynamical trajectories $y(t)$ and $y^{(\ell, \alpha)}(t)$ under the true and SEC-approximated vector fields, respectively, visualized as trace

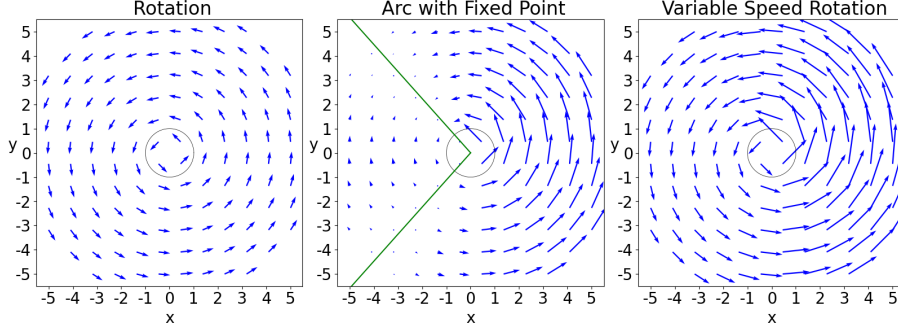


Figure 3. As in Figure 2, but for quiver plots of the SEC-approximated vector fields $\vec{V}^{(\ell, \alpha)}$ on the box $[-5, 5] \times [-5, 5] \subset \mathbb{R}^2$. The unit circle $Y = F(\mathcal{M})$ supporting the sampling distribution of the training data is drawn in each panel for reference. The green lines in panel (b) emanating from the origin depict the angular coordinates of the fixed points of V_2 from (6.3). Notice that $\vec{V}^{(\ell, \alpha)}$ consistently extends the pushforward \vec{V} from Y to other locations in \mathbb{R}^2 .

plots in \mathbb{R}^2 (Figure 4) and time series of the coordinates (Figure 5). To obtain $y^{(\ell, \alpha)}(t)$ we solve (6.1) for an arbitrary initial condition $y_0 = F(\theta_0)$. We obtain $y(t)$ by first solving (6.1) (analytically, in the case of V_1 , and numerically for V_2 and V_3) with initial condition $\theta_0 \in \mathcal{M}$. It is clear that the SEC provides a faithful reconstruction of the true dynamics in all three cases. It should be noted that qualitative reconstruction of connecting arcs under V_2 is non trivial; in particular, vector field approximations that do not exhibit fixed points would result in periodic trajectories that are qualitatively different from the true dynamics. In separate calculations, we have used the Newton method to verify that the SEC-approximation of V_2 does indeed exhibit fixed points (to numerical precision) in locations close to the true fixed points in (6.3).

6.2. Dynamical systems on the 2-torus. We consider the torus \mathbb{T}^2 embedded in \mathbb{R}^3 via the embedding $F : [0, 2\pi) \times [0, 2\pi) \rightarrow \mathbb{R}^3$,

$$(6.5) \quad F(\theta_1, \theta_2) = \begin{pmatrix} (a + b \cos \theta_2) \cos \theta_1 \\ (a + b \cos \theta_2) \sin \theta_1 \\ b \sin \theta_2 \end{pmatrix},$$

Here, $a > 0$ is the radius of the latitude circle with angle coordinate θ_1 and $b > 0$ is the radius of the meridian circle with coordinate angle θ_2 . We set $a = 5/3$ and $b = 3/5$ in all of our numerical experiments. Note that, unlike the circle embedding from (6.4) the components of the torus embedding (6.5) are not finite linear combinations of Laplace–Beltrami eigenfunctions—this is due to non-translation-invariance of the induced Riemannian metric \mathbf{g} from the torus embedding in \mathbb{R}^3 . Moreover, \mathbf{g} has nonzero curvature in this example.

Analogously to (6.2), every continuous vector field on \mathbb{T}^2 can be expressed as

$$(6.6) \quad V = h_1(\theta_1, \theta_2) \frac{\partial}{\partial \theta_1} + h_2(\theta_1, \theta_2) \frac{\partial}{\partial \theta_2}$$

for continuous functions h_1 and h_2 . The corresponding pushforward vector field under F

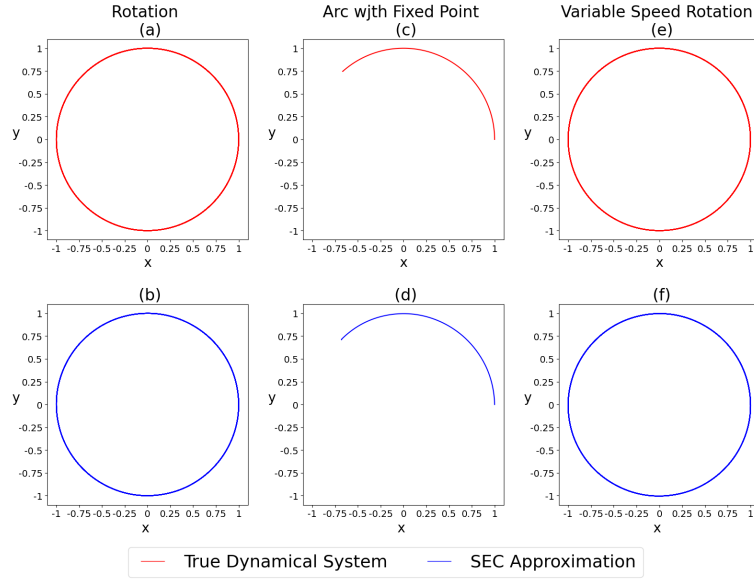


Figure 4. Dynamical trajectories in \mathbb{R}^2 under the true $(y(t))$; a, c, e) and SEC-approximated vector fields $(y^{(\ell, \alpha)})$; b, d, f) of our circle examples, V_1 (a, b), V_2 (c, d), and V_3 (e, f). All trajectories evolve in a counter-clockwise direction.

from (6.5) is given by

$$(6.7) \quad \vec{V}(y) = (F_*V)(y) = \begin{pmatrix} -h_1(\theta_1, \theta_2)(a + b \cos \theta_2) \sin \theta_1 - h_2(\theta_1, \theta_2)(a + b \sin \theta_2) \cos \theta_1 \\ h_1(\theta_1, \theta_2)(a + b \cos \theta_2) \cos \theta_1 - h_2(\theta_1, \theta_2)(a + b \sin \theta_2) \sin \theta_1 \\ h_2(\theta_1, \theta_2)b \cos \theta_2 \end{pmatrix},$$

where $y = F(\theta_1, \theta_2)$. In what follows, we consider three vector fields in this family:

- A rational rotation with $h_1(\theta_1, \theta_2) = h_2(\theta_1, \theta_2) = 1$.
- An irrational rotation with $h_1(\theta_1, \theta_2) = 1$ and $h_2(\theta_1, \theta_2) = \alpha \in \mathbb{R} \setminus \mathbb{Q}$.
- A Stepanoff flow [59] with

$$\begin{aligned} h_1(\theta_1, \theta_2) &= h_2(\theta_1, \theta_2) + (1 - \alpha)(1 - \cos \theta_2), \\ h_2(\theta_1, \theta_2) &= \alpha(1 - \cos(\theta_1 - \theta_2)), \end{aligned}$$

for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

We set $\alpha = \sqrt{20}$ in both of the irrational rotation and Stepanoff flow examples. Similarly to the circle examples from subsection 6.1, we denote the corresponding vector fields as V_1 , V_2 , and V_3 , respectively. Figure 6 visualizes these vector fields through quiver plots on the periodic domain $[0, 2\pi)^2$.

All of V_1 , V_2 , and V_3 have 0 divergence with respect to the Haar measure on \mathbb{T}^2 (i.e., $\partial_1 h_1 + \partial_2 h_2 = 0$), which implies that the corresponding flows are measure-preserving for the Haar measure. However, the divergence of these vector fields is nonzero with respect to the Riemannian volume measure ν induced from F , meaning that ν is not preserved by the dynamics. It should also be noted that the Stepanoff vector field V_3 has a fixed point at $(0, 0)$

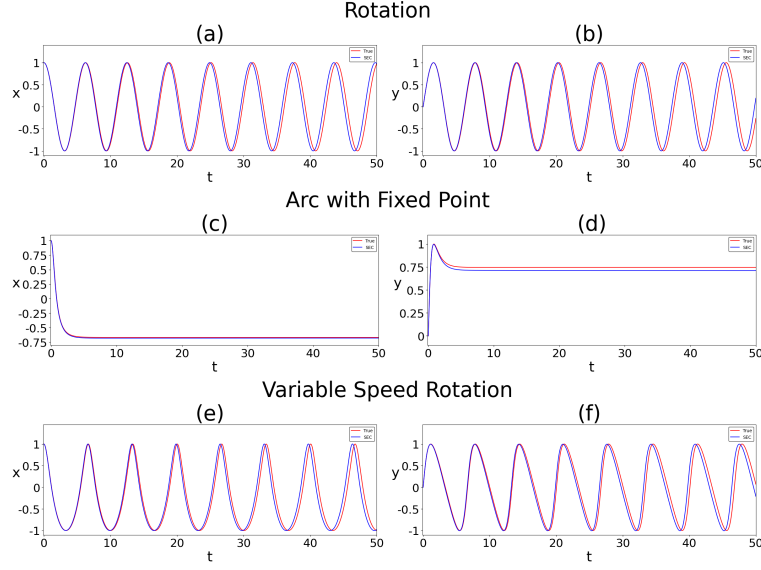


Figure 5. As in [Figure 4](#), but showing time series of the x and y components of the $y(t)$ and $y^{(\ell, \alpha)}(t)$ trajectories.

