ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/369832736

Double Diffusion Maps and their Latent Harmonics for Scientific Computations in Latent Space

Article *in* Journal of Computational Physics · April 2023

DOI: 10.1016/j.jcp.2023.112072

CITATIONS
9

READS
41

6 authors, including:



Nikolaos Evangelou
Johns Hopkins University
13 PUBLICATIONS **30** CITATIONS
SEE PROFILE



Felix Dietrich
Technische Universität München
66 PUBLICATIONS **992** CITATIONS
SEE PROFILE



Eliodoro Chiavazzo
Politecnico di Torino
131 PUBLICATIONS **2,611** CITATIONS
SEE PROFILE



Daniel Lehmberg
Technical University of Munich
11 PUBLICATIONS **65** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

iMat: Machine Learning for Materials Data extraction View project

Modified cost effective fibre based structures with improved multi-functionality and performance (MODCOMP) View project

All content following this page was uploaded by Eliodoro Chiavazzo on 18 April 2023.

The user has requested enhancement of the downloaded file.

arXiv:2204.12536v1 [stat.ML] 26 Apr 2022

DOUBLE DIFFUSION MAPS AND THEIR LATENT HARMONICS FOR SCIENTIFIC COMPUTATIONS IN LATENT SPACE

A PREPRINT

Nikolaos Evangelou
Johns Hopkins University
Baltimore, MD 21218, USA

Felix Dietrich
Technical University of Munich
Munich, 80333, Germany

Eliodoro Chiavazzo
Polytechnic University of Turin
Turin, 10129, Italy

Daniel Lehmberg
Technical University of Munich
Munich, 80335, Germany

Marina Meila
University of Washington
Seattle, WA 98195, USA

Ioannis G. Kevrekidis*
Johns Hopkins University
Baltimore, MD 21218, USA
yannisk@jhu.edu

ABSTRACT

We introduce a data-driven approach to building reduced dynamical models through manifold learning; the reduced latent space is discovered using Diffusion Maps (a manifold learning technique) on time series data. *A second round of Diffusion Maps on those latent coordinates* allows the approximation of the reduced dynamical models. This second round enables mapping the latent space coordinates back to the full ambient space (what is called *lifting*); it also enables the approximation of full state functions of interest in terms of the reduced coordinates. In our work, we develop and test three different reduced numerical simulation methodologies, either through pre-tabulation in the latent space and integration on the fly or by going *back and forth* between the ambient space and the latent space. The data-driven latent space simulation results, based on the three different approaches, are validated through (a) the latent space observation of the full simulation through the Nyström Extension formula, or through (b) lifting the reduced trajectory back to the full ambient space, via *Latent Harmonics*. Latent space modeling often involves additional regularization to favor certain properties of the space over others, and the mapping back to the ambient space is then constructed mostly independently from these properties; here, we use *the same* data-driven approach to construct the latent space and then map back to the ambient space.

1 Introduction

Differential equations (either ordinary or partial) are ubiquitous in science and engineering. They are often used to model spatiotemporal variations of quantities such as temperature, mass, energy, momentum, electric charge, stresses and strains. Being able to efficiently (accurately and parsimoniously) solve differential equations underpins modeling in physics and engineering [40].

Numerically solving a time-dependent partial differential equation (PDE) relies on approximation through discretization techniques, such as finite differences, finite elements or spectral methods, which leads to finite dimensional dynamical systems [23]. This discretization "transforms" the PDE to a system of Ordinary Differential Equations (ODEs). The more detailed tracking of the spatiotemporal evolution is required, the finer the spatial discretization needs to be and the larger the number of discretization ODEs that must be integrated [23].

Many PDEs arising in physicochemical modeling, however, exhibit long-time behavior that is relatively low-dimensional. From a mathematical point of view the long-term dynamics for those systems concentrate in an *absorbing ball*, $\mathbb{B}_0$ and ultimately lie on a low-dimensional manifold [23, 14]. For a class of such dissipative PDEs, including the Kuramoto-Sivashinsky equation (KSE) [10], the Ginzburg Landau [12, 20], Swift-Hohenberg [41] and several reaction-diffusion equations [10, 21, 39, 18, 22] an object called *inertial manifold* exists [23]. Inertial manifolds are exponentially attracting, "slow" invariant manifolds that capture the overall long-term behavior of the system

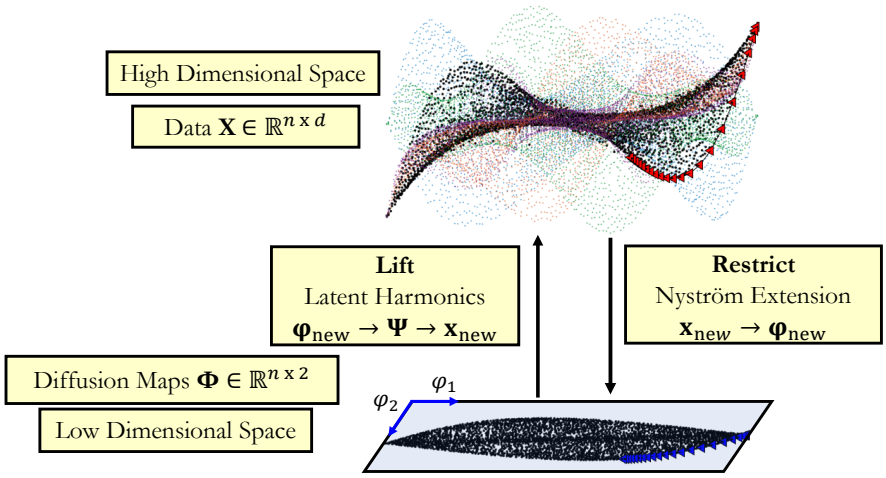


Figure 1: An illustration of our computational tools (Diffusion Maps, Latent Harmonics and Nyström Extension) for building reduced dynamical models.

[39]. Furthermore, restricting the dynamics to these slow inertial manifolds reduces dimensionality and mitigates stiffness of the initial, detailed discretized ODE system. This can significantly accelerate scientific computations.

Large stiff systems of ODEs arise also in other modeling contexts: control theory, where controlling forces may be turned in and off again suddenly, or circuit components having widely differing natural frequencies [43]. Another field in which stiff ODEs appear is chemical kinetics, in which the ordinary differential equations describe reactions of various chemical species to form other species. Here the stiffness is a consequence of the fact that different reactions take place on vastly different time scales (time scales ratios of 10^6 or more are common). Due to the presence of fast and slow dynamics, the trajectories are initially attracted towards a slow manifold, and subsequently proceed to steady state, or to thermodynamic equilibrium, always remaining close to the manifold. The presence of such an attracting slow manifold offers the possibility of constructing a consistent reduced description of the process, which accurately retains the slow dynamics along the manifold while neglecting the initial short transient approaching it in transverse directions [6, 43].

Even when the existence of an *inertial* manifold can be established, the theory usually does not provide of an explicit closed formula for it, and only approximations of the inertial manifold can be attempted [23]. Data Mining can be employed to discover such low dimensional inertial manifolds, given a collection of points sampled on it. Furthermore, machine learning algorithms can be employed to *learn* the dynamics on the manifold without the need of closed form expressions.

This paper is organized as follows: In Section 2 we describe the Double Diffusion Maps scheme. In Section 3 the proposed algorithms for the construction of reduced data-driven dynamical models are introduced. In Sections 4.1 and 4.2 we introduce our two examples: the Chafee-Infante PDE and a kinetic system comprised of stiff ODEs that models the combustion of hydrogen. In those Sections we discuss how sampling for these dynamical systems was performed and show the ability of Diffusion Maps to extract a useful set of latent coordinates, Algorithm 1 and a more detail description in Section B.1 of the Appendix. We further illustrate the ability of our data-driven scheme to learn the dynamics in the latent coordinates. We compare the data-driven trajectories with the solution computed by the equations in the ambient space. The comparison is made either by mapping the full trajectories in the reduced Diffusion Maps coordinates with Nyström Extension (discussed in Section B.2.1 of the Appendix) –what we call *Restriction*– or by mapping the reduced data-driven trajectories back to the full space (*lifting*) by using a regression scheme that operates in the latent coordinates that we call *Latent Harmonics*, described in Algorithm 2. In Section 5, we show that the latent coordinates obtained with the Diffusion Maps algorithm are one-to-one with the latent coordinates of an autoencoder. We also show the ability of two other regression schemes (Gaussian Processes and Neural Networks) to learn the dynamics in the Diffusion Maps coordinates. We conclude by summarizing the main advantages and disadvantages of our approach and discuss current research directions that we believe can improve the construction of accurate data-driven models.

We propose a data-driven framework for the construction of reduced data-driven surrogate models for dynamical systems whose long term dynamics are intrinsically lower dimensional. We combine the ability of the Diffusion Maps algorithm to (a) reduce the ambient dimension of a data set by discovering the intrinsic lower-dimensional geometry of the data manifold and (b) to extend functions defined on this manifold to perform scientific computations in latent space. We present a new use case of the Geometric Harmonics algorithm [9], *Latent Harmonics*, that allow us to extend functions out of sample in the latent space.

We develop and test three multiscale numerical simulation methodologies by using tabulation in the low-dimensional space and integration on the fly or by mapping back and forth between the ambient space and the latent one, in the spirit of equation-free computing [25, 24, 26] We also address the interpretability of our reduced models by mapping back to the ambient space (lifting) through the Latent Harmonics.

The construction of Reduced Order Models (ROMs) typically involves using a dimensionality reduction algorithm that discovers a set of latent coordinates, followed by learning the dynamics in these coordinates. In the early 90s Proper Orthogonal Decomposition (POD) was used to find a set of reduced variables for chemical reacting systems or flows in complicated geometries [11, 28, 1]. Learning the dynamics in terms of those reduced variables was done either with projection in the leading POD modes [11] or by training a feed forward neural network [28] for the evolution of the POD-Galerkin projection coefficients. Beyond these references, POD has been proven useful in deriving reduced order models for a long list of applications in the literature [44, 38, 37, 16].

However, since POD achieves dimensionality reduction by approximately passing hyperplanes through the data points, one might envision intrinsically low-dimensional *but curved* inertial manifolds that require much higher dimensional hyperplanes. To circumvent this issue autoencoders have been used that project dynamical-systems onto non-linear manifolds [31] and that simultaneously learn the governing equations and the reduced coordinates [3] by minimizing a combined loss functions. Our work deviates from Champion *et al.* [3] since no regularization is needed to discover the latent space coordinates and learn the dynamics. However, our method -even though based on harmonic analysis- retains a certain resemblence to autoencoders in that the same data-driven approach is used to both discover the latent space and then map back to the ambient space. In our case this is not an optimization problem (as it is for autoencoders), but rather a deterministic data-driven algorithm. Indeed, our group has used autoencoders to discover reduced latent spaces since the early 1990s [36]; but at that time the nonlinear, noncnvex problem of training the autoencoder was much less straightforward than it is today, due to both software and hardware reasons.

The work of Chiavazzo *et al.* [6] and of Papaioannou*et al.* [34] are perhaps the closest to ours. Chiavazzo *et al.* use the example discussed also in Section 4.2 from Chemical Kinetics for the construction of the reduced data-driven models. They mainly examine how different regression algorithms can be used to map from the Diffusion Maps coordinates back to the ambient space. For the construction of the reduced data-driven models they use a similar approach to our Grid Tabulation (Section 3.2) algorithm. Papaioannou *et al.* [34] used Diffusion Maps as one of the dimensionality reduction techniques to find a set of latent coordinates. They learn the dynamics in those coordinates with Multivariate Autoregressive (MVAR) and Gaussian Processes. Similar to our work they lift by using a version of Geometric Harmonics. The key difference between Chiavazzo *et al.* [6] and Papaioannou *et al.* [34] and our current work is that, in order to lift with Geometric Harmonics, they use, at every new computational point, only the nearest few neighbors; we develop, implement and use a sensible global interpolation scheme here.

