Reduced Models in Chemical Kinetics via Nonlinear Data-Mining

Processes **2014**, *2*, 112-140; doi:10.3390/pr2010112

OPEN ACCESS

processes

ISSN 2227-9717

www.mdpi.com/journal/processes

Article

Reduced Models in Chemical Kinetics via Nonlinear Data-Mining

Eliodoro Chiavazzo [1,2], Charles W. Gear [1], Carmeline J. Dsilva [1], Neta Rabin [3] and Ioannis G. Kevrekidis [1,4,*]

[1] Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA; E-Mails: eliodoro.chiavazzo@polito.it (E.C.); wgear@princeton.edu (C.W.G.); cdsilva@princeton.edu (C.J.D.)

[2] Energy Department, Politecnico di Torino, Torino 10129, Italy

[3] Department of Exact Sciences, Afeka Tel-Aviv Academic College of Engineering, Tel-Aviv 69107, Israel; E-Mail: neta.rabin@gmail.com

[4] Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA

* Author to whom correspondence should be addressed; E-Mail: yannis@princeton.edu; Tel.: +1-609-258-2818; Fax: +1-609-258-0211.

Received: 23 September 2013; in revised form: 10 December 2013 / Accepted: 19 December 2013 / Published: 23 January 2014

Abstract: The adoption of detailed mechanisms for chemical kinetics often poses two types of severe challenges: First, the number of degrees of freedom is large; and second, the dynamics is characterized by widely disparate time scales. As a result, reactive flow solvers with detailed chemistry often become intractable even for large clusters of CPUs, especially when dealing with direct numerical simulation (DNS) of turbulent combustion problems. This has motivated the development of several techniques for reducing the complexity of such kinetics models, where, eventually, only a few variables are considered in the development of the simplified model. Unfortunately, no generally applicable *a priori* recipe for selecting suitable parameterizations of the reduced model is available, and the choice of slow variables often relies upon intuition and experience. We present an automated approach to this task, consisting of three main steps. First, the low dimensional manifold of slow motions is (approximately) sampled by brief simulations of the detailed model, starting from a rich enough ensemble of admissible initial conditions. Second, a *global* parametrization of the manifold is obtained through the Diffusion Map (DMAP) approach, which has recently emerged as a powerful tool in data analysis/machine learning. Finally, a simplified model

is constructed and solved *on the fly* in terms of the above reduced (slow) variables. Clearly, closing this latter model requires nontrivial interpolation calculations, enabling restriction (mapping from the full ambient space to the reduced one) and lifting (mapping from the reduced space to the ambient one). This is a key step in our approach, and a variety of interpolation schemes are reported and compared. The scope of the proposed procedure is presented and discussed by means of an illustrative combustion example.

Keywords: model reduction; data mining; combustion

1. Introduction

The solution of detailed models for chemical kinetics is often dramatically time consuming, owing to a large number of variables evolving in processes with a wide range of time and space scales. As a result, fluid dynamic flow solvers coupled with detailed chemistry still present a challenge, even for modern clusters of CPUs, especially when dealing with direct numerical simulation (DNS) of turbulent combustion systems. Here, a large number of governing equations for chemical species (a few hundred for mechanisms of standard hydrocarbon fuels) are to be solved at (typically) millions of distinct discretization points in the computational domain. This has motivated the development of a plethora of approaches aiming at reducing the computational complexity of such detailed combustion models, ideally by recasting them in terms of only a few new reduced variables. (see, e.g., [1] and references therein). The implementation of many of these techniques typically involves three successive steps. First, a large set of stiff ordinary differential equations (ODEs) is considered for modeling the temporal evolution of a spatially homogenous mixture of chemical species under specified stoichiometric and thermodynamic conditions (usually fixed total enthalpy and pressure for combustion in the low Mach regime). It is well known that, due to the presence of fast and slow dynamics, the above systems are characterized by *low dimensional manifolds* in the concentration space (or phase-space), where a typical solution trajectory is initially rapidly attracted towards the manifold, while afterwards, it proceeds to the thermodynamic equilibrium point, always remaining in close proximity to the manifold. Clearly, the presence of a manifold forces the ODEs state to visit mostly a low dimensional region of the entire phase-space, thus offering the premise for constructing a consistent reduced description of the process, which accurately retains the slow dynamics along the manifold, while neglecting the initial short transient towards the manifold. In a fluid dynamic simulation, stoichiometry and thermodynamic conditions may vary throughout the computational domain. Hence, when implementing reduction techniques, the second step consists of parameterizing and tabulating the manifolds arising in the homogeneous reactor for a variety of stoichiometric and thermodynamic conditions. Finally, as a third step, the fluid dynamic equations are reformulated in terms of the new variables, with the latter tables utilized to close the new reduced set of equations (see, e.g., [2]). It is worth stressing that the above description briefly outlines only one possible approach for coupling a model reduction method to a flow solver: the case where the low dimensional manifolds of the homogeneous problem are identified in advance *in the entire*

phase-space. For completeness, it is important to mention that, due to the rapidly increasing difficulty in storing and interpolating data in high dimensions, this approach remains viable in cases with *a few* reduced variables. As an alternative to this global method, techniques have been introduced for *locally* constructing the low dimensional manifold only in the (tiny) region of interest in the phase-space, as demanded by a reacting flow code during simulations [3–5]. Local constructions can certainly cope with higher dimensional manifolds. However, their usage seems computationally advantageous only in combination with efficient algorithms for adaptive tabulation, where data is computed when needed, stored and re-utilized if necessary (see, e.g., [6]).

In this work, we focus on the global construction and parameterization of slow invariant manifolds arising in the modeling of spatially homogeneous reactive mixtures. In particular, upon identification of the slow manifold, we propose a generally applicable methodology for selecting a suitable parameterization; we also investigate various interpolation/extrapolation schemes that need to be used in the solution of a reduced dynamical system expressed in terms of the variables learned. This latter step, and its integration with the previous two in an overall computer-assisted approach, constitute the methodological novelty of the paper.