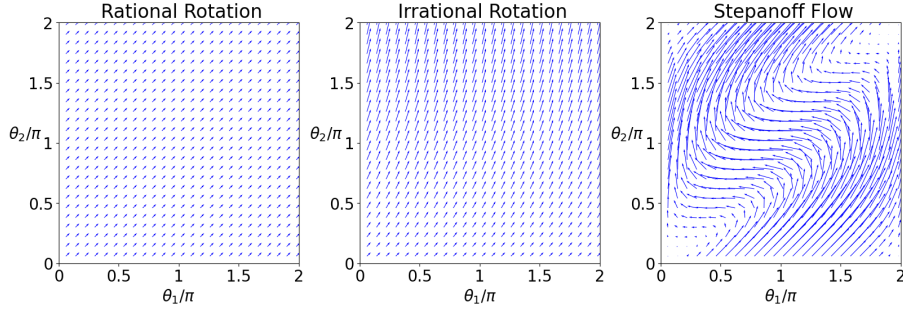


Figure 6. Quiver plots of the vector fields V_1 (rational rotation with frequencies $(1, 1)$), V_2 (irrational rotation with frequencies $(1, \sqrt{20})$), and V_3 (Stepanoff flow with parameter $\alpha = \sqrt{20}$). Arrows represent the components of the vector fields with respect to the coordinate basis fields $\frac{\partial}{\partial \theta_1}$ and $\frac{\partial}{\partial \theta_2}$ associated with the latitude (θ_1) and meridian (θ_2) angles on \mathbb{T}^2 .

(see [Figure 6](#)), making the corresponding flow weak-mixing in a topological sense. In [\[59\]](#) it is shown that the Stepanoff flow is topologically conjugate to a time-reparameterized irrational rotation with basic frequencies $(1, \alpha)$, for a singular time-reparameterization function at $(0, 0)$. Since we use the same value $\alpha = \sqrt{20}$ in both of the irrational rotation and Stepanoff flow examples, the two systems are related in this sense.

Our training dataset for each torus example consists of $N = 150 \times 150 = 22,500$ points sampled on a uniform grid in \mathcal{M} with respect to the angle coordinates (θ_1, θ_2) . Note that the sampling is non-uniform with respect to Riemannian measure but this does not affect the asymptotic convergence of the diffusion maps algorithm (see [Appendix A.1](#)). Throughout, we

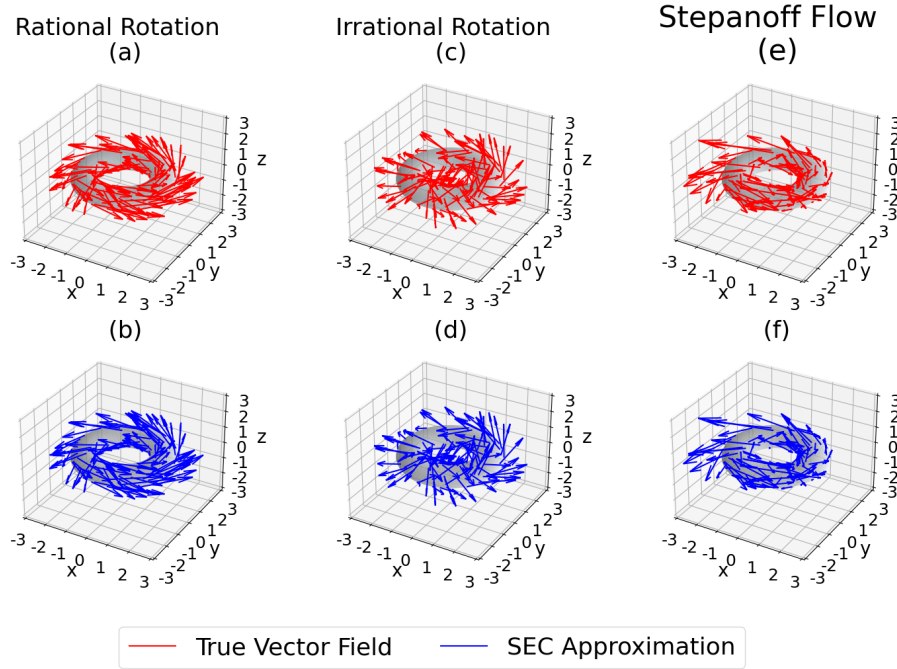


Figure 7. Quiver plots of the true (\vec{V} ; a, c, e) and SEC-approximated vector fields ($\vec{V}^{(\ell, \alpha)}$; b, d, f) on the 2-torus embedded in \mathbb{R}^3 . (a, b) Rational rotation, V_1 ; (c, d) irrational rotation, V_2 ; and (e, f) Stepanoff flow, V_3 . The embedded torus $Y = F(\mathbb{T}^2)$ is shown in each panel as a shaded surface for reference. Arrows show tangent vectors on Y for a subset of points in the training dataset.

use the kernel bandwidth parameter $\epsilon = 0.0966$ and the SEC spectral resolution parameters $L = L_1 = 120$ and $L_D = L_2 = 280$. For these parameter choices the R^2 coefficients for the SEC approximations $\vec{V}^{(\ell, \alpha)}$ of V_1 , V_2 , and V_3 are 0.999975, 0.999945, and 0.999855, respectively, indicating high approximation accuracy with respect to the Hodge norm as in the circle examples. The reconstructed vector fields $\vec{V}^{(\ell, \alpha)}$ are plotted together with the true pushforwards \vec{V} in the embedding space \mathbb{R}^3 in Figure 7, where the high approximation accuracy is visually evident.

In Figures 8 and 9 we compare traceplots in \mathbb{R}^3 and the corresponding coordinate time series, respectively, of true dynamical trajectories, $y(t)$, generated by the vector fields V_1 , V_2 , and V_3 , with the corresponding SEC approximations $y^{(\ell, \alpha)}(t)$. These trajectories are computed using similar numerical methods as in the circle examples (see subsection 6.1). As initial conditions, we use $(\theta_{1,0}, \theta_{2,0}) = (0, 0)$ for the rational and irrational rotations. The Stepanoff flow trajectories are initialized at $(\theta_{1,0}, \theta_{2,0}) = (\pi + 0.3, \pi + 0.5)$ to avoid the trivial solution starting at the fixed point. The results demonstrate accurate trajectory reconstruction for the rational and irrational rotation systems over the entire integration intervals. In the case of the Stepanoff flow, while $y^{(\ell, \alpha)}(t)$ accurately tracks the true trajectory $y(t)$ for $t \lesssim 1.5$, at later times we see significant discrepancy buildup, leading to an apparent decorrelation of the true trajectories. While this behavior illustrates the challenges of orbit reconstruction in spite of high vector field reconstruction accuracy in the Hodge norm (see subsection 5.2), we

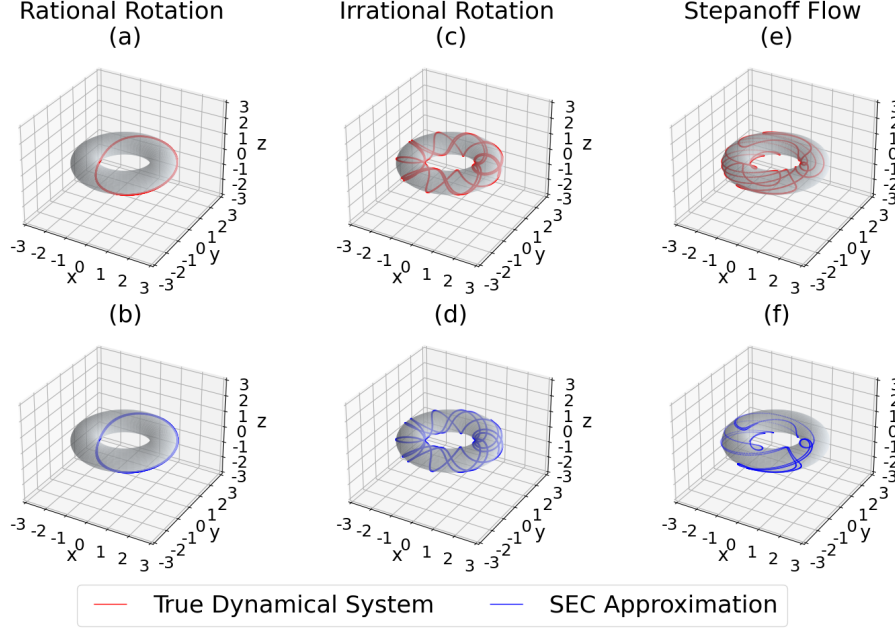


Figure 8. Dynamical trajectories in \mathbb{R}^3 under the true $(y(t); a, c, e)$ and SEC-approximated vector fields $(y^{(\ell, \alpha)}(t); b, d, f)$ of the torus examples V_1 (a, b), V_2 (c, d), and V_3 (e, f).

should note that trajectory deviation does not necessarily indicate failure of the SEC in this system that exhibits sensitive dependence on initial conditions. In particular, the fact that $y^{(\ell, \alpha)}$ remains tangent to the embedded torus in \mathbb{R}^3 (see Figure 8), suggests that the SEC consistently represents this torus as an invariant set, despite nonlinearity and curvature of the embedding F .