2 Double Diffusion Maps

The Diffusion Maps algorithm [8] can be used in two ways: (1) for dimensionality reduction of a finite data set, $\mathbf{X} = \{\boldsymbol{x}_i\}_{i=1}^N$ where $\boldsymbol{x}_i \in \mathbb{R}^m$, sampled from a manifold, $\mathcal{M} \subset \mathbb{R}^m$; or (2) to construct a function space basis called "Geometric Harmonics" that allows to extend functions defined on $\mathbf{X}$, $f(\mathbf{X}) : \mathbf{X} \to \mathbb{R}$ [30, 8, 9]. In this section, we describe the proposed Double Diffusion Maps scheme that employs two successive applications of Diffusion Maps.

Before we describe our Double Diffusion Maps Algorithm is worth mentioning briefly how the dimensionality reduction and the extension of a function can be achieved with Diffusion Maps. The Diffusion Maps algorithm applied to a data set $\mathbf{X}$ provides a parametrization for it. This parametrization aims to reveal the intrinsic geometry of $\mathbf{X}$. If the number of the significant non-harmonic (see Section B.2 in the Appendix for a discussion of harmonic and non-harmonic eigenvectors) Diffusion Map coordinates is smaller than the dimension of the ambient space, the algorithm achieves dimensionality reduction. Given a data set $\mathbf{X}$ sampled from a manifold $\mathcal{M}$, and a function f defined on the sampled data set $\mathbf{X}$, Geometric Harmonics aims to extend f for new points $\boldsymbol{x}_{new} \notin \mathbf{X}$. This map is constructed directly in the ambient space $\mathbb{R}^m$, Figure 2.

In our suggested scheme we take advantage of the two use cases of Diffusion Maps (dimensionality reduction and extending a function), and show that it is possible to extend the function of interest f by operating in only a few latent

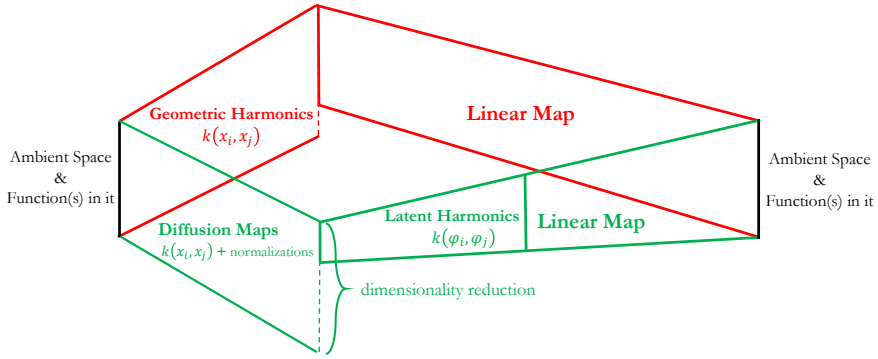


Figure 2: The upper, red schematic shows the mapping constructed from the ambient space and functions in it, via Geometric Harmonics, back to the ambient space and functions in it. The lower, green schematic illustrates our proposed approach that involves (a) first computing Diffusion Maps to discover the intrinsic geometry of a given data set (that might also lead to dimensionality reduction) and then (b) mapping from the *reduced* obtained latent space back to the ambient space and functions in it.

coordinates that parametrize the manifold, Figure 2. Our scheme proceeds as follows; first, we use Algorithm 1 to obtain a few, reduced, non-harmonic coordinates $\mathbf{\Phi}$ (if $\mathbf{X}$ is lower dimensional); then we use Algorithm 2 to construct a function space basis that we call Latent Harmonics ($\mathbf{\Psi}$) by operating only on the leading non-harmonic eigenvectors, Figure 2.

It is worth discussing why this second round of Diffusion Maps is indeed capable of extending functions defined on the original manifold. We remind the reader, that the eigenvectors $\mathbf{\Phi}$ were selected as the leading few non-harmonic eigenvectors that parametrize the (assumed low-dimensional) manifold; of course those were not the only computed eigenvectors resulting from the eigendecomposition in Algorithm 1. The discarded eigenvectors were functions of the ones selected, and therefore did not provide additional information in parametrizing the manifold. Yet, they are still important in extending a function using Geometric Harmonics - by discarding them we succeed in reducing dimensionality but lose the ability of extending functions defined on the ambient data manifold. In order to regain, our ability to accurately extend functions defined in the ambient data set taking advantage of the intrinsic manifold low dimensionality, the reconstruction of the discarded harmonics becomes necessary. This reconstruction is achieved through Algorithm 2 on $\mathbf{\Phi}$ through the computation of Latent Harmonics ($\mathbf{\Psi}$).

A particular example that we discuss in our numerical experiments is the extension of the state coordinates *in the high dimensional ambient space* for new points *in the latent space*, $f : \boldsymbol{\phi}_{new} \to \boldsymbol{x}_{new}$, what we refer to as *Lifting*.

Algorithm 1: Dimensionality Reduction

Diffusion Maps - Dimensionality Reduction ($\mathbf{X}$):

1. For each $\{\boldsymbol{x}_i\}_{i=1}^N \in \mathbf{\Phi}$ where $\boldsymbol{x}_i \in \mathbb{R}^m$ compute

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2}{2\epsilon_1}\right).$$

2. If $\mathbf{X}$ wasn't sampled uniformly:
   (a) Compute $P_{ii} = \sum_{j=1}^N K_{ij}$.
   (b) Apply the normalization $\widetilde{\mathbf{K}} = \mathbf{P}^{-\alpha}\mathbf{K}\mathbf{P}^{-\alpha}$ where $\alpha = 1$ to factor out the density.
3. Normalize the transition matrix $\widetilde{\mathbf{K}}$,

$$A(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{\widetilde{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)}{\sum_{j=1}^N \widetilde{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)}.$$

3. Solve the eigendecomposition $\mathbf{A}\boldsymbol{\phi}_i = \lambda_i \boldsymbol{\phi}_i$ for each i and then sort based on λ_i.
4. Select the set of non-harmonic eigenvectors $\mathbf{\Phi} \in \mathbb{R}^{N\times k}$, where $k \leq m$ [13].

return $\mathbf{\Phi}$;

Algorithm 2: Double Diffusion Maps & Latent Harmonics

Double Diffusion Maps $(\boldsymbol{\Phi}, f)$**:**

1. For each $\{\boldsymbol{\phi}_i\}_{i=1}^N \in \boldsymbol{\Phi}$ compute

$$K^\star(\boldsymbol{\phi}_i, \boldsymbol{\phi}_j) = \exp\left(-\frac{||\boldsymbol{\phi}_i - \boldsymbol{\phi}_j||_2^2}{2\epsilon_2}\right)$$

2. Compute the d first eigenvectors $\boldsymbol{\Psi} = \{\boldsymbol{\psi}_0, \ldots, \boldsymbol{\psi}_{d-1}\}$ where $\boldsymbol{\psi}_i \in \mathbb{R}^N$ and eigenvalues, $\boldsymbol{\sigma} = \{\sigma_0, \ldots, \sigma_{d-1}\}$, of $\mathbf{K}^\star$
3. **Training**: Project f on those eigenvectors

$$f \to P_\delta f = \sum_{j=1}^{d} \langle f, \boldsymbol{\psi}_j \rangle \boldsymbol{\psi}_j,$$

return;

Latent Harmonics $(\boldsymbol{\phi}_{new})$**:**

1. Compute the GH functions for ϕ_{new}

$$\Psi_j(\boldsymbol{\phi}_{new}) = \sigma_j^{-1} \sum_{i=1}^{m} K^\star(\boldsymbol{\phi}_{new}, \boldsymbol{\phi}_i)\psi_j(\phi_i)$$

where $\psi_j(\boldsymbol{\phi}_i)$ is the i^{th} component of the j^{th} eigenvector.
2. **Extension**: Estimate the value of the function for ϕ_{new}

$$(Ef)(\boldsymbol{\phi}_{new}) = \sum_{j=1}^{d} \langle f, \boldsymbol{\psi}_j \rangle \Psi_j(\boldsymbol{\phi}_{new})$$

return $f(\boldsymbol{\phi}_{new})$;

The computation of the latent coordinates discussed in Algorithm 1 and the construction of the functional basis from Double Diffusion Maps Algorithm 2 need to be performed only once.

Then the Latent Harmonics (Algorithm 2) can be used to extend the values of any function originally defined in the ambient space for any $\boldsymbol{\phi}_{new} \notin \boldsymbol{\Phi}$ in the neighborhood of the training set.

In our work, we further exploit the *Double Diffusion Maps-Latent Harmonics* scheme to perform scientific computations in the latent space.

3 Construction and Integration of Reduced Dynamical Models

In this section three different approaches are discussed for the construction and integration of reduced dynamical models. For the described algorithms below, it is assumed that the latent Diffusion Maps coordinates in which the models are constructed have already been obtained, Algorithm 1. To construct and validate those reduced dynamical models, global interpolation schemes are important that allows us to map new unobserved data from the ambient space coordinates to the reduced space original variables, restriction, but also from the reduced coordinates back to the original ones, lifting. In our work we *restrict* by using the Nyström Extension, described in Section B.2.1 of the Appendix. The inverse mapping of *lifting* is achieved by performing Algorithm 2. In general the Latent Harmonics allows to construct any smooth function of interest defined in the ambient or reduced space in terms of the few parsimonious reduced Diffusion Maps coordinates in which we aim to construct the data-driven surrogate model.

The aforementioned scientific tools are important to construct reduced models in terms of the leading non-harmonic Diffusion Maps coordinates. Of course building dynamical models in the Diffusion Maps coordinates means that we should be able to compute the evolution of $\boldsymbol{\phi}(t)$ with time. Expressing the evolution of $\boldsymbol{\phi}(t)$ through a simple numerical scheme such as Forward Euler we have:

$$\boldsymbol{\phi}_{t_{i+1}} = \boldsymbol{\phi}_{t_i} + h\left(\frac{d\boldsymbol{\phi}_{t_i}}{dt}\right) \tag{1}$$

where $\boldsymbol{\phi}_{t_i} = \boldsymbol{\phi}(t = i)$ and h is the step size. The computation of the dynamics in terms of $\boldsymbol{\phi}$ is non trivial, since we do not have access to closed form expressions of the right hand side of the evolution equations in the low dimensional space. It becomes possible to compute $\frac{d\boldsymbol{\phi}_{t_i}}{dt}$ since differentiating $\boldsymbol{\phi}(\boldsymbol{x}(t))$ through the chain rule gives

$$\frac{d\boldsymbol{\phi}_{t_i}}{dt} = \frac{d\boldsymbol{\phi}_{t_i}}{d\boldsymbol{x}_{t_i}} \cdot \frac{d\boldsymbol{x}_{t_i}}{dt}, \tag{2}$$

where $\frac{d\boldsymbol{x}_{t_i}}{dt}$ are the known dynamics in the ambient space coordinates for which we have analytical expressions. The term $\frac{d\boldsymbol{\phi}_{t_i}}{d\boldsymbol{x}_{t_i}}$ is computed by symbolic differentiation of the Nyström Extension Formula described in Section B.2.1 of the Appendix. These derivatives could be also computed by "automatic differentiation" that nowadays is available for scientific computations even when the function is only available in the form of a black box.