The manuscript is organized as follows. In Section 2, Diffusion Maps are briefly reviewed. In Section 3 and the subsections therein, we discuss the computation of points on the manifold, their embedding in a reduced (here, two-dimensional) space, the formulation of a reduced set of equations and their solution through several interpolation/extension techniques. Results are reported and discussed in Section 4, where the proposed approach is applied to a reactive mixture of hydrogen and air at stoichiometric proportions with fixed enthalpy and pressure. The reader may prefer a quick glance at Section 4 before the detailed presentation of the procedure in Section 3. Finally, we conclude with a summary and brief discussion of open issues in Section 5.

2. Diffusion Maps

The Diffusion Map (DMAP) approach has emerged as a powerful tool in data analysis and dimension reduction [7–9]. In effect, it can be thought of as a nonlinear counterpart of Principal Component Analysis (PCA) [10] that can be used to search for a low-dimensional embedding of a high-dimensional point set $\{\mathbf{y}_1,...,\mathbf{y}_M\}$, if an embedding exists. For completeness, we present a simple description of the DMAP process. The points, $\mathbf{y}_i$, could exist in some n-dimensional Cartesian space (as they are in our combustion example) or they could be more abstract objects, such as images. What is important is that there exists a *dissimilarity function*, $d_{ij} = d_{ji}$, between any pair of points, $\mathbf{y}_i$ and $\mathbf{y}_j$, such that the dissimilarity is zero only if the points are identical (in those aspects that are important to the study) and gets larger the more dissimilar they are. Although, for points in $\Re^n$, an obvious choice for d_{ij} is the standard Euclidean distance, this is not necessarily the best option. For instance, a weighted Euclidean norm may be considered when different coordinates are characterized by disparate orders of magnitude. As discussed below, this is indeed the case encountered in many combustion problems, where the data are composition vectors in concentration space and major species (*i.e.*, reactants and products) are characterized by much higher concentrations compared to minor species (*i.e.*, radicals). From d_{ij},

a pairwise affinity function $w_{ij} = w(d_{ij})$ is computed, where $w(0) = 1$ and $w(d)$ is monotonically decreasing and non-negative for $d > 0$. A popular option is the heat kernel:

$$w_{ij}\left(d_{ij}\right) = \exp\left[-\left(\frac{d_{ij}}{\varepsilon}\right)^2\right] \tag{1}$$

The model parameter, ε, specifies the level below, whose points are considered similar, whereas points more distant than a small multiple of ε are, effectively, not linked directly. For this presentation, we will assume that d is a distance measure in (suitably scaled) Cartesian coordinates, so that each point, $\mathbf{y}_i$, is specified by its coordinates, $y_{i,\alpha}$ with $\alpha = 1, ..., n$, in n-dimensional space.

In the DMAP approach, starting from the $M \times M$ (not $n \times n$, as in PCA) symmetric matrix $\mathbf{W} = \{w_{ij}\}$, a Markov matrix, $\mathbf{K}$, is constructed through the row normalization:

$$\mathbf{K} = \mathbf{D}^{-1}\mathbf{W} \tag{2}$$

with the diagonal matrix, $\mathbf{D}$, collecting all the row sums of matrix $\mathbf{W}$. Owing to similarity with a symmetric matrix, $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, $\mathbf{K}$ has a complete set of real eigenvectors, $\{\phi_i\}$, and eigenvalues, $\{\lambda_i\}$. Moreover, a projection of the high-dimensional points $\{\mathbf{y}_1,...,\mathbf{y}_M\}$ into an m-dimensional space (hopefully, $m << n$) can be established through the components of m appropriately selected eigenvectors (not necessarily the m leading ones, as in PCA). Specifically, let the eigenvalues be sorted in decreasing order: $1 = \lambda_1 \geq |\lambda_2| \geq ... \geq |\lambda_M|$. The diffusion map, Ψ_t, is defined based on the right eigenvectors of K, $K\phi_l = \lambda_l\phi_l$, with $\phi_l = (\phi_{1,l}, ..., \phi_{M,l})$, for $t > 0$, as follows:

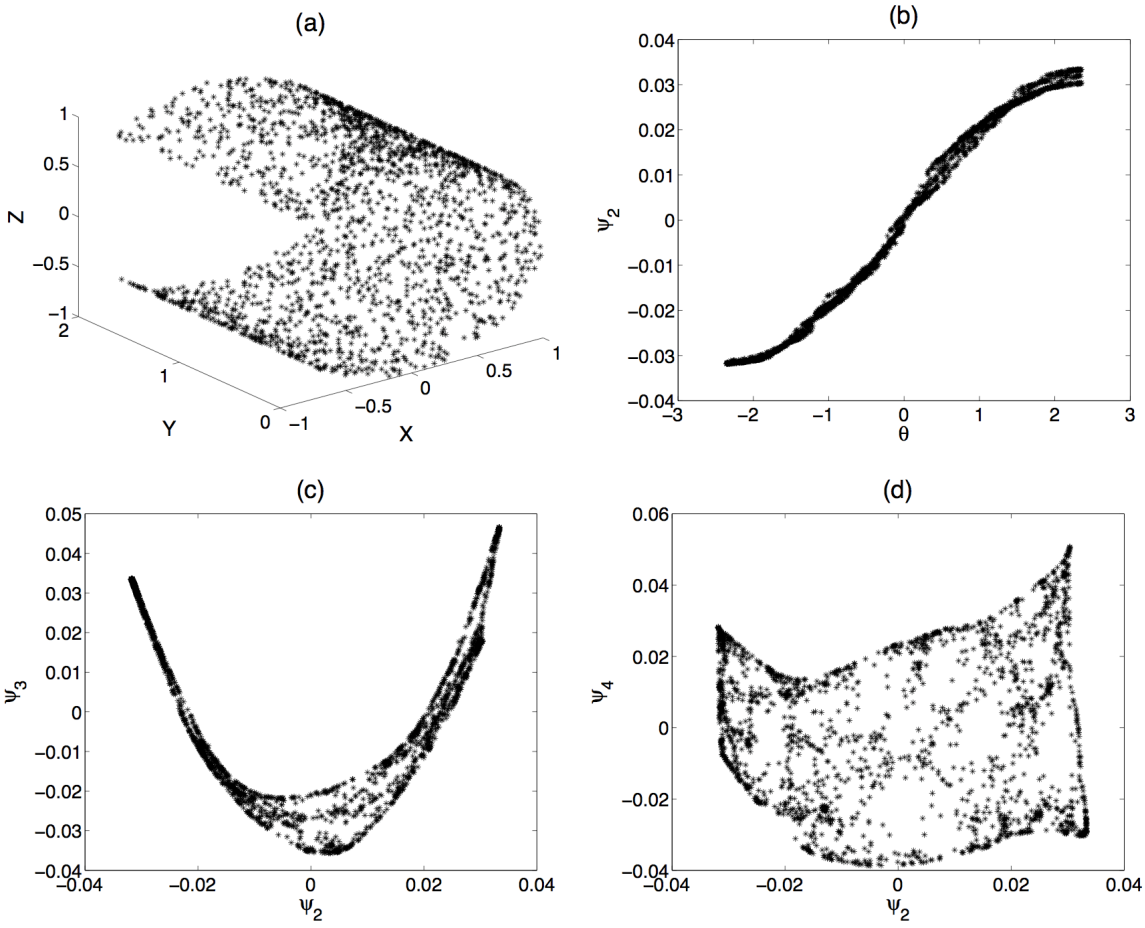
$$\Psi_t\left(\mathbf{y}_i\right) = \begin{pmatrix} \lambda_1^t\phi_{i,1} \\ \lambda_2^t\phi_{i,2} \\ \vdots \\ \lambda_M^t\phi_{i,M} \end{pmatrix} \tag{3}$$

and it assigns a vector of M new coordinates to each data point, $\mathbf{y}_i$. Here, K represents a (unit time) application of a diffusion process on the data points, so that t (in K^t) indicates the length of time over which this process is applied. Notice that all points have the same first coordinate in Equation (3), since ϕ_1 is proportional to the all-ones vector (with eigenvalue one). Notice that the diffusion map coordinates are time-dependent; using longer times in the diffusion process damps high frequency components, so that fewer coordinates suffice for an approximation of a given accuracy. However, in order to achieve a drastic dimension reduction, for a fixed threshold $0 < \delta < 1$, it is convenient to define a *truncated* diffusion map:

$$\Psi_t^{\delta}\left(\mathbf{y}_i\right) = \begin{pmatrix} \lambda_2^t\phi_{i,2} \\ \lambda_3^t\phi_{i,3} \\ \vdots \\ \lambda_{m+1}^t\phi_{i,m+1} \end{pmatrix} \tag{4}$$

where $m + 1$ is the largest integer for which $|\lambda_{m+1}|^t > \delta$. Below, we will consider only the eigenvector entries (*i.e.*, take $t = 0$) and will separately discuss using the eigenvalues (and their powers) to ignore noise.

Figure 1. Data manifold, dimensionality and independence of DMAP eigenvectors (**a**) 2000 uniformly random points initially placed in a unit square are stretched and wrapped around three-fourths of a cylinder; (**b**) the entry in the first non-trivial eigenvector of the Markov matrix, $\mathbf{K}$, *vs.* the first cylindrical coordinate, θ, for each data point; (**c**) entry in the second non-trivial eigenvector of $\mathbf{K}$ *vs.* the first one; the quasi-one-dimensionality of the plot implies strong eigenvector correlation; (**d**) entry in the third non-trivial eigenvector of $\mathbf{K}$ *vs.* the first one. The evident two-dimensional scatter implies that a new direction on the data manifold has been detected.



If the initial data points, $\{\mathbf{y}_1,...,\mathbf{y}_M\}$, are located on a (possibly non-linear) low dimensional manifold with dimension m, one might expect (by analogy to PCA) that a procedure exists to systematically select m diffusion map eigenvectors for embedding the data. If the points are fairly evenly distributed across the low-dimensional manifold, it is known that the principal directions of the manifold are spanned by some of the leading eigenvectors (*i.e.*, those corresponding to larger eigenvalues) of the DMAP operator [11,12]. In the illustrations below (Figure 1), the corresponding eigenvalues are approximately:

$$\lambda = 1 - \delta[k\pi d/L_\alpha]^2 \tag{5}$$

where $\delta \approx \exp(-d/\varepsilon^2)$, d is the typical spacing between neighbors and L_α is the length of the α-th principal direction. Here, $k = 1, 2, \cdots$ indicates the successive harmonics of the eigenvectors. (This approximation can be obtained by considering the regularly-spaced data case, assuming that ε