7. Discussion. In this paper, we have presented a geometrical technique for data-driven approximation of dynamical systems on Riemannian manifolds. Central to our approach is a formulation of supervised learning of dynamical systems as an approximation problem in a hypothesis space of vector fields (generators of dynamics), constructed from data sampled on the manifold using the spectral exterior calculus (SEC) [7] technique. The SEC builds frames for L^2 spaces of vector fields consisting of products of eigenfunctions of the Laplace–Beltrami operator and their gradients, estimated using kernel techniques along with a carré du champ identity (1.4) for gradient computation. The learned dynamical vector field, $V^{(\ell, \alpha)}$, is determined by solving a least-squares regression problem in the SEC frame using tangent vectors sampled on the manifold, \mathcal{M} , as training data, and can be represented intrinsically as a linear operator acting on functions by directional derivatives. This operator-theoretic representation allows reconstructing $V^{(\ell, \alpha)}$ as an embedded vector field $\vec{V}^{(\ell, \alpha)}$ (“arrow field”) in Euclidean data space \mathbb{R}^d through the pushforward map, $\vec{V}^{(\ell, \alpha)} = V^{(\ell, \alpha)}F$, associated with the embedding F of the manifold. The reconstructed vector field $\vec{V}^{(\ell, \alpha)}$ can then be used for a variety of tasks, including approximation of dynamical trajectories (initial-value prediction)

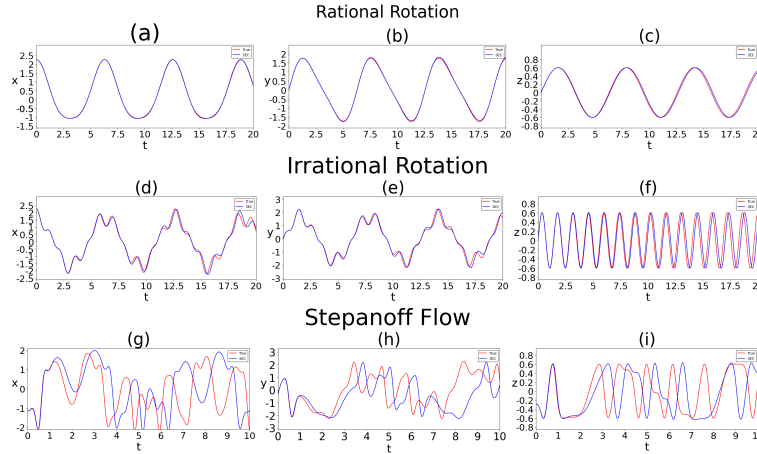


Figure 9. As in Figure 4, but showing time series of the x , y , and z components of the $y(t)$ and $y^{(\ell, \alpha)}(t)$ trajectories.

which was studied in this paper.

Our analytical results include a proof of convergence of the SEC approximation $\vec{V}^{(\ell, \alpha)}$ in a Hodge norm on vector fields (Theorem 5.1), as well as a bound on trajectory reconstruction error averaged over initial conditions on the manifold (Proposition 5.6). These results were obtained under general regularity assumptions on the dynamical system, sampling distribution of the training data, and spectral accuracy of the kernel-approximated Laplacian (Assumptions 1–4), without requiring foreknowledge of the dynamical system and/or manifold in parametric form.

A novelty of our technique over the conventional approach of representing vector fields componentwise in scalar-valued function spaces is that learning takes place in an intrinsic space of vector fields that can generate arbitrary (nonlinear) flows on the manifold in a manner compatible with its (nonlinear) embedding in data space. In particular, any loss of tangency of $\vec{V}^{(\ell, \alpha)}$ to the embedded image $F(\mathcal{M}) \subset \mathbb{R}^d$ occurs only in the final step of our reconstruction procedure due to finite-rank operator projection, avoiding spurious results due to learning in a hypothesis space containing \mathbb{R}^d -valued functions that are not tangent to $F(\mathcal{M})$. We believe that the carefully chosen hypothesis space, leveraging the SEC and kernel methods, aids the consistency and stability of the learned dynamical system with our approach. In this paper, we presented a suite of numerical experiments involving dynamical systems on the circle and 2-torus that provide encouraging results on vector field reconstruction and approximation of dynamical orbits.

There are several avenues of further investigation stemming from this work. First, besides the local structure of the vector field, a dynamical system exhibits several global properties such as attractors, invariant measures, periodic orbits, and Lyapunov spectra. An interesting topic for future work would be to characterize the extent at which such properties can be reconstructed experimentally with the SEC-based approach. For example, co-existing periodic orbits reflect the complexity of the dynamics. A close match in the periodic orbits up to a certain period of the true and reconstructed systems indicate a good reconstruction. In such

studies, a challenge that one invariably encounters is that global properties of dynamics can be extremely sensitive to and sometimes change discontinuously with changes in the vector field (e.g., [15, 5, 6]), so the focus would be on parameter regimes with sufficient structural stability for data-driven analysis.

Second, the bounds in Proposition 5.6 can be potentially improved by working in spaces of vector fields equipped with stronger norm than the Hodge norm of \mathbb{H} . A good candidate is a combined Hodge–Sobolev norm that admits a weighted SEC frame construction [7]. Recent work [61] has demonstrated accurate reconstruction of Lyapunov spectra with neural network models trained with a loss function that incorporates the Jacobian of the dynamics. It is possible that a Hodge–Sobolev SEC frame is similarly beneficial to statistical reconstruction accuracy by augmenting the pointwise comparison in the current formulation by a C^1 comparison.

Finally, it would be useful to explore ways of reducing computational cost. In our current implementation, both training and prediction phases are carried out using brute-force linear algebra methods (see Figure 1). For a training dataset of size N these methods materialize rank- N kernel matrices with $O(N^2)$ and $O(N)$ time/space cost during training and prediction, respectively. Low-rank approximation methods, such as random feature projection and subsampling [50, 29, 26], provide promising routes for improving efficiency, particularly in timestepping problems requiring large numbers of sequential evaluations of $V^{(\ell, \alpha)}$.

Appendix A. Algorithms.

A.1. Diffusion maps. We give an outline of the diffusion maps algorithm [16] as used in the numerical experiments of section 6.

Our starting point is a Gaussian radial basis function (RBF) kernel $k_\epsilon: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_{>0}$ on data space,

$$(A.1) \quad k_\epsilon(y, y') = \exp\left(-\frac{\|y - y'\|_2^2}{\epsilon^2}\right),$$

parameterized by a bandwidth (lengthscale) parameter $\epsilon > 0$. Let $\alpha = (\epsilon, N)$ as in section 4. Associated with k_ϵ are right and left normalization functions $\deg_r^{(\alpha)}, \deg_l^{(\alpha)} \in C^\infty(\mathbb{R}^d)$, where

$$(A.2) \quad \begin{aligned} \deg_r^{(\alpha)}(y) &= \int_{\mathcal{M}} k_\epsilon(y, F(x)) d\mu_N(x) = \frac{1}{N} \sum_{n=1}^N k_\epsilon(y, y_n), \\ \deg_l^{(\alpha)}(y) &= \int_{\mathcal{M}} \frac{k_\epsilon(y, F(x))}{\deg_r^{(\alpha)}(x)} d\mu_N(x) = \frac{1}{N} \sum_{n=1}^N \frac{k_\epsilon(y, y_n)}{\deg_r^{(\alpha)}(y_n)}. \end{aligned}$$

By construction, both normalization functions are strictly positive with images bounded away from zero on compact sets. Defining $\rho_\alpha: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}_{>0}$ as

$$\rho_\alpha(y, y') = \frac{k_\epsilon(y, y')}{\deg_l^{(\alpha)}(y) \deg_r^{(\alpha)}(y')},$$

we then have that this kernel is C^∞ and Markovian in the sense of (4.1), and induces a Markov operator $P_\alpha: L^2(\mu_N) \rightarrow L^2(\mu_N)$. Because $p_\alpha(x, x') > 0$, this operator is ergodic, i.e., $P_\alpha f = f$ iff $f(x_n)$ is constant for all $n \in \{1, \dots, N\}$.