In our described algorithms below we show how the reduced dynamics are estimated by repeatedly going *Back and Forth* (BF) between the ambient space coordinates and the Diffusion Maps coordinates, Section 3.1. Connections between this BF algorithm and the Equation-Free modeling approach of [25] are described in Section C the Appendix. A second algorithm, described in Section 3.2 (GT) has many similarities with [6] in which the reduced dynamics are estimated and tabulated in the nodes of a Cartesian grid. The difference between our approach and the one discussed in [6] has to do with the computation of the reduced vector field on the nodes of the grid; further details are given in Section 3.2. In our last algorithm *Tabulation with Latent Harmonics Interpolation* (TaLHI), described in Section 3.3 $\frac{d\phi}{dt}$ is computed for *all the collected data points* $\boldsymbol{x}_i \in \mathbf{X}$ and globally tabulated.

3.1 Back and Forth

The BF approach allow us to integrate in the reduced coordinates, $\boldsymbol{\Phi}$, by repeatedly going Back and Forth between the reduced space and the ambient space. Algorithm 3 below assumes as input a point in the reduced space. If an initial condition in the ambient space is given, the corresponding value in the Diffusion Maps coordinates can be computed through the Nyström Extension formula, Section B.2.1.

Algorithm 3: Back and Forth

Input $(\boldsymbol{\phi}_{t_i})$**:**

1. Use the extension scheme from Algorithm 2 to *lift* $\boldsymbol{\phi}_{t_i}$.
2. Compute $\frac{d\boldsymbol{x}_{t_i}}{dt} = \boldsymbol{f}(\boldsymbol{x}_{t_i})$ with the original model.
3. Compute $\frac{d\boldsymbol{\phi}_{t_i}}{dt}$ from the chain rule in Equation 2.

return $\frac{d\boldsymbol{\phi}_{t_i}}{dt}$

Given $\frac{d\boldsymbol{\phi}_{t_i}}{dt}$, the value $\boldsymbol{\phi}_{t_{i+1}}$ can be estimated through any ODE solver. Figure 3 illustrates a caricature of the algorithm for a toy example in which the ambient space is $\mathbb{R}^2$ and the reduced space $\mathbb{R}$.

3.2 Grid Tabulation

The GT algorithm tabulates the reduced vectorfield only for the nodes $\widetilde{\boldsymbol{\phi}}_i \in \widetilde{\boldsymbol{\Phi}}$ in a Cartesian $(N \times N)$ grid. Given the

Algorithm 4: Grid Tabulation

Input $(\widetilde{\boldsymbol{\Phi}}$,N**):**

1. Lift by using the extension scheme from Algorithm 2 for each $\widetilde{\boldsymbol{\phi}}_i \in \widetilde{\Phi}$.
2. Compute $\frac{d\widetilde{\boldsymbol{x}}_i}{dt} = \boldsymbol{f}(\widetilde{\boldsymbol{x}}_i)$ from the full Equations.
3. For each $\widetilde{\boldsymbol{\phi}}_i$ compute $\frac{d\widetilde{\boldsymbol{\phi}}_{t_i}}{dt}$, from Equation (2).

return $\frac{d\widetilde{\boldsymbol{\Phi}}}{dt}$

tabulated vector field, the integration of the model is performed in the reduced space through a multi-linear interpolation, without the need to map back to the ambient space at each integration step. The use of multi-linear interpolation

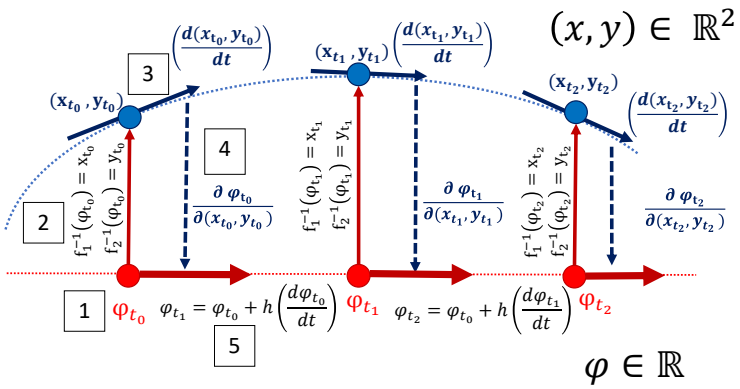


Figure 3: 1. Start integration at $t = 0$ from an initial condition in the reduced space ϕ_{t_0}. 2. Use Algorithm 2 to compute the corresponding values of the high dimensional space variables, x_{t_0}, y_{t_0}. 3. From x_{t_0}, y_{t_0} by using the expressions of the dynamics in the high dimensional space compute the derivatives $\frac{d(x_{t_0}, y_{t_0})}{dt}$. 4. To obtain the reduced dynamics $\frac{d\phi_{t_0}}{dt}$ the calculation of the chain rule is needed $\frac{d\phi_{t_0}}{dt} = \frac{\partial \phi_{t_0}}{\partial x_{t_0}} \cdot \frac{dx_{t_0}}{dt} + \frac{\partial \phi_{t_0}}{\partial y_{t_0}} \cdot \frac{dy_{t_0}}{dt}$, where the terms $\frac{\partial \phi_{t_0}}{\partial (x_{t_0}, y_{t_0})}$ are computed by differentiation of the Nyström Extension formula. With known the value of $\frac{d\phi_{t_0}}{dt}$ the value of ϕ_{t_1} is estimated. By repeating the steps 2-5 integration in the Diffusion Maps coordinates is achieved.

is not significant: any interpolation scheme could have been used instead. The GT algorithm is similar to the one presented by Chiavazo *et al.* [6]; the main difference in our work is the use of a *global interpolation scheme* for the entire manifold during the computation of the first step in Algorithm 4. In [6] the authors estimate the reduced derivative for each node in the grid via the nearest k neighbors, and then lift through Geometric Harmonics *only on those k neighbors* (what we here term as Latent Harmonics). This then involves training a regression scheme for every point in the grid separately and not all at once, as in our case.

3.3 Tabulation with Latent Harmonics Interpolation (TaLHI)

Our last proposed algorithm, TaLHI, learns the reduced vector field as a function of the reduced Diffusion Maps coordinates by using the Latent Harmonics regression scheme. For TaLHI the estimation of the $\frac{d\mathbf{\Phi}}{dt}$ is performed as follows, Algorithm 5:

Algorithm 5: TaLHI

function `Input (X):`

1. Compute $\frac{d\boldsymbol{x}_i}{dt} = f(\boldsymbol{x}_i)$ for each $\boldsymbol{x}_i \in \mathbf{X}$.
2. Compute $\frac{d\boldsymbol{\phi}_i}{dt}$ by using Equation (2).

return $\frac{d\mathbf{\Phi}}{dt}$

The main difference of TaLHI with the other three discussed algorithms is that it doesn't utilize a *lifting* step for the estimation of the reduced derivative since, for each $\boldsymbol{\phi}_i \in \mathbf{\Phi}$ we already have the corresponding data point $\boldsymbol{x}_i$ in the ambient space. Latent Harmonics regression is only used to construct a mapping from the Diffusion Maps coordinates to their reduced vector field $\frac{d\mathbf{\Phi}}{dt}$.

4 Results

We test our suggested algorithms described in Sections 3.1-3.3 through two examples. The first example in Section 4.1 is the Chafee-Infante reaction diffusion PDE. The second example, Section 4.2, comes from Chemical Kinetics and more precisely from the combustion of hydrogen [6].

4.1 The Chafee-Infante PDE

To illustrate how reduced dynamical models can be obtained for dissipative PDEs, we utilize the Chafee-Infante reaction-diffusion equation

$$u_t = u - u^3 + \nu u_{xx} \tag{3}$$

with ν = 0.16, and boundary conditions $u(0,t) = u(\pi,t) = 0$. This PDE for $\nu = 0.16$ it is known that has a two-dimensional inertial manifold [39, 18, 22]. Diffusion Maps can reveal this two-dimensional manifold embedded in a higher dimensional space given data $\mathbf{X}$ sampled from simulations of this PDE [39, 18]. More precisely, to sample data, $\mathbf{X}$, firstly we approximate u with a discrete sine transformation:

$$u(x,t) \approx \sum_{k=1}^{10} \alpha_k(t) sin(kx) \tag{4}$$

The choice of 10 terms is not significant, results are similar for other levels of accuracy in the approximation [18]. The Galerkin projection onto those 10 Fourier modes $\{\text{sin(x)}, \ldots, \text{sin(10x)}\}$, result in a system of equations of the form:

$$\frac{d\boldsymbol{\alpha}}{dt} = \boldsymbol{f}(\boldsymbol{\alpha}) \tag{5}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{10}$. To sample points on the inertial manifold we integrate this system of ODEs, Equation 5, for different initial conditions $\boldsymbol{\alpha}_{t_0}$ and keep only the long-term dynamics. Since we would like to ensure that our sample is as uniform as possible in the ambient space we perform a subsampling where data within a threshold distance d are removed. Thus a uniform data set $\mathbf{X}$ with ~ 2800 data points is obtained.

Algorithm 1 computed on $\mathbf{X}$ reveals a 2-D manifold embedded in $\mathbb{R}^{10}$. The Diffusion Maps independent coordinates, ϕ_1, ϕ_2, that fully parametrize this 2-D manifold are also one-to-one with the first two Fourier modes α_1,α_2 as it can be seen in Figure 4. The selected hyperparameters used for Diffusion Maps are reported in Section E of the Appendix. For the Diffusion Maps computations the *datafold* package was used [32].

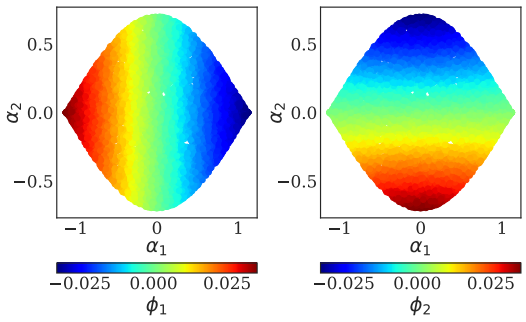
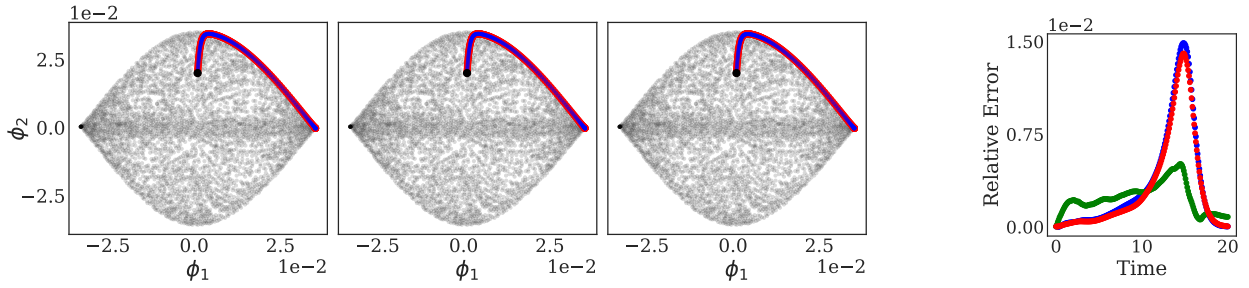


Figure 4: The first two Fourier modes α_1 and α_2 are colored with the two independent Diffusion Maps coordinates ϕ_1 (left) and ϕ_2 (right). The Fourier mode α_1 is one-to-one with ϕ_1 and α_2 is one-to-one with ϕ_2.