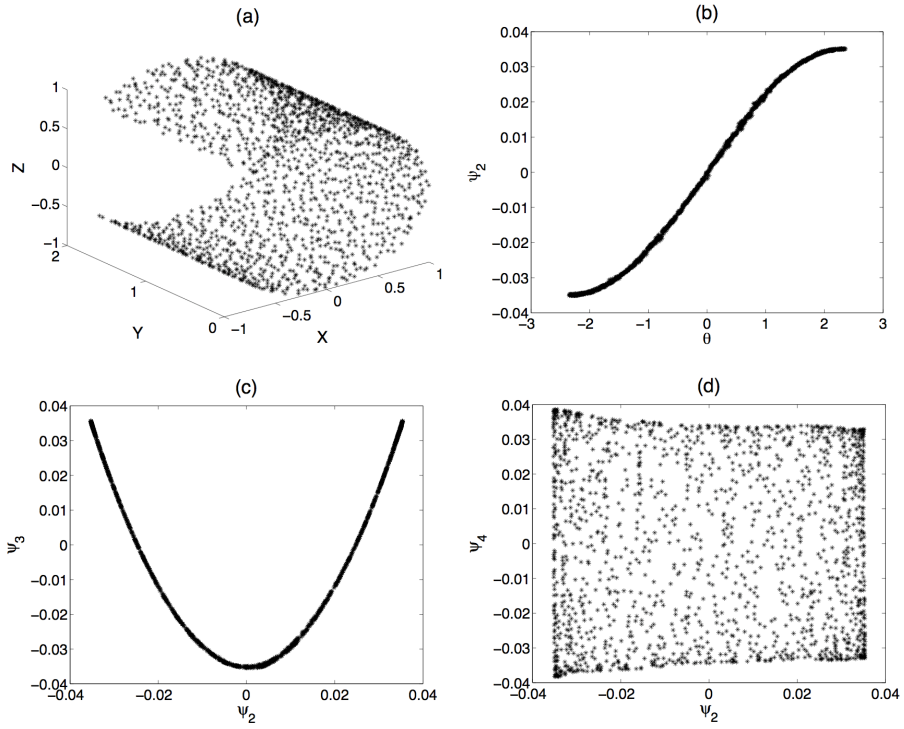
is comparable to d and that δ is small enough, so that higher powers can be ignored.) Section 2.1 below discusses how to ignore eigenvectors that are harmonics of previous ones by checking for dependence. Equation (5) provides a tool for deciding when to ignore the smaller eigenvalues. Suppose, for example, that we know that our data accuracy is approximately a fraction, γ, of the range of the data. This range roughly corresponds to the longest principal direction, say L_1. There is little point in considering manifold directions of the order of γL_1, since they are of the order of the errors in the data. Hence, by applying Equation (5), we should ignore any eigendirections whose eigenvalue is less than $1 - (1 - \lambda_2)\gamma^{-2}$, where λ_2 is the first non-trivial eigenvalue.

2.1. Issues in the Implementation of the Algorithm

While the formulas above appear to provide a simple recipe, a number of important, problem-dependent issues arise, having to do with the sampling of the points to be analyzed, the choice of the parameter, ε, *etc*.; we now discuss these issues through illustrative caricatures. Consider 2000 uniformly random points initially placed in a unit square, then stretched and wrapped around three fourths of a cylinder of radius one and length two (see Figure 1a). In Figure 1b, the first non-trivial eigenvector, ψ_2, is reported against the first cylindrical coordinate, θ: the i-th component of this eigenvector is plotted against the θ angle of the *i*-th point. The clearly apparent one-dimensional nature of the plot confirms that ψ_2 parametrizes this principal geometric direction. However, a plot of ψ_3, the eigenvector corresponding to the next leading eigenvalue, against ψ_2 clearly shows a strong correlation: ψ_3 is not representative of a new, *independent* direction on the data manifold. In Figure 1d, the two-dimensional scatter of the plot of the entries of the fourth eigenvector *vs*. the entries of the second one indicates independence between ψ_2 and ψ_4; ψ_4 does represent a new, independent direction along the data manifold and becomes our second embedding coordinate. Visually testing independence between two DMAP eigenvectors is relatively easy: we can agree that Figure 1b,c appear one-dimensional and Figure 1d appears two-dimensional. However, testing independence in higher dimensions (for subsequent DMAP eigenvectors) becomes quickly visually impossible and even computationally nontrivial. Subsequent eigenvectors should be plotted against previous important ones (in the example just considered, against ψ_2 and ψ_4), and the dimensionality of the plot should be assessed; this is still visually doable for, say, ψ_5, and the plot appears as a 2D surface in 3D: ψ_5 is not a new data coordinate. Beyond visual assessment (and in higher dimensions), one can use the sorted edge-length algorithm for dimensionality assessment: a log-log plot of the graph edge-length *vs*. edge number is constructed, with the manifold dimension being the slope in the middle part of the plot. Algorithms for detecting the dimension of attractors in chaotic dynamical systems can also find use here [13,14].

The irregularity of sample points can be easily seen to lead to problems in this simple example. Consider, however, an additional case where 1600 points are initially randomly placed in each of the 40 by 40 array of small squares, forming the unit square, and, afterwards, bent around the cylinder (Figure 2a). As clearly visible in Figure 2b–d, these time dependencies between eigenvectors are very well defined.

Figure 2. More on the effect of data sampling: noise. (**a**) One thousand six hundred points are initially randomly placed in each of the 40 by 40 array of small squares, forming the unit square, and, afterwards, bent around a cylinder; (**b**) entry in the first non-trivial eigenvector of the Markov matrix, $\mathbf{K}$, *vs.* the first cylindrical coordinate, θ, for each data point; (**c**) entry in the second non-trivial eigenvector of $\mathbf{K}$ *vs.* the first one; (**d**) entry in the third non-trivial eigenvector of $\mathbf{K}$ *vs.* the first one.