Let $p_\alpha: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$ be the pullback kernel on the manifold, $p_\alpha(x, y) = \rho_\alpha(F(x), F(y))$ as in [subsection 4.1](#), and define $d_\alpha \in C^\infty(\mathcal{M})$ as

$$d_\alpha(x) = \sqrt{\deg_l^{(\alpha)}(F(x)) / \deg_r^{(\alpha)}(F(x))}.$$

One then verifies that $\tilde{p}_\alpha: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}_{>0}$ with

$$\tilde{p}_\alpha(x, x') = d_\alpha(x)p_\alpha(x, x')/d_\alpha(x')$$

is a symmetric kernel, and thus induces a self-adjoint integral operator $\tilde{P}_\alpha: L^2(\mu_N) \rightarrow L^2(\mu_N)$, where

$$\tilde{P}_\alpha f = \int_{\mathcal{M}} \tilde{p}_\alpha(\cdot, x) f(x) d\mu_N(x) \equiv \frac{1}{N} \sum_{n=1}^N \tilde{p}_\alpha(\cdot, x_n) f(x_n).$$

This operator is related to P_α by the similarity transformation

$$(A.3) \quad \tilde{P}_\alpha = D_\alpha P_\alpha D_\alpha^{-1},$$

where $D_\alpha: L^2(\mu_N) \rightarrow L^2(\mu_N)$ is the multiplication operator by d_α . Computationally, D_α is represented by the $N \times N$ diagonal matrix $\mathbf{D}_\alpha = [D_{ij}^{(\alpha)}]$ with diagonal entries $D_{ii}^{(\alpha)} = d_\alpha(x_i)$. Moreover, we have $\tilde{\mathbf{P}}_\alpha = \mathbf{D}_\alpha \mathbf{P}_\alpha \mathbf{D}_\alpha^{-1}$, where $\tilde{\mathbf{P}}_\alpha = [\tilde{P}_{ij}^{(\alpha)}]$ and $\mathbf{P}_\alpha = [P_{ij}^{(\alpha)}]$ are the matrix representations of \tilde{P}_α and P_α , respectively, with $\tilde{P}_{ij}^{(\alpha)} = \tilde{p}_\alpha(x_i, x_j)/N$ and $P_{ij}^{(\alpha)} = p_\alpha(x_i, x_j)/N$.

By [\(A.3\)](#), P_α and \tilde{P}_α have the same eigenvalues $\Lambda_j^{(\alpha)}$. The eigenvalues are therefore real by self-adjointness of \tilde{P}_α , and satisfy $|\Lambda_j^{(\alpha)}| \leq 1$ with equality for a simple eigenvalue equal to 1 by Markovianity and ergodicity of P_α . It can also be shown that $\Lambda_j^{(\alpha)} > 0$ since the Gaussian kernel k_ϵ is strictly positive-definite. We can thus conclude that the eigenvalues admit the ordering $1 = \Lambda_0^{(\alpha)} > \Lambda_1^{(\alpha)} \geq \Lambda_2^{(\alpha)} \geq \dots \geq \Lambda_{N-1}^{(\alpha)} > 0$.

Next, consider an eigendecomposition of \tilde{P}_α ,

$$\tilde{P}_\alpha e_j^{(\alpha)} = \Lambda_j^{(\alpha)} e_j^{(\alpha)},$$

where the eigenvectors $e_j^{(\alpha)}$ are chosen to be orthogonal in $L^2(\mu_N)$. By the similarity relation [\(A.3\)](#), it follows that

$$\phi_j^{(\alpha)} = D_\alpha^{-1} e_j^{(\alpha)} = e_j^{(\alpha)} / d_\alpha$$

are eigenvectors of P_α corresponding to $\Lambda_j^{(\alpha)}$ and $\{\phi_j^{(\alpha)}\}_{j=0}^{N-1}$ forms a basis of $L^2(\mu_N)$. As a result, the data-driven Laplacian [\(4.2\)](#) is defined uniquely from the eigendecomposition

$$\Delta_\alpha \phi_j^{(\alpha)} = \lambda_j^{(\alpha)} \phi_j^{(\alpha)}, \quad \lambda_j^{(\alpha)} = -\frac{\log \Lambda_j^{(\alpha)}}{c\epsilon^2}.$$

For the Gaussian RBF kernel [\(A.1\)](#) the choice $c = 1/4$, which we adopt throughout, leads to a consistent approximation of the Laplace–Beltrami operator Δ by Δ_α [\[16\]](#).

Let now $w_\alpha \in L^2(\mu_N)$ be a top eigenvector of P_α^* , i.e., $P_\alpha^* w_\alpha = w_\alpha$. Every such eigenvector satisfies $w_\alpha(x_n) = \left(e_0^{(\alpha)}(x_n)\right)^2 / Z_\alpha$ for a normalization constant Z_α . To fix this constant, suppose that $Z_\alpha > 0$, and observe that the eigenvectors $\phi_j^{(\alpha)}$ are orthogonal with respect to the inner product

$$(A.4) \quad \langle \phi_i^{(\alpha)}, \phi_j^{(\alpha)} \rangle_{L^2(\nu_\alpha)} := \int_{\mathcal{M}} \phi_i^{(\alpha)} \phi_j^{(\alpha)} d\nu_\alpha \equiv \int_{\mathcal{M}} \phi_i^{(\alpha)} \phi_j^{(\alpha)} w_\alpha d\mu_N$$

associated with the weighted sampling measure $\nu_\alpha = \sum_{n=1}^N w_\alpha(x_n) \delta_{x_n}$. Similarly, we can express inner products of the Laplace–Beltrami eigenfunctions ϕ_j as

$$(A.5) \quad \langle \phi_i, \phi_j \rangle_{L^2(\nu)} := \int_{\mathcal{M}} \phi_i \phi_j d\nu = \int_{\mathcal{M}} \phi_i \phi_j \frac{1}{\sigma} d\mu,$$

where $\frac{1}{\sigma} \in C^\infty(\mathcal{M})$ is the reciprocal of the sampling density σ (see [Assumption 3](#)). By convergence of P_α and P_α^* to the heat operator $e^{-c\epsilon^2\Delta}$ (implied by [Assumption 4](#)), it follows that w_α approximates $\frac{1}{\sigma}$ up to normalization. Given, moreover, that $\int_{\mathcal{M}} \frac{1}{\sigma} d\mu = \nu(\mathcal{M})$, we can ensure asymptotic consistency between the $L^2(\nu_\alpha)$ and $L^2(\nu)$ inner products by choosing the normalization constant Z_α such that

$$(A.6) \quad \int_{\mathcal{M}} w_\alpha d\mu_N = \mathcal{V}_\alpha,$$

for an estimate \mathcal{V}_α of the Riemannian volume $\nu(\mathcal{M})$ that converges in the large-data limit $\alpha \rightarrow (0^+, \infty)$.

To obtain such a volume estimate \mathcal{V}_α we use the Minakshisundaram–Pleijel formula for the heat trace [\[57\]](#),

$$\int_{\mathcal{M}} h_\tau(x, x) d\nu(x) \simeq (4\pi\tau)^{-\dim(\mathcal{M})/2} \nu(\mathcal{M}), \quad \tau \ll 1,$$

where $h_\tau: \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}_+$ is the time- τ heat kernel associated with the Riemannian metric \mathbf{g} (that is; $e^{-\tau\Delta} f = \int_{\mathcal{M}} h_\tau(\cdot, x) f(x) d\nu(x)$ for $\tau > 0$ and $f \in L^2(\nu)$). In our setting, P_α with $\alpha = (\epsilon, N)$ approximates the heat operator at time $\tau = c\epsilon^2 = \epsilon^2/4$, and thus $\int_{\mathcal{M}} p_\alpha(x, x) d\mu_N(x)$ approximates the corresponding heat trace $\int_X h_\tau(x, x) d\nu(x)$. This leads to the following asymptotic estimate for $\nu(\mathcal{M})$,

$$(A.7) \quad \nu(\mathcal{M}) \simeq \mathcal{V}_\alpha = (\pi\epsilon^2)^{\dim(\mathcal{M})/2} \int_{\mathcal{M}} p_\alpha(x, x) d\mu_N(x) = (\pi\epsilon^2)^{\dim(\mathcal{M})/2} \text{tr } \mathbf{P}_\alpha.$$

Note that the formula above requires knowledge of the manifold dimension $\dim(\mathcal{M})$, or a sufficiently accurate estimate thereof. In the examples of [section 6](#), $\dim(\mathcal{M})$ is assumed to be known. Numerical methods for estimating $\dim(\mathcal{M})$ can be found, e.g., in [\[49, 8, 17\]](#).

Having computed \mathcal{V}_α , we set the normalization coefficient Z_α and weight w_α so as to satisfy [\(A.6\)](#). We then normalize our diffusion eigenvectors such that they are orthonormal in $L^2(\nu_\alpha)$, $\langle \phi_i^{(\alpha)}, \phi_j^{(\alpha)} \rangle_{L^2(\nu_\alpha)} = \delta_{ij}$.

A.2. Pseudocode. This appendix includes pseudocode for computation of the diffusion maps eigenpairs (Algorithm A.1), out-of-sample evaluation of the diffusion eigenvectors (Algorithm A.2), and vector field regression problem in the SEC frame (Algorithm A.3). In the pseudocode, $\mathbf{u}^{\odot 2}$ and \mathbf{u}/\mathbf{v} denote elementwise squaring and division of vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$, respectively. Moreover, $\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{w}} = \mathbf{u}^\top (\text{diag } \mathbf{w}) \mathbf{v}$ is the weighted Euclidean inner product induced by a vector $\mathbf{w} \in \mathbb{R}^N$ with strictly positive elements, and $\|\mathbf{v}\|_{\mathbf{w}} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{w}}}$ the corresponding norm. The sampled points $\{y_n \in \mathbb{R}^d\}_{n=1}^N$ on the embedded manifold $F(\mathcal{M})$ and corresponding samples $\{\vec{v}_n\}_{n=1}^N$ of the pushforward vector field F_*V are as per the description in section 2.