In Figure 5a, we illustrate a qualitative comparison between the proposed data-driven integration approaches and the restricted (with Nyström extension) solution computed from integrating the original spectral ODEs for a single trajectory. In addition, in Figure 5b the relative error between the restricted trajectory and the three data-driven trajectories is shown in time.

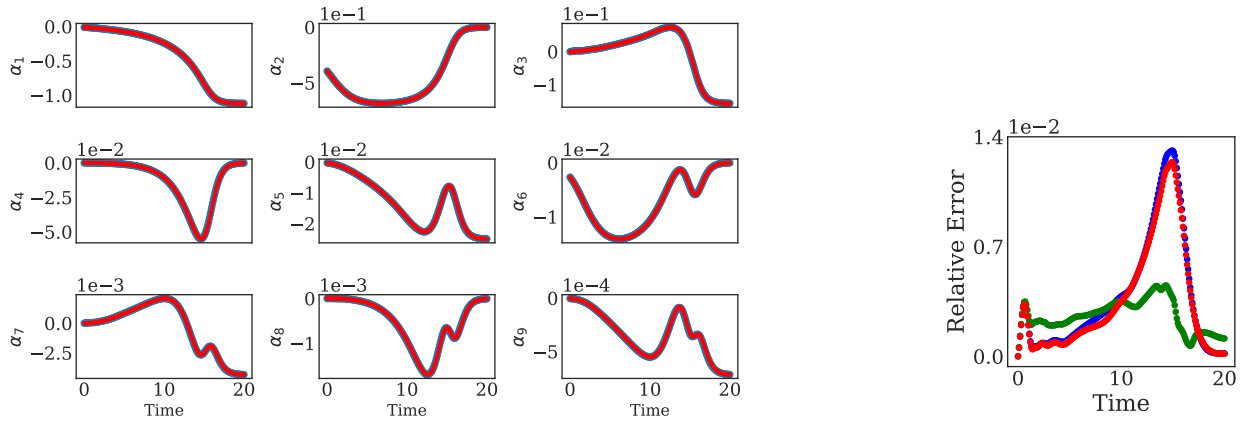
A comparison between the data-driven trajectories in the ambient space is also possible. Figure 6a shows the comparison between the lifted trajectory computed with TaLHI algorithm and the first 9 components for a trajectory coming from the full, unreduced equations. Qualitative comparisons in the ambient space for the BF and GT algorithms are shown in Section D of the Appendix. In Figure 6b also the relative error in time between the trajectories computed with the data-driven algorithms and the trajectory from the full, unreduced equations is shown in time.

The reconstruction from the latent coordinates back to the original $u(x,t)$ space of the PDE in Equation (3) is also possible through our scheme. In Figure 7 the comparison between a trajectory of the Chafee-Infante PDE is contrasted to a lifted and reconstructed trajectory with the same initial condition, integrated with the data-driven scheme TaLHI in the latent coordinates.



(a) (b)

Figure 5: (a) A comparison between a trajectory computed with the three data-driven frameworks TaLHI, BF, and GT, respectively is shown (blue color) against with a restricted (with Nyström Extension) trajectory obtained from the spectral ODEs (red color) for the Chafee-Infante example. (b) The relative error between the data-driven trajectories and the reduced trajectory from the unreduced equations.



(a) (b)

Figure 6: (a) A comparison between the lifted with Latent Harmonics trajectory computed from the data-driven integration scheme TaLHI (blue color) and the trajectory computed from the full, unreduced equations (red color) is shown for the first nine Fourier Modes. (b) The relative errors in time between the data-driven trajectories and the trajectory from the full, unreduced equations, BF (blue), GT (green), TaLHI (red).

4.2 Chemical Kinetics

The second application with which we want to test our algorithms is a system of stiff ODES. This example involves constructing a reduced kinetic model for a combustion example. This system of equations describes the combustion of hydrogen and air at stoichiometric proportions under fixed total enthalpy (H = 300 $\frac{kJ}{kg}$) and pressure (P = 1 bar). For this dynamical system the long term dynamics live in a two-dimensional manifold [6]. We used the detailed mechanism of Li et al. [33], in which nine chemical species (H_2, N_2, H, O, OH, O_2, H_2O, HO_2, H_2O_2) and three elements (H, O, N) are participating, with the concentration of N_2 to remain constant during the reaction. The selection of this second example on Chemical Kinetics and our Grid Tabulation (GT) discussed in Section 3.2 approach aims to give a comparison between our proposed scheme and the work of Chiavazzo et al [6]. We have used the same model and also the same sampling strategy trying to mimic the work of Chiavazzo et al [6] to collect a data set $\mathbf{X}$ of $\sim$ 3800 data points.

Algorithm 1 applied on our collected data discovers a two-dimensional manifold embedded in $\mathbb{R}^9$, Figure 8. On those latent coordinates we test our schemes for constructing reduced dynamical models. It is worth mentioning that this system of equations is stiff because of the widely disparate time scales, and also because the concentrations of some chemical species change by multiple orders of magnitude. This makes the global interpolation between the Diffusion Maps space and the ambient space numerically challenging. It is worth reminding the reader that the main difference between our lifting scheme and the lifting used with Latent Harmonics in [6] is that in our approach we build our extension scheme by parametrizing *the entire manifold* in contrast to the local approach (nearest neighbors) used

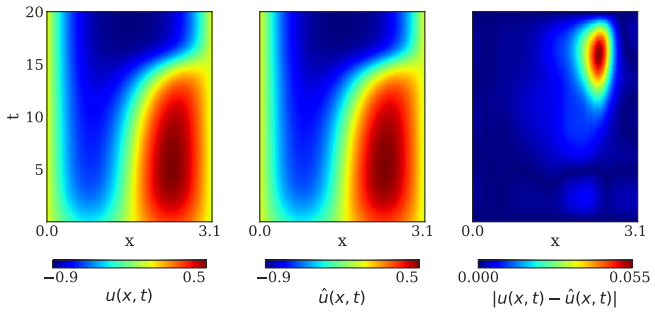


Figure 7: A trajectory $u(x,t)$ of the Chafee Infante PDE, Equation (3), integrated with Finite Differences (left) compared to a lifted & reconstructed trajectory $\hat{u}(x,t)$ obtained with the data-driven scheme TaLHI. The absolute error of $u(x,t)$ and the data-driven trajectory $\hat{u}(x,t)$ is shown in the right.

in [6]. Therefore Latent Harmonics in our case attempts a more challenging approach, namely global learning and extension of stiffer functions. In our case to learn the dynamics the *training* discussed in Algorithm 2 needs to be performed only once given all the available data points sampled on the manifold; then any point in the *neighborhood* can be *lifted*. In the work of Chiavazzo et al. [6] the *training* of the for *lifting* any point of interest needs to be performed separately for each point. A comparison between the computational time needed to construct a 60×60 grid with our GT approach (12.50 mins) and the approach discussed in Chiavazzo et al (45.10 mins) was performed (both experiments conducted on a Intel(R) Xeon(R) Silver 4112 CPU @ 2.60GHz Processor) and shows a speed up of for our tabulation scheme with GT.

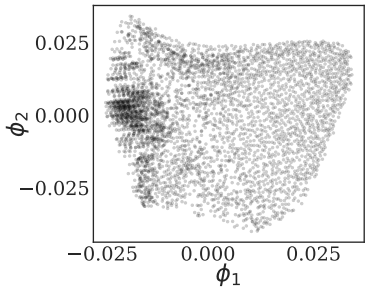
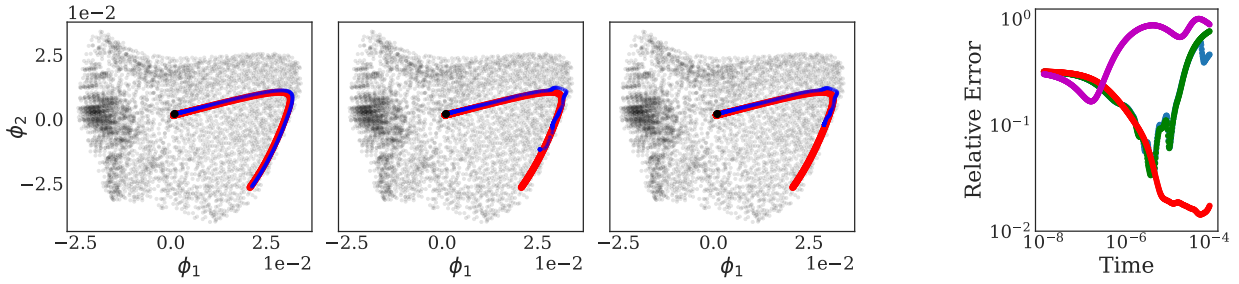


Figure 8: The two dimensional manifold in terms of the independent eigenvectors ϕ_1 and ϕ_2 discovered on Chemical Kinetics example.

The goal for this example remains the same as before, to construct a data-driven integration scheme in the latent coordinates. Figure 9a illustrates a qualitative comparison between the restricted with Nyström full, unreduced equations and the three proposed data-driven approaches. In Figure 9b the relative error between the trajectory of the restricted equations and the data-driven trajectories is shown. We computed also the trajectory with the approach discussed in Chiavazzo et al. [6] to give a more quantitative comparison between our suggested approach an the one in [6].Figure 10a shows the comparison between the lifted data-driven trajectory integrated with TaLHI and a trajectory of the full system of equations. Qualitative comparisons between the trajectory computed from the full model and trajectories computed with the data-driven integration schemes BF and GT in the ambient space are given in Section D of the Appendix. The relative error between the trajectory of the full equations and lifted trajectories from our suggested data-driven schemes including the one from Chiavazzo et al [6] for each point along the simulated trajectories are shown in Figure 10b.

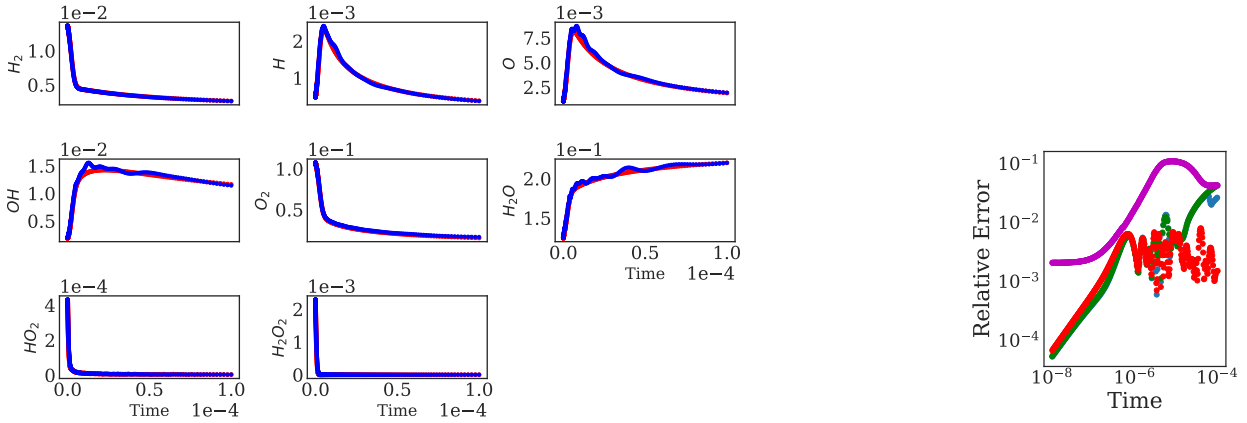
5 Discussion-Conclusions

We presented three different approaches that allow us to construct reduced data-driven models in the Diffusion Maps coordinates. The key elements for constructing reduced data-driven models thought the Double Diffusion Maps scheme were described in Algorithms 1 and 2 along with the Nyström Extension, discussed in detail in Section B.2.1 of the Appendix. We tested our algorithms in two different examples in which the long term dynamics lie on a low-



(a) (b)

Figure 9: (a) A comparison between a trajectory computed from the three data-driven frameworks TaLHI, BF, and GT, respectively is shown (blue color) and a restricted (with Nyström Extension) trajectory obtained from the spectral ODEs (red color) for the example from Chemical Kinetics.(b) The relative error in time of the four data driven approaches for each point along the simulated trajectory, BF (blue), GT (green), TaLHI (red), Chiavazzo et al [6] (mangenta), each compared to the restricted with Nyström Extension trajectory of the full dynamics.