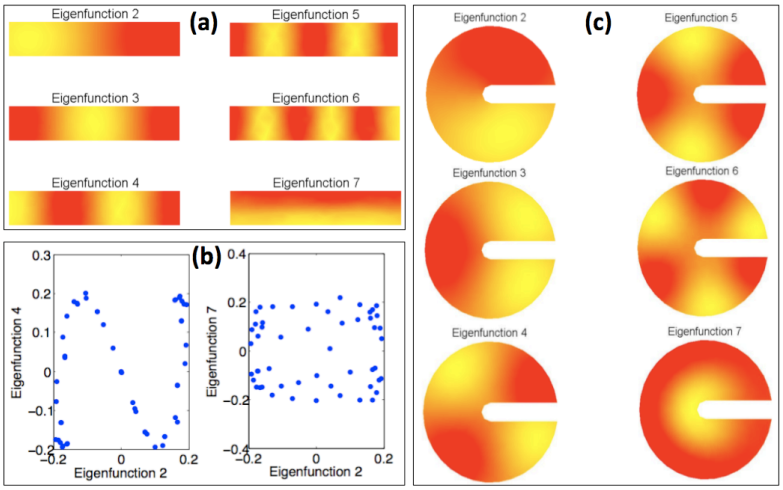
While the first non-trivial eigenvector, ψ_2, always characterizes the principal direction on the manifold, no general recipe can be formulated for an *a priori* identification of the subsequent uncorrelated eigenvectors parameterizing other dimensions. We have already seen that eigenvectors in Equation (3) are often dependent; this implies that they do not encode new directions along the data manifold; in this sense, they are redundant for our embedding. In order to obtain more insight into eigenvector dependency (and, in other words, in how diffusion is linked with manifold parametrization), consider, as our domain of interest, a narrow two-dimensional stripe, or, in our case, data points densely sampled from it. Figure 3a reports the solution to the discretized (through the finite element method, FEM) eigenvalue problem $\nabla^2\phi = \lambda\phi$ with Neumann boundary conditions. The first non-trivial eigenfunction is analytically given by $\cos(\bar{x})$, where $\bar{x}$ denotes the horizontal space direction, and is very well approximated by the FEM numerics; the point to notice is that $cos(\bar{x})$ is one-to-one with $\bar{x}$ between zero and 2π; so, the first nontrivial diffusion eigenvector parameterizes one manifold direction (the $\bar{x}$). Several subsequent eigenfunctions still correlate with the $\bar{x}$ direction: they are simply higher harmonics ($\cos(2\bar{x})$, $\cos(3\bar{x})$,...). We have to go as high as the seventh eigenfunction (which analytically is $\cos(\bar{y})$) to find something that is one-to-one with the second, independent, vertical direction, $\bar{y}$ (see Figure 3b, where the first non-trivial eigenfunction is plotted against both the fourth and seventh eigenfunction at scattered locations). A more complex two-dimensional geometry is considered in Figure 3c. Similarly to the above example, the first non-trivial eigenfunction parameterizes one of the manifold "principal

dimensions" (the angular coordinate), while the next (seventh) uncorrelated eigenfunction can be used to parameterize the other relevant (radial) coordinate (it is just an accident that we had to go to the *seventh* eigenfunction in both cases). In practical applications, only a discrete set of sample points on the manifold in question is available as an input. Starting from those points, the Diffusion Maps create a graph, where the points are the graph nodes and the edges are weighted on the basis of point distances, as described above. Noticing that the (negatively defined) normalized graph Laplacian, L, is given by [15]:

$$\mathbf{L} = \mathbf{D}^{-1}\mathbf{W} - \mathbf{I} \tag{6}$$

with I being the $M \times M$ identity matrix, we immediately recognize the link between the eigenvalue problem in Figure 3 and mapping Equation (3) based on the spectrum of Markov matrix Equation (2).

Figure 3. The analogy between traditional diffusion on domains and diffusion on graphs from sampled data. (**a**) The solution to the finite element method (FEM) formulation of the PDE (partial differential equation) eigenvalue problem $\nabla^2\phi = \lambda\phi$ with no flux boundary conditions is reported for a narrow two-dimensional rectangular stripe. The second and seventh eigenfunctions are found to be uncorrelated and suitable to parametrize the two relevant dimensions of the manifold; (**b**) Entries in the first non-trivial eigenfunction of the problem in figure (a) *vs.* entries in the fourth eigenfunction (sampled at scattered locations of the computational domain) reveals a strong correlation between those two functions: the fourth eigenvector (which we know corresponds to the third harmonic, $cos(3\bar{x})$) does not encode a new direction on the data manifold. **Right-hand side**: Entries in the first non-trivial eigenfunction of the problem in (a) *vs.* entries in the seventh eigenfunction (at the same scattered locations) confirms that the seventh eigenvector (which we know corresponds to $cos(\bar{y})$, encodes a new, second direction; (**c**) A different domain, but the same premise: The solution to the finite element method (FEM) formulation of the PDE (partial differential equation) eigenvalue problem $\nabla^2\phi = \lambda\phi$ with no flux boundary conditions is reported for a two-dimensional manifold with a complicated boundary. The second and seventh eigenfunctions are found to be uncorrelated and suitable to parametrize the two relevant dimensions of the manifold (an "angular" and a "radial" one).



Diffusion on this graph (*i.e.*, obtaining the spectrum of the graph Laplacian) approximates, at the appropriate limit [7], the usual diffusion in the original domain; it provides an alternative, different from our FEM, irregular mesh-discretization of the Laplace equation eigenproblem in the original domain, and asymptotically recovers the spectrum of the Laplace operator there.

3. The Proposed Approach

We demonstrate the feasibility of constructing reduced kinetics models for combustion applications, by extracting the slow dynamics on a manifold globally parameterized by a truncated diffusion map. We focus on spatially homogeneous reactive mixtures of ideal gases under fixed total enthalpy H and pressure P. Such a set-up is relevant for building up tables to be used in reactive flow solvers in the low Mach number regime. In such systems, a complex reaction occurs with n chemical species $\{A_1,...,A_n\}$ and q chemical elements involved in a (typically) large number, r, of elementary steps:

$$\sum_{p=1}^{n} \iota_{sp} A_p \rightleftarrows \sum_{p=1}^{n} \kappa_{sp} A_p, \quad s = 1, ..., r \tag{7}$$

where ι_{sp} and κ_{sp} represent the stoichiometric coefficients of the p-th species in the s-th step. Time evolution of chemical species can be modeled by a system of ordinary differential equations (ODEs) cast in the general form:

$$\frac{dy}{dt} = \sum_{s=1}^{r} \gamma_s \Omega_s \left(y, T\right) \tag{8}$$

with $\gamma_s = (\kappa_{s1} - \iota_{s1}, ..., \kappa_{sn} - \iota_{sn})$, while the reaction rate function, Ω_s, is usually expressed in terms of the concentration vector, y, by mass action laws and Arrhenius dependence on the temperature, T. Clearly, a constraint on a thermodynamic potential is required in order to close the system Equation (8), thus providing an additional equation for temperature. Below (without loss of generality), we consider reactions under fixed total enthalpy, H.