Algorithm A.1 Eigenpairs $(\Lambda_j^{(\alpha)}, \phi_j^{(\alpha)})$ of the heat operator matrix \mathbf{P}_α from Appendix A.1.

- **Inputs:** Dataset $\{y_n \in \mathbb{R}^d\}_{n=1}^N$; kernel bandwidth $\epsilon > 0$; manifold dimension $m \in \mathbb{N}$; number of eigenpairs $M \in \mathbb{N}$.
- **Outputs:** Inner product weights $\mathbf{w}_\alpha \in \mathbb{R}^N$; top eigenvalues $\{\Lambda_j^{(\alpha)}\}_{j=0}^{M-1}$, sorted in descending order; corresponding eigenvectors $\{\phi_j^{(\alpha)}\}_{j=0}^{M-1}$, orthonormalized with respect to \mathbf{w}_α ; corresponding Laplace–Beltrami eigenvalues $\{\lambda_j^{(\alpha)}\}_{j=0}^{M-1}$.
- **Steps:** With k_ϵ from (A.1) as the kernel function and $\deg_r^{(\alpha)}, \deg_l^{(\alpha)}$ from (A.2) as the degree (normalization) functions:

1. Compute the values $r_n = \deg_r^{(\alpha)}(y_n)$ and $l_n = \deg_l^{(\alpha)}(y_n)$ of the degree functions for $n \in \{1, \dots, N\}$.
2. Form the $N \times N$ symmetric matrix $\tilde{\mathbf{P}}_\alpha = [\tilde{P}_{ij}^{(\alpha)}]_{i,j=1}^N$ with entries

$$\tilde{P}_{ij}^{(\alpha)} = \frac{k_\epsilon(y_i, y_j)/N}{\sqrt{l_i r_i r_j l_j}}.$$

3. Compute the largest M eigenvalues $\Lambda_0^{(\alpha)}, \dots, \Lambda_{M-1}^{(\alpha)} \in [0, 1]$ of $\tilde{\mathbf{P}}_\alpha$ and corresponding eigenvectors $\mathbf{e}_j \in \mathbb{R}^N$. Sort the $\Lambda_j^{(\alpha)}$ in descending order.
4. Compute the volume estimate \mathcal{V}_α using (A.7) with $\dim(\mathcal{M}) = m$.
5. Compute the weight vector $\mathbf{w}_\alpha = \mathcal{V}_\alpha \mathbf{e}_0^{\odot 2} / \|\mathbf{e}_0^{\odot 2}\|_1$.
6. For $j \in \{0, \dots, M-1\}$, compute the eigenvectors

$$\phi_j^{(\alpha)} = \mathbf{v}_j / \|\mathbf{v}_j\|_{\mathbf{w}_\alpha}, \quad \mathbf{v}_j = \mathbf{e}_j \cdot / \mathbf{e}_0.$$

7. For $j \in \{0, \dots, M-1\}$, compute the Laplace–Beltrami eigenvalues $\lambda_j^{(\alpha)} = -\frac{4 \log \Lambda_j^{(\alpha)}}{\epsilon^2}$.
 8. Return $\mathbf{w}_\alpha, \{\Lambda_j^{(\alpha)}\}_{j=0}^{M-1}, \{\phi_j^{(\alpha)}\}_{j=0}^{M-1}, \{\lambda_j^{(\alpha)}\}_{j=0}^{M-1}$.
-

Appendix B. Proof of Proposition 5.6. We shall make use of two auxiliary results, Lemmas B.1 and B.2. In what follows, $W_{-1}: [1/e, 0) \rightarrow \mathbb{R}_{<0}$ is the -1 branch of the Lambert W function (also known as product logarithm function), defined as the multivalued inverse of the function $w \mapsto we^w$; see e.g., [20].

Algorithm A.2 Out-of-sample extension of diffusion eigenfunctions.

• **Inputs:** Dataset $\{y_n \in \mathbb{R}^d\}_{n=1}^N$; kernel bandwidth $\epsilon > 0$; eigenpairs $\{\Lambda_j^{(\alpha)}, \phi_j^{(\alpha)}\}_{j=0}^{M-1}$ from Algorithm A.1.

• **Outputs:** Out-of-sample extensions $\{\varphi_j^{(\alpha)}: \mathbb{R}^d \rightarrow \mathbb{R}\}_{j=0}^{M-1}$ of the ϕ_j .

1. Build the vector-valued function $\rho: \mathbb{R}^d \rightarrow \mathbb{R}^N$, where

$$\rho(y) = \left(\frac{k_\epsilon(y, y_1)}{\deg_l^{(\alpha)}(y) \deg_r^{(\alpha)}(y_1)}, \dots, \frac{k_\epsilon(y, y_N)}{\deg_l^{(\alpha)}(y) \deg_r^{(\alpha)}(y_N)} \right).$$

2. For $j \in \{0, \dots, M-1\}$, build the functions $\varphi_j^{(\alpha)}: \mathbb{R}^d \rightarrow \mathbb{R}$, where

$$\varphi_j^{(\alpha)}(y) = \frac{1}{N \Lambda_j^{(\alpha)}} \rho(y) \cdot \phi_j^{(\alpha)}.$$

3. Return $\{\varphi_j^{(\alpha)}\}_{j=0}^{M-1}$.

Lemma B.1 (“Running-away inequality”). *Let E be a separable Banach space, $T > 0$, and $f: [0, T) \rightarrow E$ differentiable at every $t \in [0, T)$. Then $t \mapsto \|f(t)\|_E$ is differentiable for every $t \in [0, T)$ for which $f(t) \neq 0$, and it is right-differentiable everywhere. Moreover, the bound*

$$\frac{d}{dt} \|f(t)\|_E \leq \left\| \frac{df(t)}{dt} \right\|_E$$

holds for every $t \in [0, T)$ with right-hand derivatives used when necessary.

Proof. See [58, Lemma 3.3.1] for a proof when E is finite-dimensional. The proof in the infinite-dimensional case proceeds similarly, using a version of Taylor’s theorem for separable infinite-dimensional Banach spaces. ■

Lemma B.2. *For every $\alpha, \beta > 0$, the solution of the initial-value problem*

$$\dot{h}(t) = \alpha + \beta \sqrt{h(t)}, \quad h(0) = 0,$$

is

$$h(t) = \left(\frac{\alpha}{\beta} (W_{-1}(\zeta(t)) + 1) \right)^2, \quad \zeta(t) = -\exp\left(-\frac{\beta^2 t}{\alpha} - 1\right).$$

Proof. The W_{-1} function on $[-1/e, 0)$ can be equivalently defined as the solution of the initial-value problem [20]

$$\dot{W}_{-1}(z) = \frac{W_{-1}(z)}{z(1 + W_{-1}(z))}, \quad W_{-1}(-1/e) = -1.$$

The claim of the lemma follows by differentiating the given expression for $h(t)$ and using this result. ■

Algorithm A.3 Regression problem for the dynamical vector field V in the SEC frame.

• **Inputs:** Manifold samples $\{y_n\}_{n=1}^N$; vector field samples $\{\vec{v}_n \in \mathbb{R}^d\}_{n=1}^N$; number of gradient fields $J \in \mathbb{N}$ in the SEC frame; spectral resolution parameters $L, L_D, L_1, L_2 \in \mathbb{N}$; spectral truncation parameter $\eta > 0$; eigenpairs $\{(\lambda_j^{(\alpha)}, \phi_j^{(\alpha)})\}_{j=0}^{M-1}$ and inner product weights \mathbf{w}_α from [Algorithm A.1](#) with $M = \max\{J, L, L_D, L_1, L_2\}$.

• **Outputs:** Frame expansion coefficients $\{b_{ij}^{(\ell, \alpha)} \in \mathbb{R}\}_{i,j=0,1}^{L-1,J}$ for the vector field $\vec{V}^{(\ell, \alpha)}$.

• **Steps:**

1. Compute c -coefficients associated with the eigenvectors $\phi_j^{(\alpha)}$,

$$c_{ijp}^{(\alpha)} = \langle \phi_p^{(\alpha)}, \phi_i^{(\alpha)} \phi_j^{(\alpha)} \rangle_{\mathbf{w}_\alpha}, \quad i, j, p \in \{1, \dots, M\}.$$

2. Using the $c_{ijp}^{(\alpha)}$ coefficients and the eigenvalues $\lambda_j^{(\alpha)}$, compute $d_{ijkl}^{(L_D, \alpha)}$ for $i \in \{0, \dots, L-1\}$, $j \in \{1, \dots, J\}$, $k \in \{0, \dots, L_1-1\}$, and $l \in \{0, \dots, L_2-1\}$ via the formula (4.7). Reshape the resulting tensor $[d_{ijkl}^{(L_D, \alpha)}]$ into an $(LJ) \times (L_1L_2)$ matrix \mathbf{D} by grouping together the (i, j) and (k, l) indices.
3. Form the $d \times N$ data matrix $\mathbf{Y} = (y_1 \cdots y_N)$ with columns y_n and compute the leading L_1 expansion coefficients $F_k^{(\alpha)} \in \mathbb{R}^d$ of the embedding,

$$F_k^{(\alpha)} = \langle F, \phi_k^{(\alpha)} \rangle_{L^2(\nu_\alpha)} = \mathbf{Y}(\text{diag } \mathbf{w}_\alpha) \phi_k, \quad k \in \{0, \dots, L_1-1\}.$$

4. Form the $d \times N$ vector field matrix $\mathbf{V} = (\vec{v}_1 \cdots \vec{v}_N)$ with columns \vec{v}_n and compute the coefficients

$$\hat{v}_{kl} = F_k^{(\alpha)} \cdot \mathbf{V}(\text{diag } \mathbf{w}_\alpha) \phi_l, \quad k \in \{0, \dots, L_1-1\}, \quad l \in \{0, \dots, L_2-1\}.$$

Reshape the resulting tensor $[\hat{v}_{kl}]$ into an (L_1L_2) -dimensional vector $\hat{\mathbf{v}}$.