(a) (b)

Figure 10: (a) Comparison between the lifted with Latent Harmonics trajectory computed with the data-driven integration scheme Back and Forth (blue trajectory) and the trajectory computed from the full, unreduced equations of the kinetics (red trajectory). (b) The relative error in the ambient space between the lifted data driven schemes and a trajectory of the full dynamics, BF (blue), GT (green), TaLHI (red), Chiavazzo et al [6] (mangenta). The lifting in our proposed data-driven schemes was performed with Algorithm 2 and the one in Chiavazzo et al. with *local* Geometric Harmonics.

dimensional, slow-inertial manifold. To test our Double Diffusion Maps scheme we discover a set of latent coordinates and test our suggested schemes for the construction of reduced dynamical models. We compared the data-driven trajectories to the dynamics of the closed form expressions either in the latent coordinates or in the ambient space, In the ambient space we used the Latent Harmonics (Algorithm 2) to lift our reduced trajectories in the ambient space. The Comparison in the reduced coordinates was achieved by restricting the full trajectories computed from the equations in the reduced space through Nyström Extension.

The proposed approaches named, BF, GT and, TaLHI have their strengths and weaknesses. The tabulation schemes GT and TaLHI are computationally expensive and need to allocate memory during the tabulation process. However, the tabulation needs to be performed only once.

Based on our examples TaLHI seems to be more accurate compared to the other two approaches. This can be perhaps attributed to the fact that this method does not acquire lifting in order to learn the reduced dynamics.

The BF approach avoids the expensive tabulation step. Integrating in the reduced coordinates without the need to tabulate in advance the vector field would be even more beneficial in cases the reduced space coordinates are higher than two. In addition, the BF approach was constructed under the assumption that by decreasing the dimensions of a

dynamical system its stiffness also decreases. This seems to be valid in our examples, and so larger time steps can be used in the reduced space compared to the ambient space.

It seems that whenever the trajectory visits the boundary of a manifold it becomes more prone to numerical errors. Additional sampling in the boundary of the manifold will, we expect, increase the interpolation accuracy of our scheme and thus give more accurate dynamic models [19] .

In our current work we chose, for discovering the latent coordinates, the manifold learning method Diffusion Maps. Of course, other non-linear dimensionality reduction methods could have been used alternatively. To justify that we trained an autoencoder neural network with two bottle neck variables for the Chafee-Infantee example discussed in Section 4.1. The two latent coordinates computed from the Autoencoder (ν_1, ν_2) are one-to-one with the Diffusion Maps variables. The one-to-one relationship between the two sets of variables was confirmed by learning a mapping $f : N \to \Phi$ with Geometric Harmonics and the inverse mapping $f^{-1} : \Phi \to N$ with Latent Harmonics. In Figure 11 we plot the relative error computed for each test set point as function of the (rescaled to the range 0 to 1) Diffusion Maps Coordinates and Autoencoders Coordinates respectively. We believe that the natural use of the intrinsic geometry dictated by the data, discovered through an eigenproblem, is "more direct" than an autoencoder-based latent space, even if both lead to comparable reduction. Indeed, the intrinsic-geometry based approach will lead always to the same reduced space parametrization for the same data; an autencoder based one will/may lead to a different (equivalent) representation for different researchers/different optimization runs, unless a systematic regularization is used.

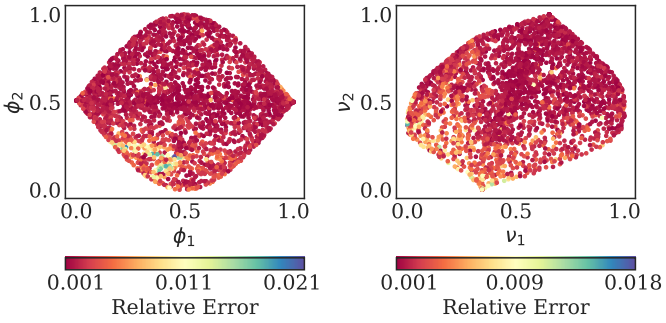


Figure 11: The relative error $\frac{|\phi_i - \hat{\phi}_i|}{\phi_i}$ of the forward mapping, $f : N \to \Phi$ is shown on the left as a function on the original values of the Diffusion Maps coordinates. The relative error $\frac{|\nu_i - \hat{\nu}_i|}{\nu_i}$ of the backward mapping $f^{-1} : \Phi \to N$ is shown on the right. In both cases the coordinates are rescaled in the range 0 to 1.

In our work since the reduced models are constructed in the Diffusion Maps coordinates it made sense to use a regression scheme based on its extension; however different regression algorithms could have also been used instead. In Figure 12 we show the ability to learn the reduced dynamics in the Diffusion Maps coordiantes and integrate with Gaussian Proccesses and also with a Neural Network.

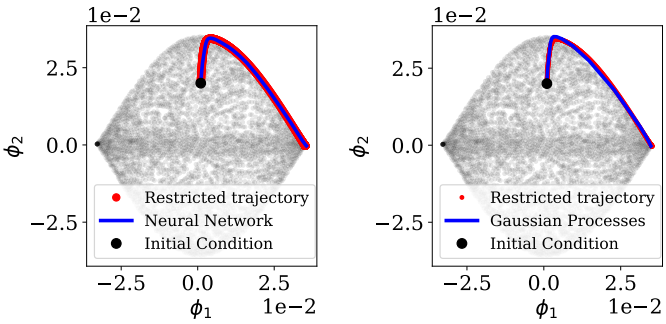


Figure 12: data-driven Integration with a Neural Network (on the left) and Gaussian Processes (on the right).

We conclude, by mentioning that the focus of this paper was on the construction of reduced models for dynamical systems whose long term behavior live in a lower dimensional manifold. We illustrated three possible algorithms with which a reduced data-assisted model can be constructed. Those suggested algorithms could be used to any

dynamical systems of ordinary or partial differential equation with reduced long term behavior. Those data-driven models can be used as a more efficient, compressed simulator simulator in cases that using the full model repeatedly is computationally inefficient (e.g. control processing). In our algorithms, we assumed that sampling the whole portion of interest of the manifold is necessary; however this might not be possible or computationally efficient. In problems in science and engineering data might be given only in a small neighborhood of their entire manifold whose dimensionality also might not be known in advance. In that direction further research might needed in order to utilize *effective* sampling algorithms combined with algorithms that aim to learn operators on the data [5].

In addition, the latent coordinates in our case were computed without taking into consideration physical constrains that may exist in their original model (e.g. mass conservation during chemical reactions). Infusing such constraints in physics-informed models [35] constitutes another promising research area.

Funding: This work was partially supported by the US AFOSR and the US Department of Energy.

Acknowledgements: N.E with the help of F.D, performed the Diffusion Maps computations and developed the code for the three suggested schemes for the construction of reduced data-driven models. N.E, F.D, and I.G.K wrote the paper. E.C, provided data and models. All authors designed and performed research. All authors contributed in manuscript revision.

References

[1] Peter Benner, Serkan Gugercin, and Karen Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review*, 57(4):483–531, 2015.

[2] Tyrus Berry, John Cressman, Z. Gregurić-Ferenček, and T. Sauer. Time-scale separation from diffusion-mapped delay coordinates. *SIAM Journal on Applied Dynamical Systems [electronic only]*, 12, 01 2013.

[3] Kathleen Champion, Bethany Lusch, J Nathan Kutz, and Steven L Brunton. Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, 116(45):22445–22451, 2019.

[4] Yu-Chia Chen and Marina Meila. Selecting the independent coordinates of manifolds with large aspect ratios. *Advances in Neural Information Processing Systems*, 32, 2019.

[5] Eliodoro Chiavazzo, Roberto Covino, Ronald R. Coifman, C. William Gear, Anastasia S. Georgiou, Gerhard Hummer, and Ioannis G. Kevrekidis. Intrinsic map dynamics exploration for uncharted effective free-energy landscapes. *Proceedings of the National Academy of Sciences*, 114(28):E5494–E5503, 2017.

[6] Eliodoro Chiavazzo, Charles Gear, Carmeline Dsilva, Neta Rabin, and Ioannis Kevrekidis. Reduced models in chemical kinetics via nonlinear data-mining. *Processes*, 2(1):112–140, Jan 2014.

[7] R. R. Coifman, Y. Shkolnisky, F. J. Sigworth, and A. Singer. Graph laplacian tomography from unknown random projections. *IEEE Transactions on Image Processing*, 17(10):1891–1899, 2008.

[8] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5 – 30, 2006. Special Issue: Diffusion Maps and Wavelets.

[9] Ronald R. Coifman and Stéphane Lafon. Geometric harmonics: A novel tool for multiscale out-of-sample extension of empirical functions. *Applied and Computational Harmonic Analysis*, 21(1):31 – 52, 2006. Special Issue: Diffusion Maps and Wavelets.

[10] Peter Constantin, Ciprian Foias, Basil Nicolaenko, and Roger Temam. *Integral manifolds and inertial manifolds for dissipative partial differential equations*, volume 70. Springer Science & Business Media, 2012.

[11] AE Deane, IG Kevrekidis, G Em Karniadakis, and SA Orszag. Low-dimensional models for complex geometry flows: application to grooved channels and circular cylinders. *Physics of Fluids A: Fluid Dynamics*, 3(10):2337–2354, 1991.

[12] C R Doering, J D Gibbon, D D Holm, and B Nicolaenko. Low-dimensional behaviour in the complex ginzburg-landau equation. *Nonlinearity*, 1(2):279–309, may 1988.

[13] Carmeline J Dsilva, Ronen Talmon, Ronald R Coifman, and Ioannis G Kevrekidis. Parsimonious representation of nonlinear dynamical systems through manifold learning: A chemotaxis case study. *Applied and Computational Harmonic Analysis*, 44(3):759–773, 2018.

[14] Ciprian Foias, George R Sell, and Roger Temam. Inertial manifolds for nonlinear evolutionary equations. *Journal of Differential Equations*, 73(2):309 – 353, 1988.

[15] C. Fowlkes, S. Belongie, and J. Malik. Efficient spatiotemporal grouping using the nystrom method. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–I, 2001.