The first step of our method consists in the identification of a discrete set of states lying in a neighborhood of the low-dimensional attracting manifold. While many possible constructions have been suggested in the literature (see, e.g., [1,4,5,16]), here, in the spirit of the *equation-free* approach [17,18], we assume that we have no access to the analytical form of the vector field; instead, we only have access to a "black box" subroutine that evaluates the rates, $f(y)$, and, when incorporated in a numerical initial value solver, can provide simulation results.

3.1. Data Collection

To start the procedure, we need an ensemble of representative data points on (close to) the manifold we wish to parametrize. To ensure good sampling, our ensemble of points comes from integrating Equation (8) starting from a (rich enough) set of random states within the admissible phase-space (a convex polytope defined by elemental conservation constraints and concentration positivity). After sufficient time to approach a neighborhood of the manifold, samples are collected from each such trajectory. As a result, a set of points $\{\mathbf{y}_i, i = 1, ..., M\}$ in $\Re^n$ (hopefully dense enough within the region of interest) becomes available for defining the manifold. To construct the required initial conditions,

we first search for all vertices of the convex polytope defined by a set of equalities and inequalities as follows:

$$\begin{aligned} &\sum_{\alpha=1}^{n} y_\alpha c_{\alpha\beta}/\bar{W}_\alpha = \sum_{\alpha=1}^{n} y_\alpha^{eq} c_{\alpha\beta}/\bar{W}_\alpha, \quad \forall\beta = 1, ..., q \\ &y_\alpha > 0, \quad \forall\alpha = 1, ..., n \end{aligned} \tag{9}$$

where $c_{\alpha\beta}$ and $\bar{W}_\alpha$ denote the number of atoms of the β-th element in the species α and the molecular weight of species α, respectively, while the state vector $y = (y_1, ..., y_n)$ expresses species concentration in terms of mass fractions. The selection of random initial conditions is performed by convex linear combinations of the v polytope vertices, $\{y_i^{pol}\}$:

$$y^{in} = \sum_{i=1}^{v} \bar{w}_i y_i^{pol} \tag{10}$$

with $\{\bar{w}_i\}$ being a set of v random weights, such as $\sum_{i=1}^{v} \bar{w}_i = 1$. Clearly, owing to convexity, Equation (10) always provides states within the admissible space. In combustion applications, the phase-space region of interest goes from the fresh mixture conditions to the thermodynamic equilibrium, y^{eq}; hence, in Equation (10), we consider a subset of the polytope vertices, $\{y_i^{pol}\}$, based on their vicinity to the mean point of the mixing line connecting the fresh mixture point to equilibrium. Due to a disparity of the magnitudes of species concentrations, d_{ij} is taken as the Euclidean distance between properly rescaled points y_i' and y_j', with $y_i' = \mathbf{R}y_i$ using the fixed diagonal matrix $\mathbf{R} = \{r_{\beta\beta}\}$, $r_{\beta\beta} = 1/max(y_\beta)$. Here, $max(y_\beta)$ represents the largest β-th coordinate among all available samples. Below, distances in the concentration space are intended in this sense. It is worth noticing that, upon the choice of v random numbers $\{\tilde{w}_i, i = 1, ..., v\}$ uniformly distributed over the range $0 \leq \tilde{w}_i \leq 1$, weights might be straightforwardly obtained by normalization: $\bar{w}_i = \tilde{w}_i / \sum_{j=1}^{v} \tilde{w}_j$. However, such an approach leads to poor sampling in the vicinity of the polytope edges and, at the same time, to oversampling within its interior. Therefore, in order to achieve a more uniform sampling in the whole phase-space region of interest, the weights are chosen as follows:

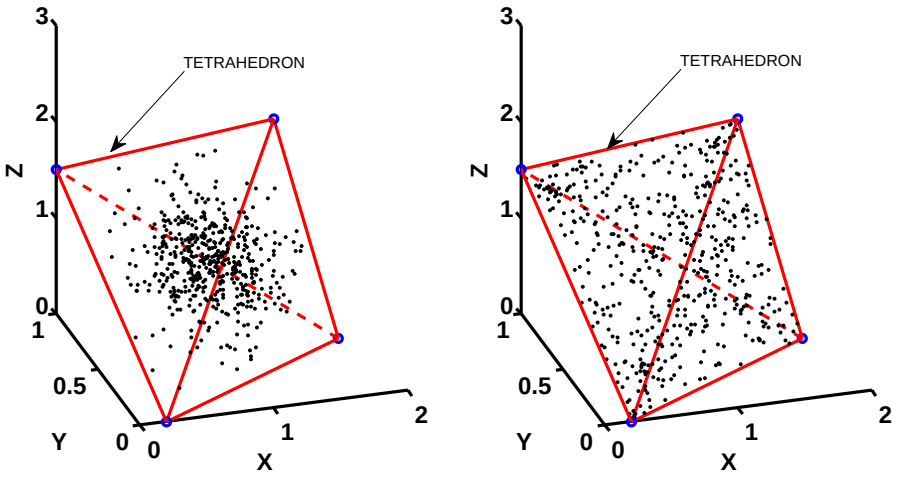
$$\tilde{w}_i = [-\ln(z_i)]^p, \quad \bar{w}_i = \tilde{w}_i / \sum_{j=1}^{v} \tilde{w}_j, \quad i = 1, ..., v \tag{11}$$

with z_i representing random values uniformly distributed within the interval $0 \leq z_i \leq 1$ and $1 \leq p \leq 2$ a free parameter (see also Figure 4).