5. Using the $c_{ijp}^{(\alpha)}$ coefficients and the eigenvalues $\lambda_j^{(\alpha)}$, compute the matrix elements $d_{ijlk}^{(L_D, \alpha)}$ of the Gram operator $G^{(L_D, \alpha)}$ in (4.8) for $i, k \in \{0, \dots, L_1\}$ and $j, l \in \{1, \dots, J\}$. Reshape the resulting tensor $[d_{ijlk}^{(L_D, \alpha)}]$ into an $(LJ) \times (LJ)$ matrix \mathbf{G} by grouping together the (i, j) and (l, k) indices.
 6. Compute an eigendecomposition $\mathbf{G}\mathbf{u}_j = \gamma_j \mathbf{u}_j$ of \mathbf{G} , where the eigenvectors \mathbf{e}_j are orthonormal in \mathbb{R}^{LJ} .
 7. Form the spectrally truncated Gram matrix $\mathbf{G}_\eta = \sum_{j: \gamma_j > \eta} \gamma_j \mathbf{u}_j \mathbf{u}_j^\top$.
 8. Compute the solution $\mathbf{b} \in \mathbb{R}^{LJ}$ of the regression problem (4.10), $\mathbf{b} = \mathbf{G}_\eta^+ \mathbf{D} \hat{\mathbf{v}}$.
 9. Return the coefficients $\{b_{ij}^{(\ell, \alpha)}\}_{i,j=0,1}^{L-1,J}$ from the elements of this vector.
-

Returning to the proof of the proposition, we use [Lemma B.1](#) to get

$$\begin{aligned} \dot{\varepsilon}(t) &\leq \left\| \frac{d}{dt} y(t, \cdot) - \frac{d}{dt} \hat{y}(t, \cdot) \right\|_{L^1(\nu; \mathbb{R}^d)} = \left\| V(F \circ \Phi^t) - \hat{V}(F \circ \hat{\Phi}^t) \right\|_{L^1(\nu; \mathbb{R}^d)} \\ &\leq \left\| V(F \circ \Phi^t) - \hat{V}(F \circ \Phi^t) \right\|_{L^1(\nu; \mathbb{R}^d)} + \left\| \hat{V}(F \circ \Phi^t) - \hat{V}(F \circ \hat{\Phi}^t) \right\|_{L^1(\nu; \mathbb{R}^d)}. \end{aligned}$$

Writing $F = (F_1, \dots, F_d)$ where $F_j \in C^\infty(\mathcal{M})$ are the component functions of the embedding, we can estimate the first term in the last line as

$$\begin{aligned}
\|V(F \circ \Phi^t) - \hat{V}(F \circ \Phi^t)\|_{L^1(\nu; \mathbb{R}^d)} &= \int_{\mathcal{M}} \|(V - \hat{V})(F \circ \Phi^t)(x)\|_1 d\nu(x) \\
&= \int_{\mathcal{M}} \sum_{j=1}^d |(V - \hat{V})(F_j \circ \Phi^t)(x)| d\nu(x) \\
&= \int_{\mathcal{M}} \sum_{j=1}^d |(V - \hat{V}) \cdot \nabla(F_j \circ \Phi^t)(x)| d\nu(x) \\
&\leq \int_{\mathcal{M}} \sum_{j=1}^d \|V - \hat{V}\|_{\mathfrak{g}(x)} \|\nabla(F_j \circ \Phi^t)\|_{\mathfrak{g}(x)} d\nu(x) \\
&\leq \|V - \hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \sqrt{\int_{\mathcal{M}} \|\nabla(F_j \circ \Phi^t)\|_{\mathfrak{g}(x)}^2 d\nu(x)} \\
&= \|V - \hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \left\| \|\nabla(F_j \circ \Phi^t)\|_{\mathfrak{g}} \right\|_{L^2(\nu)} \\
&\leq \|V - \hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \|\nabla(F_j \circ \Phi^t)\|_{\mathbb{H}} \\
&= \|V - \hat{V}\|_{\mathbb{H}} |F \circ \Phi^t|_{1,2} \leq \alpha.
\end{aligned}$$

For the second term, letting $\Xi_j = F_j \circ \Phi^t - F_j \circ \hat{\Phi}^t$ and proceeding as above, we get

$$\begin{aligned}
\|\hat{V}(F \circ \Phi^t) - \hat{V}(F \circ \hat{\Phi}^t)\|_{L^1(\nu; \mathbb{R}^d)} &\leq \|\hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \left\| \|\nabla(F_j \circ \Phi^t - F_j \circ \hat{\Phi}^t)\|_{\mathfrak{g}} \right\|_{L^2(\nu)} \\
&\leq \|\hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \sqrt{\int_{\mathcal{M}} \nabla \Xi_j \cdot \nabla \Xi_j d\nu} \\
&\leq \|\hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \sqrt{-\int_{\mathcal{M}} (\Delta \Xi_j) \Xi_j d\nu} \\
&\leq \|\hat{V}\|_{\mathbb{H}} \sum_{j=1}^d \sqrt{\|\Delta \Xi_j\|_{L^\infty(\nu)} \|\Xi_j\|_{L^1(\nu)}} \\
&\leq \|\hat{V}\|_{\mathbb{H}} \left(\sum_{i=1}^d \|\Delta \Xi_i\|_{L^\infty(\nu)} \right)^{1/2} \left(\sum_{j=1}^d \|\Xi_j\|_{L^1(\nu)} \right)^{1/2} \\
&= \|\hat{V}\|_{\mathbb{H}} |F \circ \Phi^t - F \circ \hat{\Phi}^t|_{2,\infty} \sqrt{\varepsilon(t)}.
\end{aligned}$$

Combining these estimates, we deduce that

$$(B.1) \quad \dot{\varepsilon}(t) \leq \alpha + \beta\sqrt{\varepsilon(t)}.$$

Next, since $\varepsilon(0) = 0$ by definition, we can use Lemma B.2 to integrate (B.2), giving

$$\varepsilon(t) \leq \left(\frac{\alpha}{\beta} (W_{-1}(\zeta(t)) + 1) \right)^2, \quad \zeta(t) = -\exp\left(-\frac{\beta^2 t}{\alpha} - 1\right).$$

Moreover, from [14, Theorem 1] we have,

$$-1 - \sqrt{2\tau} - \tau < W_{-1}(-\exp(-(\tau + 1))) < -1 - \sqrt{2\tau} - \frac{2\tau}{3}, \quad \forall \tau > 0.$$

With $\tau = \beta^2 t / (2\alpha)$, the lower bound gives

$$-W_{-1}(\zeta(t)) - 1 \leq \sqrt{\frac{\beta^2 t}{\alpha}} + \frac{\beta^2 t}{2\alpha}.$$

As a result, we obtain

$$(B.2) \quad \sqrt{\varepsilon(t)} \leq \sqrt{\alpha t} + \frac{\beta t}{2}.$$

The claim of the proposition follows from (B.1) and (B.2). ■

REFERENCES

- [1] R. ALEXANDER AND D. GIANNAKIS, *Operator-theoretic framework for forecasting nonlinear time series with kernel analog techniques*, Phys. D, 409 (2020), 132520, <https://doi.org/10.1016/j.physd.2020.132520>.
- [2] D. N. ARNOLD, R. S. FALK, AND R. WINTHER, *Finite element exterior calculus, homological techniques, and applications*, Acta Numerica, **15** (2006), pp. 1–155, <https://doi.org/10.1017/S0962492906210018>.
- [3] N. BAILEY, *The mathematical theory of infectious diseases and its applications*, Charles Griffin & Co., 1975.
- [4] D. BAKRY, I. GENTIL, AND M. LEDOUX, *Analysis and Geometry of Markov Diffusion Operators*, Springer, Cham, 2014, <https://doi.org/10.1007/978-3-319-00227-9>.
- [5] T. BERRY AND S. DAS, *Learning theory for dynamical systems*, SIAM J. Appl. Dyn., 22 (2023), pp. 2082 – 2122, <https://doi.org/10.1137/22M1516865>.
- [6] T. BERRY AND S. DAS, *Limits of learning dynamical systems*, SIAM review, 16 (2025), <https://doi.org/10.1137/24M1696974>, <https://epubs.siam.org/doi/abs/10.1137/24M1696974?journalCode=siread>.
- [7] T. BERRY AND D. GIANNAKIS, *Spectral exterior calculus*, Commun. Pure Appl. Math., 73 (2019), pp. 689–770, <https://doi.org/10.1002/cpa.21885>.
- [8] T. BERRY, D. GIANNAKIS, AND J. HARLIM, *Nonparametric forecasting of low-dimensional dynamical systems*, Phys. Rev. E., 91 (2015), 032915, <https://doi.org/10.1103/PhysRevE.91.032915>.
- [9] K. BHATTACHARYA, B. HOSSEINI, N. B. KOVACHKI, AND A. M. STUART, *Model reduction and neural networks for parametric pdes*, SIAM J. Comput. Math., 7 (2021), pp. 121–156, <https://doi.org/10.5802/smai-jcm.74>.
- [10] F. BRAUER AND C. CASTILLO-CHÁVEZ, *Mathematical models in population biology and epidemiology*, vol. 2, Springer, 2012.
- [11] S. L. BRUNTON, B. W. BRUNTON, J. L. PROCTOR, E. KAISER, AND J. N. KUTZ, *Chaos as an intermittently forced linear system*, Nat. Commun., 8 (2017), 19, <https://doi.org/10.1038/s41467-017-00030-8>.