[16] C.E. Frouzakis, Y.G. Kevrekidis, J. Lee, K. Boulouchos, and A.A. Alonso. Proper orthogonal decomposition of direct numerical simulation data: Data reduction and observer construction. *Proceedings of the Combustion Institute*, 28(1):75–81, 2000.

[17] C. W. Gear, E. Chiavazzo, and I. G. Kevrekidis. Manifolds defined by points: Parameterizing and boundary detection (extended abstract). *AIP Conference Proceedings*, 1738(1):020005, 2016.

[18] C. W. Gear, I. G. Kevrekidis, and B. E. Sonday. Slow manifold integration on a diffusion map parameterization. *AIP Conference Proceedings*, 1389(1):13–16, 2011.

[19] Anastasia Georgiou, Juan Bello-Rivas, Charles Gear, Hau-Tieng Wu, Eliodoro Chiavazzo, and Ioannis Kevrekidis. An exploration algorithm for stochastic simulators driven by energy gradients. *Entropy*, 19(7):294, Jun 2017.

[20] J. M. Ghidaglia and B. Herón. Dimension of the attractors associated to the ginzburg-landau partial differential equation. *Phys. D*, 28(3):282–304, October 1987.

[21] Michael S. Jolly. Explicit construction of an inertial manifold for a reaction diffusion equation. *Journal of Differential Equations*, 78(2):220 – 261, 1989.

[22] Michael S. Jolly. Explicit construction of an inertial manifold for a reaction diffusion equation. *Journal of Differential Equations*, 78(2):220–261, January 1989.

[23] M.S. Jolly, I.G. Kevrekidis, and E.S. Titi. Approximate inertial manifolds for the kuramoto-sivashinsky equation: Analysis and computations. *Physica D: Nonlinear Phenomena*, 44(1):38 – 60, 1990.

[24] Ioannis G Kevrekidis, C William Gear, and Gerhard Hummer. Equation-free: The computer-aided analysis of complex multiscale systems. *AIChE Journal*, 50(7):1346–1355, 2004.

[25] Ioannis G Kevrekidis, C William Gear, James M Hyman, Panagiotis G Kevrekidid, Olof Runborg, Constantinos Theodoropoulos, et al. Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis. *Communications in Mathematical Sciences*, 1(4):715–762, 2003.

[26] Ioannis G Kevrekidis and Giovanni Samaey. Equation-free multiscale computation: Algorithms and applications. *Annual review of physical chemistry*, 60:321–344, 2009.

[27] Samson Koelle, Hanyu Zhang, Marina Meila, and Yu-Chia Chen. Manifold coordinates with physical meaning. *arXiv e-prints*, pages arXiv–1811, 2018.

[28] K Krischer, R Rico-Martínez, IG Kevrekidis, HH Rotermund, G Ertl, and JL Hudson. Model identification of a spatiotemporally varying catalytic reaction. *AIChE Journal*, 39(1):89–98, 1993.

[29] J. Nathan Kutz. *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*. Oxford University Press, Inc., USA, 2013.

[30] S.S. Lafon. *Diffusion Maps and Geometric Harmonics*. PhD dissertation, Yale University, 2004.

[31] Kookjin Lee and Kevin T Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *Journal of Computational Physics*, 404:108973, 2020.

[32] Daniel Lehmberg, Felix Dietrich, Gerta Köster, and Hans-Joachim Bungartz. datafold: data-driven models for point clouds and time series on manifolds. *Journal of Open Source Software*, 5(51):2283, 2020.

[33] Juan Li, Zhenwei Zhao, Andrei Kazakov, and Frederick L Dryer. An updated comprehensive kinetic model of hydrogen combustion. *International journal of chemical kinetics*, 36(10):566–575, 2004.

[34] Panagiotis Papaioannou, Ronen Talmon, Ioannis Kevrekidis, and Constantinos Siettos. Time series forecasting using manifold learning, 2021.

[35] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[36] R Rico-Martinez, K Krischer, IG Kevrekidis, MC Kube, and JL Hudson. Discrete-vs. continuous-time nonlinear signal processing of cu electrodissolution data. *Chemical Engineering Communications*, 118(1):25–48, 1992.

[37] Eli Shlizerman, Edwin Ding, Matthew O Williams, and J Nathan Kutz. The proper orthogonal decomposition for dimensionality reduction in mode-locked lasers and optical systems. *International Journal of Optics*, 2012, 2012.

[38] Stanislav Y Shvartsman, C Theodoropoulos, Roberto Rico-Martınez, IG Kevrekidis, Edriss S Titi, and TJ Mountziaris. Order reduction for nonlinear dynamic models of distributed reacting systems. *Journal of Process Control*, 10(2-3):177–184, 2000.

[39] B. Sonday. *Systematic Model Reduction for Complex Systems Through Data Mining and Dimensionality Reduction*. PhD dissertation, Princeton University, 2011.

[40] Steven H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Westview Press, 2000.

[41] Mario Taboada. Finite-dimensional asymptotic behavior for the swift-hohenberg model of convection. *Nonlinear Analysis: Theory, Methods & Applications*, 14(1):43 – 54, 1990.

[42] Ronen Talmon, Israel Cohen, Sharon Gannot, and Ronald R Coifman. Diffusion maps for signal processing: A deeper look at manifold-learning techniques based on kernels and graphs. *IEEE signal processing magazine*, 30(4):75–86, 2013.

[43] Lloyd Nicholas Trefethen. *Finite difference and spectral methods for ordinary and partial differential equations*. Cornell University-Department of Computer Science and Center for Applied, 1996.

[44] Yongchun Zhang, Michael A. Henson, and Yannis G. Kevrekidis. Nonlinear model reduction for dynamic analysis of cell population models. *Chemical Engineering Science*, 58(2):429–445, 2003.

A Appendix

B Methodology

B.1 Diffusion Maps

Dimensionality reduction In this use case, the Diffusion Maps algorithm can reveal the intrinsic geometry of a given data set $\mathbf{X}$. This is achieved by first defining the local similarity between the sampled data points on $\mathbf{X}$. With this similarity, a weighted graph $\mathbf{K} \in \mathbb{R}^{N\times N}$ is constructed between all data points. The entries of this graph are often computed with a kernel function. In our case, the diffusion kernel was used:

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2}{2\epsilon}\right), \tag{6}$$

where ϵ is a hyper-parameter that specifies the (square of the) kernel bandwidth [17]. If points $\boldsymbol{x}_i, \boldsymbol{x}_j$ satisfy $||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2 < \epsilon$, they are considered similar (connected), whereas points with distance larger than ϵ^2 are, effectively, not linked directly. Even though in this paper only the Euclidean distance of the points was used, in general, different metrics are also possible.

It is worth mentioning that Diffusion Maps is based on the approximation of the heat kernel $e^{-t\Delta}$, where Δ is the Laplace-Beltrami operator on a manifold M, in the limit $n \to \infty, \epsilon \to 0$ [8, 2]. Because the sampling density affects such a kernel estimate, a normalization is needed to recover the correct operator:

$$P_{ii} = \sum_{j=1}^{N} K_{ij}, \tag{7}$$

$$\widetilde{\mathbf{K}} = \mathbf{P}^{-\alpha}\mathbf{K}\mathbf{P}^{-\alpha}. \tag{8}$$

The bias parameter α is being introduced to control the influence that the sampling density has on the geometry. When $\alpha = 0$ is chosen the influence of density is maximal and we get the approximation of the Laplace-Beltrami operator only in the case of uniform sampling [7]. In the case $\alpha = 1$ the density effects are factored out and the Laplace-Beltrami operator approximation is obtained. A second normalization is applied,

$$A(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{\widetilde{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)}{\sum_{j=1}^{N} \widetilde{K}(\boldsymbol{x}_i, \boldsymbol{x}_j)}, \tag{9}$$

which leads to the construction of $\mathbf{A}$, a row-stochastic, Markovian matrix. The eigendecomposition of $\mathbf{A}$ has a set of eigenvectors, $\boldsymbol{\phi}_i$ and eigenvalues λ_i.

$$\mathbf{A}\boldsymbol{\phi}_i = \lambda_i \boldsymbol{\phi}_i \tag{10}$$

Appropriate selection of the eigenvectors $\boldsymbol{\phi}_i, i \geq 1$, that correspond to unique eigendirections yields a more parsimonious representation of the data [13]. More precisely, if the number of those independent non-harmonic eigenvectors (e.g. k) is smaller than the dimensions of the ambient space ($k < m$) of $\mathbf{X}$ those k new coordinates, $\boldsymbol{\Phi} = \{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, ..., \boldsymbol{\phi}_k\}$, can be used for a new lower dimensional parametrization of the original high dimensional data set.

B.2 Harmonic and Non-Harmonic Eigenvectors

This section aims to provide a more clear theoretical justification of our techniques and discuss the selection of the *non-harmonic* Diffusion Maps coordinates. We start by reminding the reader how the selection of the important *modes* is performed with Proper Orthogonal Decomposition (P.O.D) to give a more clear comparison with the selection with Diffusion Maps. In P.O.D. one firstly samples the behavior of the underlying dynamical system to obtain a representative data ensemble. Next, one applies Singular Value Decomposition (S.V.D.) to this data set. The spectrum of the singular values computed gives an indication of how many dimensions are necessary to capture the important *modes* of a dynamical system. This is based on the assumption that the leading singular vectors that capture most of the variance capture also the most important dynamics. A useful reduced dimensionality of a given data set with P.O.D. is chosen as a compromise between reduction of complexity and reconstruction error. From here, expressing the reduced dynamics can be done by performing a projection of the original dynamics to the chosen important modes [11, 29].

For Diffusion Maps the selection of the latent coordinates that parametrize the data set should not be done based (only) on the information provided by the eigenvalues even though in some case an eigengap can give a hint [42]. The

reason is that some eigenvectors are parametrizing the same direction. We refer to those eigenvectors, as *harmonic* eigenvectors and for the set of eigenvectors that parametrize different directions as *non-harmonic*. Therefore our goal of discovering the minimal set of Diffusion Maps eigenvectors is restricted on selecting the *non-harmonic* eigenvectors.

In the remaining of this section we provide a toy example to elucidate the concepts of *harmonic* and *non-harmonic* eigenvectors . We sampled uniformly 10,000 data points on a rectangle with ratio of it's length to width equal to 4:1, Figure 13. For this toy example since the Diffusion Maps algorithm (for $\alpha = 1$ or for data with uniform sampling density) approximates the Laplace-Beltrami operator with Neumann boundary conditions the eigenvalues and the eigenvectors can be computed also analytically [4, 13] and it is known that the order of the Diffusion Maps coordinates that parametrize this data set is based on the ratio of it's sides [4].

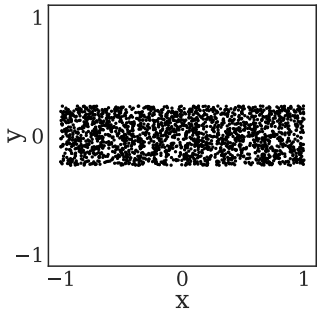


Figure 13: Data points collected from rectangle with sides' ratio 4:1.

We performed Diffusion Maps on the sampled data points and rank the eigenvectors based on their eigenvalues. The first Diffusion Maps eigenvector ϕ_0 with $\lambda_0 = 1$ is the *trivial* constant eigenvector therefore is discarded. In Figure 14 we color the first four (non-trivial) Diffusion Maps coordinates as functions on the data set. As it can be seen from Figure 14 ϕ_1 parametrize the same direction as the x coordinate; and is actually one to one with x. The next two Diffusion Maps coordinates, ϕ_2 and ϕ_3, parametrize also the same direction as ϕ_1 in different "frequencies", therefore they are harmonics of ϕ_1. However, ϕ_4 appears to parametrize the y direction (is one to-one) and thus is non-harmonic to ϕ_1.