It is worth stressing that Equation (9) always defines a convex polytope; hence, all samples computed by Equation (10) necessarily belong to the admissible domain, and no sample rejection is needed at any time. Clearly, such a procedure is only one possible approach for uniformly sampling a high-dimensional convex polytope. While a detailed study of alternative methods are beyond the scope of this work, we remind that other algorithms are also possible, such as Gibbs or hit-and-run sampling [19,20]. Trajectories starting at the random initial conditions, y^{in}, computed by Equation (10) are evolved for τ_f, after which, samples are collected as they proceed towards the equilibrium point, y^{eq}. Samples from the same trajectory are retained if their distance exceeds a fixed threshold. We would like the sample to be as uniform as possible in the original space (which we will call the ambient space), because doing so yields a better parameterization with Diffusion Maps [21,22]. However, such a condition is not naturally fulfilled by samples of time integration: the trajectories (hence, also, our sampled points)

often show a tendency to gather in narrow regions (especially close to the equilibrium point, governed by the eigenvalue differences in the linearized dynamics). Hence, we also performed an *a posteriori* data filtering (subsampling), where neighbors within a minimum distance, d^{min}, are removed.

Figure 4. On sampling initial conditions in a convex polytope in $\Re^3$ with vertices $A = (1.8, 0.5, 0)$, $B = (1, 0, 3)$, $C = (0, 1, 1.5)$ and $D = (0.2, 0, 0)$. (**Left**) Five hundred points are generated by Equation (10) with uniformly random values $0 \leq \tilde{w}_i \leq 1$ and $\bar{w}_i = \tilde{w}_i / \sum_{j=1}^{4} \tilde{w}_j$; notice the poor sampling close to the boundaries; (**Right**) Five hundred points are generated by Equation (10) with uniformly random values $0 \leq z_i \leq 1$, $\tilde{w}_i = [-\ln(z_i)]^{1.5}$ (*i.e.*, $p = 1.5$) and $\bar{w}_i = \tilde{w}_i / \sum_{j=1}^{4} \tilde{w}_j$. The latter approach generates a more uniform sampling of the polytope interior.



The Diffusion Map approach is performed as outlined in Section 2, whereas the parameter, ε, in Equation (1) can be chosen as a multiple of the quantity, $\max_j \min_{i \neq j} d_{ij}$ [21–23]. A better choice for ε is to make it a multiple of what we will call the *critical diffusion distance*: the maximum edge length such that, if all edges of at least that length are deleted in the distance graph, the graph becomes disconnected. The reason this distance is important is that if ε is much smaller than this, the diffusion map will find disjoint sets.

The Diffusion Map process provides a mapping from each point, y_i, in the ambient space to the reduced representation $u_i = [\psi_{i,2}, \psi_{i,3}, ..., \psi_{i,m+1}]^T$ in the m-dimensional reduced space. We will refer to this as the u-space. The manifold, $\mathbf{\Omega}$, in the ambient (y) space is known only by the finite set of points, $\{\mathbf{y}_i\}$, on $\mathbf{\Omega}$, and its mapping to u-space is known only up to the mapping of that set of points to the corresponding set, $\{\mathbf{u}_i\}$. Clearly, we can use any interpolation technique to compute y for any other value of u. Let us call this $y = \Theta(u)$. If u is in an m-dimensional space, this mapping defines an m-dimensional manifold in y-space, $\mathbf{\Omega}_c$. If we chose an interpolation method, such that $y_i = \Theta(u_i)$, then $\mathbf{\Omega}_c$ contains the original set, $\{\mathbf{y}_i\}$, but is an approximation to the slow manifold, $\mathbf{\Omega}$.

We will also assume that we can construct a mapping in the other direction, $u = \psi(y)$, where $u_i = \psi(y_i)$ for all i. Finally, in the third step, we need to conceptually recast Equation (8), which has the form $dy/dt = f(y)$, into the reduced space as:

$$\frac{du}{dt} = g(u) \tag{12}$$

In other words, given a value of u, we need a computational method to evaluate $g(u)$; yet, all we have available is a method to compute $f(y)$. To do this, we have to execute the following three sub-steps:

1. Compute the y on $\mathbf{\Omega}_c$ corresponding to the current u (using whatever form of interpolation we chose earlier);
2. Compute $dy/dt = f(y)$;
3. Compute the equivalent du/dt.

Since $\mathbf{\Omega}_c$ is only an approximation to $\mathbf{\Omega}$, it is highly unlikely that dy/dt lies in the tangent plane of $\mathbf{\Omega}_c$ at the point, y. (If it did, the problem of computing an equivalent du/dt would be straightforward.) Two possible solutions to this dilemma are: (i) project dy/dt onto the tangent plane; or (ii) extend the mapping $u = \psi(y)$ to include a neighborhood of $\mathbf{\Omega}_c$ (a many-to-one map). If we do the latter, we can write:

$$\frac{du}{dt} = \frac{\partial \psi}{\partial y}\frac{dy}{dt} \tag{13}$$

These two approaches are really the same, since a local extension of ψ to a neighborhood on $\mathbf{\Omega}_c$ implies a local foliation, and Equation (13) is simply a projection along that foliation. If an orthogonal projection is used, we simply write:

$$\frac{du}{dt} = \left(\mathbf{J}^T\mathbf{J}\right)^{-1}\mathbf{J}\frac{dy}{dt} \tag{14}$$

where $\mathbf{J} = \frac{\partial \Theta}{\partial u}$ and $\psi(y)$ is possibly needed only for initializing Equation (13) in the case initial conditions are available in the ambient space.