- [12] S. L. BRUNTON, J. L. PROCTOR, AND J. N. KUTZ, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, Proc. Natl. Acad. Sci., 113 (2016), pp. 3932–3937, <https://doi.org/10.1073/pnas.1517384113>.
- [13] J. CALDER AND N. TRILLOS, *Improved spectral convergence for graph Laplacians on ϵ -graphs and k -NN graphs*, Appl. Comput. Harmon. Anal., 60 (2022), pp. 123–175, <https://doi.org/10.1016/j.acha.2022.02.004>.
- [14] I. CHATZIGEORGIOU, *Bounds on the Lambert function and their application to the outage analysis of user cooperation*, IEEE Commun. Lett., 17 (2013), pp. 1505–1508, <https://doi.org/10.1109/LCOMM.2013.070113.130972>.
- [15] M. D. CHEKROUN, J. D. NEELIN, D. KONDRASHOV, J. C. MCWILLIAMS, AND M. GHIL, *Rough parameter dependence in climate models and the role of Ruelle-Pollicott resonances*, Proc. Natl. Acad. Sci., 111 (2013), pp. 1684–1690, <https://doi.org/10.1073/pnas.1321816111>.
- [16] R. R. COIFMAN AND S. LAFON, *Diffusion maps*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 5–30, <https://doi.org/10.1016/j.acha.2006.04.006>.
- [17] R. R. COIFMAN, Y. SHKOLNISKY, F. J. SIGWORTH, AND A. SINGER, *Graph Laplacian tomography from unknown random projections*, IEEE Trans. Image Process., 17 (2008), pp. 1891–1899, <https://doi.org/10.1109/tip.2008.2002305>.
- [18] M. J. COLBROOK, I. MEZIĆ, AND A. STEPANENKO, *Limits and powers of Koopman learning*, 2024, <https://arxiv.org/abs/2407.06312>.
- [19] M. J. COLBROOK AND A. TOWNSEND, *Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems*, Commun. Pure Appl. Math., 77 (2024), pp. 221–283, <https://doi.org/10.1002/cpa.22125>.
- [20] R. M. CORLESS, G. H. GONNET, D. E. G. HARE, D. J. JEFFREY, AND D. E. KNUTH, *On the Lambert W function*, Adv. Comput. Math., 5 (1996), pp. 329–359, <https://doi.org/10.1007/BF02124750>.
- [21] S. DAS, *Reconstructing dynamical systems as zero-noise limits*, 2024, <https://arxiv.org/abs/math/2407.16673>.
- [22] S. DAS, D. GIANNAKIS, AND J. SLAWINSKA, *Reproducing kernel Hilbert space compactification of unitary evolution groups*, Appl. Comput. Harmon. Anal., 54 (2021), pp. 75–136, <https://doi.org/10.1016/j.acha.2021.02.004>.
- [23] S. DAS AND J. YORKE, *Super convergence of ergodic averages for quasiperiodic orbits*, Nonlinearity, 31 (2018), p. 391, <https://doi.org/10.1088/1361-6544/aa99a0>.
- [24] M. DELLNITZ AND O. JUNGE, *On the approximation of complicated dynamical behavior*, SIAM J. Numer. Anal., 36 (1999), p. 491, <https://doi.org/10.1137/S0036142996313002>.
- [25] M. DESBRUN, E. KANSO, AND Y. TONG, *Discrete differential forms for computational modeling*, vol. 38, Oberwolfach Seminars, Birkhäuser, Basel, 2008, pp. 287–324, https://doi.org/10.1007/978-3-7643-8621-4_16.
- [26] E. N. EPPERLY, J. A. TROPP, AND R. J. WEBBER, *Embrace rejection: Kernel matrix approximation by accelerated randomly pivoted Cholesky*, 2025, <https://arxiv.org/abs/2410.03969>.
- [27] L. FINESSO, A. GRASSI, AND P. SPREIJ, *Approximation of stationary processes by hidden Markov models*, Mathematics of control, signals, and systems, 22 (2010), pp. 1–22, <https://doi.org/10.1007/s00498-010-0050-7>.
- [28] G. FROYLAND AND P. KOLTAI, *Estimating long-term behavior of periodically driven flows without trajectory integration*, Nonlinearity, 30 (2017), p. 1948–1986, <https://doi.org/10.1088/1361-6544/aa6693>.
- [29] D. GIANNAKIS, A. HENRIKSEN, J. A. TROPP, AND R. WARD, *Learning to forecast dynamical systems from streaming data*, SIAM J. Appl. Dyn. Syst., 22 (2023), pp. 527–558, <https://doi.org/10.1137/21M144983X>.
- [30] L. GONON AND J. ORTEGA, *Fading memory echo state networks are universal*, Neural Networks, 138 (2021), pp. 10–13, <https://doi.org/10.1016/j.neunet.2021.01.025>.
- [31] R. GONZÁLEZ-GARCIA, R. RICO-MARTINEZ, AND I. G. KEVREKIDIS, *Identification of distributed parameter systems: A neural net based approach*, Comput. Chem. Eng., 22 (1998), pp. S965–S968, [https://doi.org/10.1016/S0098-1354\(98\)00191-4](https://doi.org/10.1016/S0098-1354(98)00191-4).
- [32] L. GRIGORYEVA, A. HART, AND J. ORTEGA, *Chaos on compact manifolds: Differentiable synchronizations beyond the Takens theorem*, Phys. Rev. E, 103 (2021), p. 062204, <https://doi.org/10.1103/PhysRevE.103.062204>.