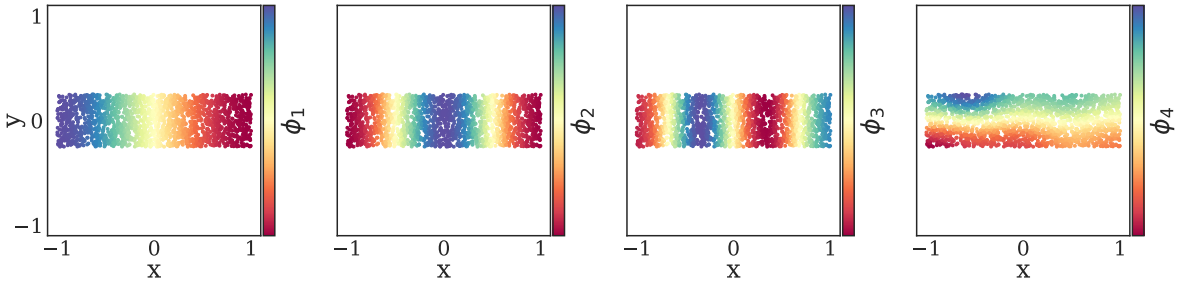


Figure 14: The first four (non-trivial) Diffusion Maps coordinates are colored as function on the data set.

In our illustrated toy example the two non-harmonic eigenvectors parametrize independently two of the original coordinates of the data set (similar to the Chafee-Infante example discussed in Section 4.1); this won't be the case in any given data set. In many cases the Diffusion Maps coordinates will parametrize combinations (linear or nonlinear) of the ambient data set's coordinates. In those cases a visual selection of the non-harmonic eigenvectors is also possible if the number of non-harmonic eigenvectors is smaller or equal to three. In these cases, the selection starts by plotting the leading non-trivial Diffusion Maps coordinate ϕ_1 versus the eigenvectors with lower values of eigenvalues. From Figure 15, ϕ_2 and ϕ_3 can be seen as functions of ϕ_1. On the contrary, the direction that ϕ_4 spans is independent of ϕ_1 since given a value of ϕ_1 there is a range of possible values for ϕ_4. In a data set where we might suspect also a third dimension constructing three dimensional visualizations of the candidate eigenvectors in terms of the first two non-harmonic eigenvectors could be performed.

In higher dimensions where visual selection of the non-harmonic eigenvectors is not feasible algorithms that compute the independent coordinates become possible. Carmeline et al. [13] proposed an automated algorithm that detects coordinates corresponding to harmonic eigenvectors based on local linear regression. For our examples illustrated in

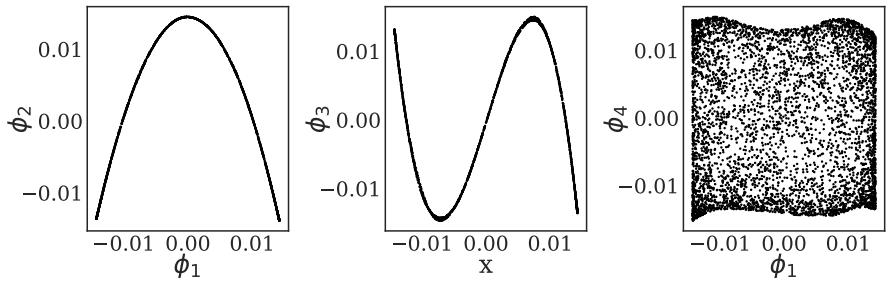


Figure 15: The first non-trivial coordinate ϕ_1 is plotted as against the eigenvectors ϕ_2, ϕ_3, ϕ_4 respectively, .

Sections 4.1 and 4.2 we selected the latent coordinates with visual inspection and as a sanity check by the local linear regression algorithm [13].

B.2.1 Nyström Extension

The Nyström method is a technique for finding numerical approximations to eigenfunction problems of the form [15]

$$\int_a^b K(\boldsymbol{x}_i, \boldsymbol{x}_j)\boldsymbol{\phi}(\boldsymbol{x}_j) = \lambda \boldsymbol{\phi}(\boldsymbol{x_i}). \tag{11}$$

In our work, the Nyström extension is mainly used to generate a Diffusion Map coordinate $\boldsymbol{\phi}_{new}$ for a sample point $\boldsymbol{x}_{new} \notin \boldsymbol{X}$. This interpolation function, $f : \boldsymbol{x}_{new} \to \boldsymbol{\phi}_{new}$, is being obtained by evaluating the kernel that was used during the dimensionality reduction step (and apply the same normalizations) on the new point $\boldsymbol{x}_{new}$. More precisely, the Euclidean distance between this new point, $\boldsymbol{x}_{new}$ and all the preexisting points in the data set $\mathbf{X}$ is computed as:

$$K(\boldsymbol{x}_{new}, \boldsymbol{x}_j) = \exp\left(-\frac{||\boldsymbol{x}_{new} - \boldsymbol{x}_j||^2}{2\epsilon}\right), \tag{12}$$

$$\widetilde{K}(\boldsymbol{x}_{new}, \boldsymbol{x}_j) = \frac{K(\boldsymbol{x}_{new}, \boldsymbol{x}_j)}{p(\boldsymbol{x}_{new})^\alpha \boldsymbol{p}(\boldsymbol{x}_j)^\alpha}, \tag{13}$$

with $p(\boldsymbol{x}_{new}) = \sum_{j=1}^N K(\boldsymbol{x}_{new}, \boldsymbol{x}_j)$ being just a scalar value and $\boldsymbol{p}(\boldsymbol{x}_i) = \sum_{j=1}^N K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ being a N-dimensional vector. Then,

$$A(\boldsymbol{x}_{new}, \boldsymbol{x}_i) = \frac{\widetilde{K}(\boldsymbol{x}_{new}, \boldsymbol{x}_j)}{\sum_{j=1}^N \widetilde{K}(\boldsymbol{x}_{new}, \boldsymbol{x}_j)}, \tag{14}$$

and the value of the reduced coordinate β^{th} is computed through

$$\phi_\beta(\boldsymbol{x}_{new}) = \frac{1}{\lambda_\beta}\sum_{i=1}^N A(\boldsymbol{x}_{new}, \boldsymbol{x}_i)\phi_\beta(x_i), \tag{15}$$

where $\phi_\beta(x_i)$ is the i-th component of the β-th eigenvector ($\boldsymbol{\phi}_\beta$) and λ_β is the β-th eigenvalue. Since the Nyström Extension provides a mapping from the ambient space coordinates $\boldsymbol{x}$ to the reduced coordinates $\boldsymbol{\phi}$, we can also get a closed form expression of derivative of the Nyström Extension. Symbolic differentiation of Equation (15) for $\alpha = 0$ gives:

$$\frac{\partial \phi_\beta}{\partial x_{new,k}} = \frac{1}{\lambda_\beta}\frac{\partial}{\partial x_{new,k}}\left(\sum_{i=1}^N A(\boldsymbol{x}_{new}, \boldsymbol{x}_j)\phi_{j,\beta}\right) = \frac{1}{\lambda_\beta}\frac{\sum_{i,j=1}^N K(\boldsymbol{x}_{new}, \boldsymbol{x}_j)\left(\frac{\partial}{\partial x_{new,k}}K(\boldsymbol{x}_{new}, \boldsymbol{x}_i)\right)(\phi_{i,\beta} - \phi_{j,\beta})}{\left(\sum_{i=1}^N K(\boldsymbol{x}_{new}, \boldsymbol{x}_i)\right)^2} \tag{16}$$

where,

$$\frac{\partial}{\partial x_{new,k}}K(\boldsymbol{x}_{new}, \boldsymbol{x}_i) = 2\epsilon^{-2}K(\boldsymbol{x}_{new}, \boldsymbol{x}_i)(x_{i,k} - x_{new,\beta}). \tag{17}$$

This derivative can then be used to map vectors in the ambient space (e.g., time derivatives) to the latent space.

C Equation-free projective integration

Similar to the aforementioned BF scheme is the *Equation-free* approach. In this paper we describe this particular scheme without using it for our examples; a more thorough explanation of this algorithm can be found in [25]. In the Equation Free approach, sampling of the entire manifold is not needed in advance. Sampling the manifold "locally" is enough in order to initiate this scheme. In this sampled data set Diffusion Maps can be performed in order to get a set of reduced latent coordinates. On those coordinates we need to estimate the gradients. This can be done by computing numerically the derivative with finite differences in terms of the reduced coordinates and perform *projective integration*. Alternatively, as suggested in [27] instead of estimating the gradients by using differences between nearby points $\phi(x_i), \phi(x_{i+1})$ one can estimate their values *after first projecting the points to the tangent space* of $\phi(x_i)$. In general a (projective) integration step leads to a new point in the reduced coordinates *outside* the sampled manifold. Given this out of sample point *lifting* in the ambient space coordinates can be achieved with Latent Harmonics (Algorithm 2). For this new coordinate in the ambient space, where closed form expressions of the dynamics are available, sampling again the manifold locally and repeating this procedure is possible.

D Additional Results

D.1 Chafee-Infante

A qualitative comparison is given between the lifted trajectories computed with the data-driven suggested approaches (BF and GT) and the integrated spectral ODEs. Figure 16 shows the comparison for the BF approach and Figure 17 shows the comparison for the GT approach.

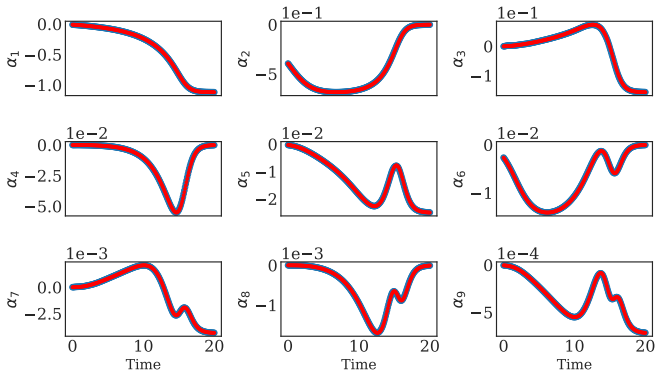


Figure 16: Comparison between the lifted with Latent Harmonics trajectory computed with the data driven integration scheme Back and Forth (blue trajectory) and the trajectory coming from the spectral odes (red trajectory).

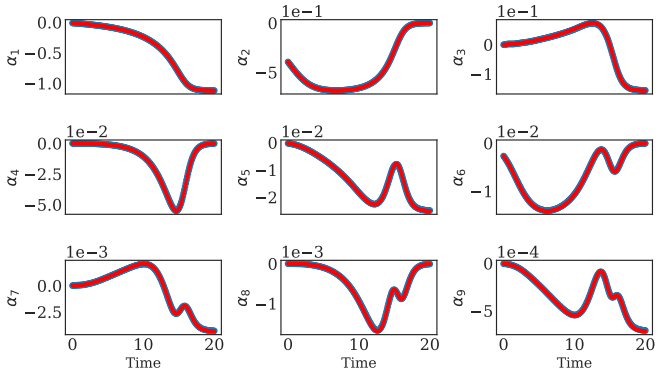


Figure 17: Comparison between the lifted with Latent Harmonics trajectory computed with the data-driven integration scheme Grid Tabulation (blue trajectory) and the trajectory coming from the spectral odes (red trajectory)

D.2 Chemical Kinetics

For the Chemical Kinetics example the qualitative comparison is given between the lifted trajectories computed with the data-driven suggested approaches (BF and GT) and the integrated system of ODEs. Figure 18 shows the comparison for the BF approach and Figure 19 shows the comparison for the GT approach.