- [33] P. HALL, J. REIMANN, AND J. RICE, *Nonparametric estimation of a periodic function*, Biometrika, 87 (2000), pp. 545–557, <https://doi.org/10.1093/biomet/87.3.545>.
- [34] B. HAMZI AND H. OWHADI, *Learning dynamical systems from data: A simple cross-validation perspective, part I: Parametric kernel flows*, Phys. D, 421 (2021), p. 132817, <https://doi.org/10.1016/j.physd.2020.132817>.
- [35] J. HARLIM, S. JIANG, S. LIANG, AND H. YANG, *Machine learning for prediction with missing dynamics*, J. Comput. Phys., 428 (2021), p. 109922, <https://doi.org/10.1016/j.jcp.2020.109922>.
- [36] J. HARLIM, S. W. JIANG, AND J. W. PEOPLES, *Radial basis function approximation of tensor fields on manifolds: From operator approximation to manifold learning*, J. Mach. Learn. Res., 2024 (2024), pp. 16521–16605, <https://dl.acm.org/doi/10.5555/3648699.3649044>.
- [37] M. HEIN, J.-Y. AUDIBERT, AND U. VON LUXBURG, *Graph Laplacians and their convergence on random neighborhood graphs*, J. Mach. Learn. Res., 8 (2007), pp. 1325–1368.
- [38] A. N. HIRANI, *Discrete exterior calculus*, PhD thesis, California Institute of Technology, 2003, https://thesis.library.caltech.edu/1885/3/thesis_hirani.pdf.
- [39] K. HORNIK, M. STINCHCOMBE, AND H. WHITE, *Multilayer feedforward networks are universal approximators*, Neural Netw., 2 (1989), pp. 359–366, [https://doi.org/https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/https://doi.org/10.1016/0893-6080(89)90020-8).
- [40] H. JAEGER AND H. HAAS, *Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication*, Science, 304 (2004), pp. 78–80, <https://doi.org/10.1126/science.1091277>.
- [41] Y. KAWAHARA, *Dynamic mode decomposition with reproducing kernels for Koopman spectral analysis*, in Advances in Neural Information Processing Systems, D. D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon, and R. Garnett, eds., Curran Associates, 2016, pp. 911–919, <http://papers.nips.cc/paper/6583-dynamic-mode-decomposition-with-reproducing-kernels-for-koopman-spectral-analysis.pdf>.
- [42] S. KLUS, I. SCHUSTER, AND K. MUANDET, *Eigendecomposition of transfer operators in reproducing kernel Hilbert spaces*, J. Nonlinear Sci., 30 (2020), pp. 283–315, <https://doi.org/10.1007/s00332-019-09574-z>.
- [43] P. KOLTAI AND P. KUNDE, *A Koopman–Takens theorem: Linear least squares prediction of nonlinear time series*, 2023, <https://arxiv.org/abs/2308.02175>.
- [44] V. R. KOSTIC, P. NOVELLI, A. MAUER, C. CILIBERTO, L. ROSASCO, AND M. PONTIL, *Learning dynamical systems with Koopman operator regression in reproducing kernel Hilbert spaces*, in Advances in Neural Information Processing Systems 35 (NeurIPS 2022), S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, eds., 2022, pp. 4017–4031.
- [45] N. KOVACHKI, Z. LI, B. LIU, K. AZIZZADENESHELI, K. BHATTACHARYA, A. STUART, AND A. ANANDKUMAR, *Neural operator: Learning maps between function spaces with applications to PDEs*, J. Mach. Learn. Res., 24 (2023), pp. 4061–4157, <https://doi.org/https://dl.acm.org/doi/10.5555/3648699.3648788>.
- [46] H. LANGE, S. L. BRUNTON, AND K. J. N, *From Fourier to Koopman: Spectral methods for long-term time series prediction.*, J. Mach. Learn. Res., 22 (2021), pp. 41–1.
- [47] C. LESSIG, *Ψ_{ec} : A local spectral exterior calculus*, Appl. Comput. Harmon. Anal., 51 (2021), pp. 56–103, <https://doi.org/10.1016/j.acha.2020.10.003>.
- [48] Y. LIN AND L. BROWN, *Statistical properties of the method of regularization with periodic gaussian reproducing kernel*, Ann. Stat., 32 (2004), pp. 1723–1743, <https://doi.org/10.1214/009053604000000454>.
- [49] A. V. LITTLE, J. LEE, Y.-M. JUNG, AND M. MAGGIONI, *Estimation of intrinsic dimensionality of samples from noisy low-dimensional manifolds in high dimensions with multiscale SVD*, in Proceedings of the 15th IEEE/SPWorkshop on Statistical Signal Processing, 2009, pp. 85–88, <https://doi.org/10.1109/SSP.2009.5278634>.
- [50] A. W. LONG AND A. L. FERGUSON, *Landmark diffusion maps (L-dMaps): Accelerated manifold learning out-of-sample extension*, Appl. Comput. Harmon. Anal., 47 (2019), pp. 190–211, <https://doi.org/10.1016/j.acha.2017.08.004>.
- [51] C. MA, J. WANG, AND W. E, *Model reduction with memory and the machine learning of dynamical systems*, Commun. Comput. Phys., 25 (2018), pp. 947–962, <https://doi.org/10.4208/CICP.OA-2018-0269>.
- [52] A. J. MAJDA AND X. WANG, *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*, Cambridge University Press, Cambridge, 2006.
- [53] S. MALLAT, *A Wavelet Tour of Signal Processing*, Academic Press, Burlington, 3 ed., 2009.

- [54] R. MAULIK, A. MOHAN, B. LUSCH, S. MADIREDDY, P. BALAPRAKASH, AND D. LIVESCU, *Time-series learning of latent-space dynamics for reduced-order model closure*, Phys. D, 405 (2020), p. 132368, <https://doi.org/10.1016/j.physd.2020.132368>.
- [55] I. MEZIĆ, *Spectral properties of dynamical systems, model reduction and decompositions*, Nonlinear Dyn., 41 (2005), pp. 309–325, <https://doi.org/10.1007/s11071-005-2824-x>.
- [56] C. A. MICCHELLI, Y. XU, AND H. ZHANG, *Universal kernels*, J. Mach. Learn. Res., 7 (2006), pp. 2651–2667.
- [57] S. MINAKSHISUNDARAM AND Å. PLEIJEL, *Some properties of the eigenfunctions of the Laplace-operator on Riemannian manifolds*, Can. J. Math., 1 (1949), pp. 242–256, <https://doi.org/10.4153/CJM-1949-021-5>.
- [58] J. A. MURDOCK, *Perturbations: Theory and Methods*, vol. 27 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1999.
- [59] J. C. OXTOPY, *Stepanoff flows on the torus*, Proc. Amer. Math. Soc., 4 (1953), pp. 982–987.
- [60] T. N. PALMER AND R. HAGEDORN, eds., *Predictability of Weather and Climate*, Cambridge University Press, Cambridge, 2006.
- [61] J. PARK, N. T. YANG, AND N. CHANDRAMOORTHY, *When are dynamical systems learned from time series data statistically accurate?*, in Advances in Neural Information Processing Systems, A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, eds., vol. 37, Curran Associates, Inc., 2024, pp. 43975–44008, https://proceedings.neurips.cc/paper_files/paper/2024/file/4dc57702c987e1e72f0dd2921edb5ded-Paper-Conference.pdf.
- [62] C. PENLAND, *Random forcing and forecasting using principal oscillation pattern analysis*, Mon. Weather Rev., 117 (1989), pp. 2165–2185, [https://doi.org/10.1175/1520-0493\(1989\)117<2165:RFAFUP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<2165:RFAFUP>2.0.CO;2).
- [63] J. W. PEOPLES AND J. HARLIM, *Spectral convergence of symmetrized graph Laplacian on manifolds with boundary*, 2023, <https://arxiv.org/abs/2110.06988>.
- [64] E. QIAN, B. KRAMER, B. PEHERSTORFER, AND K. WILLCOX, *Lift & Learn: Physics-informed machine learning for large-scale nonlinear dynamical systems*, Phys. D, 406 (2020), 132401, <https://doi.org/10.1016/j.physd.2020.132401>.
- [65] M. RAISSI, P. PERDIKARIS, AND G. E. KARNIADAKIS, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, J. Comput. Phys., 378 (2019), pp. 686–707, <https://doi.org/10.1016/j.jcp.2018.10.045>.
- [66] V. RAKOČEVIĆ, *On continuity of the Moore–Penrose and Drazin inverses*, Mat. vesn., 49 (1997), pp. 163–172, <http://dml.mathdoc.fr/item/01677966/>.
- [67] C. W. ROWLEY, I. MEZIĆ, S. BAGHERI, P. SCHLATTER, AND D. S. HENNINGSON, *Spectral analysis of nonlinear flows*, J. Fluid Mech., 641 (2009), pp. 115–127, <https://doi.org/10.1017/S0022112009992059>.
- [68] C. SCHÜTTE, W. HUISINGA, AND P. DEUFLHARD, *Transfer operator approach to conformational dynamics in biomolecular systems*, in Ergodic Theory, Analysis, and Efficient Simulation of Dynamical Systems, B. Fiedler, ed., Springer-Verlag, Berlin, 2001, pp. 191–223, https://doi.org/10.1007/978-3-642-56589-2_9.
- [69] B. SILVERMAN, *Spline smoothing: the equivalent variable kernel method*, Ann. Stat., (1984), pp. 898–916, <https://doi.org/10.1214/aos/1176346710>.
- [70] A. M. STUART, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559, <https://doi.org/10.1017/S0962492910000061>.
- [71] W. TAO AND Z. SHI, *Convergence of Laplacian spectra from random samples*, J. Comput. Math., 38 (2020), pp. 952–984, <https://doi.org/10.4208/jcm.2008-m2018-0232>.
- [72] A. TOMPKINS AND F. RAMOS, *Periodic kernel approximation by index set Fourier series features*, in Uncertainty in Artificial Intelligence, PMLR, 2020, pp. 486–496.
- [73] N. TRILLOS, M. GERLACH, M. HEIN, AND D. SLEPČEV, *Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace–Beltrami operator*, Found. Comput. Math., 20 (2020), pp. 827–887, <https://doi.org/10.1007/s10208-019-09436-w>.
- [74] M. VIDYASAGAR, *The realization problem for hidden Markov models: The complete realization problem*, in Proceedings of the 44th IEEE Conference on Decision and Control, IEEE, 2005, pp. 6632–6637, <https://doi.org/10.1109/CDC.2005.1583227>.
- [75] M. O. WILLIAMS, I. G. KEVREKIDIS, AND C. W. ROWLEY, *A data-driven approximation of the Koopman*

- operator: Extending dynamic mode decomposition*, J. Nonlinear Sci., 25 (2015), pp. 1307–1346, <https://doi.org/10.1007/s00332-015-9258-5>.
- [76] C. L. WORMELL AND S. REICH, *Spectral convergence of diffusion maps: Improved error bounds and an alternative normalisation*, SIAM J. Numer. Anal., 59 (2021), pp. 1687–1734, <https://doi.org/10.1137/20M1344093>.
- [77] E. YEUNG, S. KUNDU, AND N. HODAS, *Learning deep neural network representations for koopman operators of nonlinear dynamical systems*, in 2019 American Control Conference (ACC), IEEE, 2019, pp. 4832–4839, <https://doi.org/10.23919/ACC.2019.8815339>.
- [78] Y. ZHU AND N. ZABARAS, *Bayesian deep convolutional encoder–decoder networks for surrogate modeling and uncertainty quantification*, J. Comput. Phys., 366 (2018), pp. 415–447, <https://doi.org/10.1016/j.jcp.2018.04.018>.