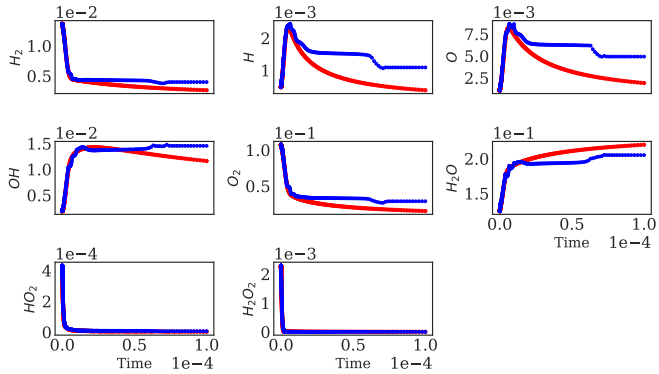


Figure 18: Comparison between the lifted with Latent Harmonics trajectory computed with the data-driven integration scheme Back and Forth (blue trajectory) and the trajectory computed from the full, unreduced equations of the kinetics (red trajectory).

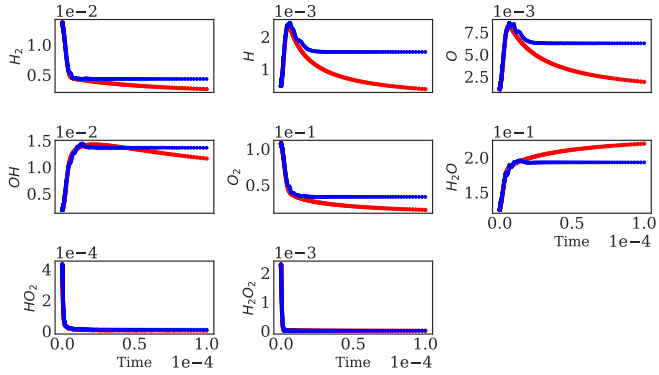


Figure 19: Comparison between the lifted with Latent Harmonics trajectory computed with the data-driven integration scheme Grid Tabulation (blue trajectory) and the trajectory computed from the full, unreduced equations of the mechanism (red trajectory). The reduced trajectory due to numerical errors is not able to continue since it *leaves* the manifold.

E Diffusion Maps and Latent Harmonics Parameters Selection

On Tables 1 and 2 we report the different parameters used for Diffusion Maps and also for Latent Harmonics for the two examples.

Table 1: Parameters used for Diffusion Maps and Latent Harmonics for the Chafee-Infante PDE

Method	ϵ	# eigenvalues	cut off	normalization (α)	Markov Matrix
Diffusion Maps	0.584	9	∞	0	Yes
Latent Harmonics	4.46×10^{-05}	300	∞	0	No

Table 2: Parameters used for Diffusion Maps and Latent Harmonics for the Combustion of Hydrogen.

Method	ϵ	# eigenvalues	cut off	normalization (α)	Markov Matrix
Diffusion Maps	18	9	∞	0	Yes
Latent Harmonics	4.11×10^{-05}	400	∞	0	No

F Errors

In this section we report the Mean Squared Errors for our interpolation schemes with Latent Harmonics. The reported errors in Tables 3 and 4 are computed on test data sets that contain 10% of the total sampled data points. Before the use of our interpolation scheme the functions of interest were scaled and translated in the range zero to one by using the *MinMaxScaler* from scipy. However, in Figures 20 and 21 the variables are being transformed back to their original values.

Chafee-Infante Example - Errors

Table 3: The Mean Squared Error of the functions of interest were computed on a test set for the Chafee-Infante PDE

	α_1	α_2	α_3	α_4	α_5
Mean Squared Error	1.41×10^{-9}	1.10×10^{-10}	2.39×10^{-6}	7.59×10^{-6}	2.68×10^{-6}
	α_6	α_7	α_8	α_9	α_{10}
Mean Squared Error	2.68×10^{-6}	1.65×10^{-5}	1.67×10^{-6}	4.37×10^{-6}	1.64×10^{-5}

	$\frac{d\phi_1}{dt}$	$\frac{d\phi_2}{dt}$
Mean Squared Error	3.40×10^{-7}	7.84×10^{-7}

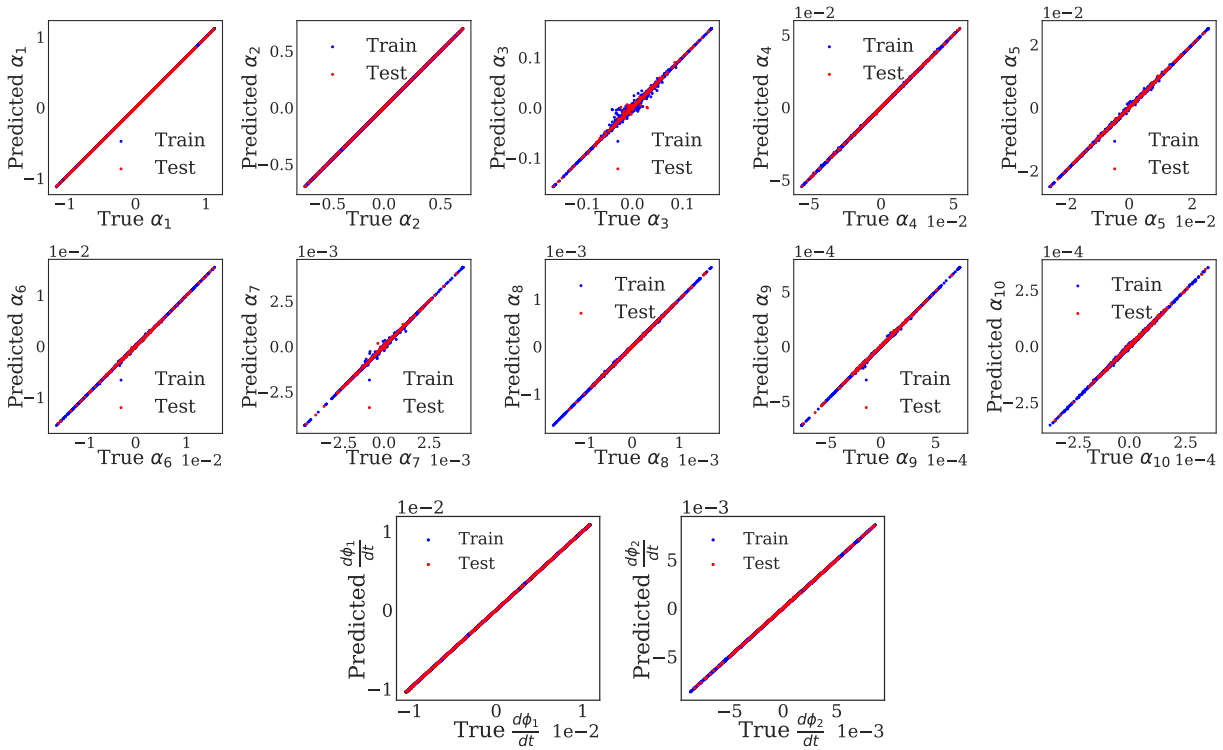


Figure 20: The true values of functions of interest plotted against the predicted with Latent Harmonics for the Chafee Infante PDE. Blue color points represent the training points and red points the test.

Chemical Kinetics - Errors

Table 4: The Mean Squared Error of the functions of interest were computed on a test set for the Chemical Kinetics example.

	H_2	H	O	OH
Mean Squared Error	1.51×10^{-5}	5.11×10^{-2}	8.41×10^{-2}	5.79×10^{-2}
	O_2	H_2O	HO_2	H_2O_2
Mean Squared Error	2.41×10^{-5}	2.24×10^{-2}	2.05×10^{-2}	1.17×10^{-1}
		$\frac{d\phi_1}{dt}$	$\frac{d\phi_2}{dt}$	
Mean Squared Error		9.07×10^{-4}	2.27×10^{-4}	

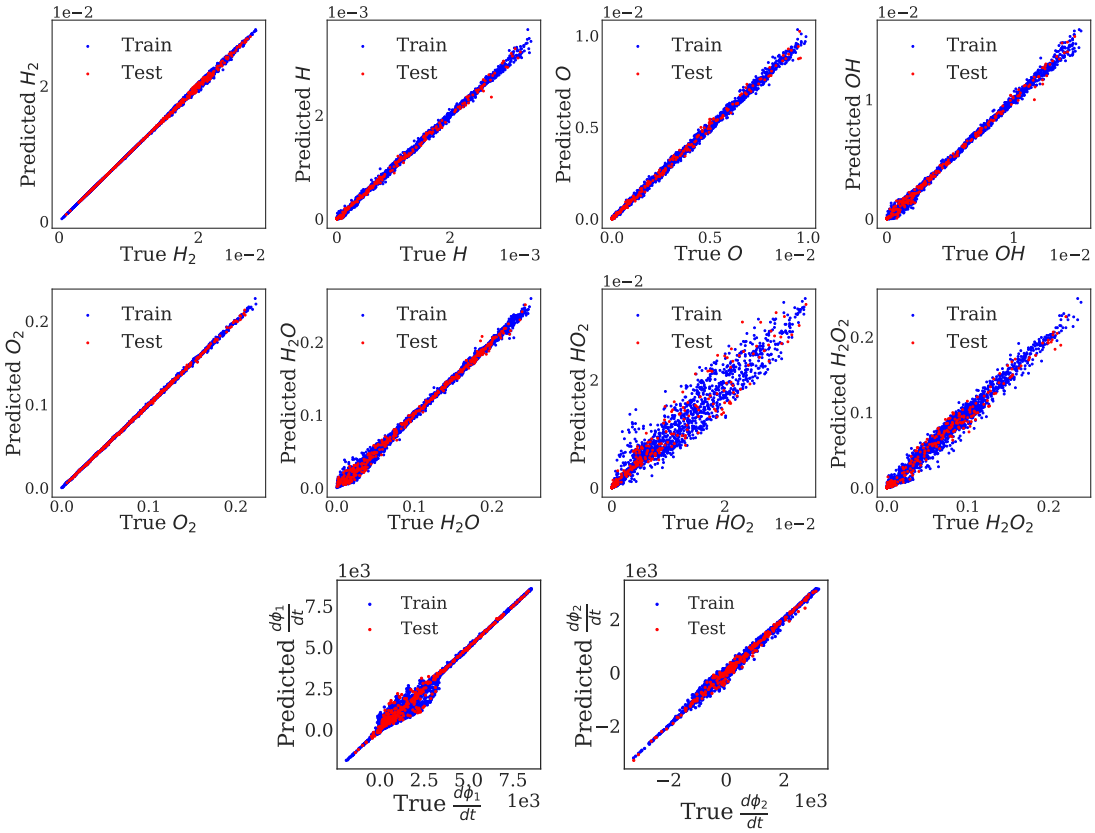


Figure 21: The true values of functions of interest plotted against the predicted with Latent Harmonics for the Kinetics example. Blue color points represent the training points and red points the test.

G Codes

All the codes that generate the figures and construct the suggested algorithms will be uploaded on a public repository on Gitlab upon acceptance.

View publication stats