3.2. Interpolation/Extension Schemes

In the following, we will review a number of possible extension (in effect, interpolation/ extrapolation) schemes that might be adopted for solving system Equation (13) on a learned low dimensional manifold.

3.2.1. Nyström Extension

An established procedure for obtaining the α-th DMAP coordinate, ψ_α, at an arbitrary state, $y \in \Re^n$, is the popular Nyström extension [24]:

$$\begin{aligned} \psi_\alpha &= {\lambda_\alpha}^{-1}\sum_{i=1}^{M} k\left(y_i, y\right)\psi_{i,\alpha} \\ k\left(y_i, y\right) &= \left(\sum_{j=1}^{M} w\left(y_j, y\right)\right)^{-1} w\left(y_i, y\right) \\ w\left(y_i, y\right) &= e^{-\left(d_i/\varepsilon\right)^2},\ d_i = \left\|y_i' - y'\right\| \end{aligned} \tag{15}$$

where λ_α and $(\psi_{1,\alpha}, ..., \psi_{M,\alpha})$ are the α-th eigenvalue and eigenvector of the Markov matrix, $\mathbf{K}$, respectively. For the combustion case below, the d_i denote the Euclidean distances between rescaled points ($y' = \mathbf{R}y$, $y_i' = \mathbf{R}y_i$). The Jacobian matrix at the right-hand side of Equation (13) can be obtained by differentiation of Equation (15) as follows [22,25]:

$$\begin{aligned} \frac{\partial \psi_\alpha}{\partial y_\beta} &= \lambda_\alpha^{-1}\left(\sum_{i=1}^{M} w\left(y_i, y\right)\right)^{-2}\sum_{i,j=1}^{M} w\left(y_j, y\right)\frac{\partial w(y_i, y)}{\partial y_\beta}\left[\psi_{i,\alpha} - \psi_{j,\alpha}\right] \\ \frac{\partial w(y_i, y)}{\partial y_\beta} &= 2\varepsilon^{-2} r_{\beta\beta}^2 w\left(y_i, y\right)\left(y_{i,\beta} - y_\beta\right) \end{aligned} \tag{16}$$

where, in the case of point rescaling, $r_{\beta\beta}$ is computed as specified above, otherwise $r_{\beta\beta} = 1, \forall\beta$. The Nyström extension can be utilized for implementing the restriction operator, as well as for computing its Jacobian matrix.

3.2.2. Radial Basis Functions

Both lifting and restriction operators may be also obtained by local interpolation through radial basis functions. Let u be a new state in the reduced space; the corresponding point in the full space $y = \Theta(u)$ can be generally expressed by the following summation:

$$y_\beta = \sum_{i=1}^{nn} \alpha_{i,\beta} \bar{\phi}\left(\|u - u_i\|\right), \quad \beta = 1, ..., n \tag{17}$$

over the nn nearest neighbors of u with the radial function, $\bar{\phi}(\bullet)$, only depending on a distance $\|\bullet\|$. In this work, we focus on the following special form of Equation (17):

$$y_\beta = \sum_{i=1}^{nn} \alpha_{i,\beta} \|u - u_i\|^p, \quad \beta = 1, ..., n \tag{18}$$

where p is an odd integer, while $\|\bullet\|$ denotes the usual Euclidean distance in the reduced space. The coefficients, $\alpha_{i,\beta}$, are computed as:

$$\left[\alpha_{1,\beta}, ..., \alpha_{nn,\beta}\right]^T = \mathbf{\Lambda}^{-1} \left[y_{1,\beta}, ..., y_{nn,\beta}\right]^T, \quad \mathbf{\Lambda}(i,j) = \|u_i - u_j\|^p, \quad i,j = 1, ..., nn \tag{19}$$

Similarly, restriction can be expressed in the form:

$$\psi_\beta = \sum_{i=1}^{nn} \alpha_{i,\beta} \|y' - y'_i\|^p, \quad j = 1, ..., k \tag{20}$$

where data in the full space have been possibly rescaled ($y' = \mathbf{R}y$, $y'_i = \mathbf{R}y_i$). The Jacobian matrix at the right-hand side of Equation (13) can be obtained by differentiation of Equation (20) as follows:

$$\frac{\partial \psi_\beta}{\partial y_\gamma} = p r_{\gamma\gamma}^2 \sum_{i=1}^{nn} \alpha_{i,\beta} \left[\sum_{\omega=1}^{n} \left(y'_\omega - y'_{i,\omega}\right)^2\right]^{\frac{p}{2}-1} \left[y_\gamma - y_{i,\gamma}\right] \tag{21}$$

3.2.3. Kriging

Kriging typically refers to a number of sophisticated interpolation techniques originally developed for geostatistics applications. Provided a function, f, known on scattered data, its extension to a new point is performed via a weighted linear combination of the values of f at known locations. A noticeable feature of Kriging is that weights may depend on both distances *and correlations* between the available samples. In fact, one possible disadvantage of schemes only based on the quantities $\|\bullet\|$ (e.g., radial basis functions) is that samples at a given distance from the location where an estimate is needed are all equally treated. In contrast, Kriging offers the possibility of performing a weighting that accounts for redundancy (*i.e.*, sample clustering) and even sample orientation. This is done by choosing an analytical model that best fits the experimental semivariogram of the dataset. More details on Kriging can be found in [26]. In this work, both interpolated points and derivatives are computed by the readily available Matlab toolbox, DACE [27].

3.2.4. Laplacian Pyramids

Laplacian Pyramids are a multi-scale extension algorithm, where a function only known at M (scattered) sample points can be estimated at a new location. Based on a chosen kernel and pair-wise distances between samples, this algorithm aims at generating a sequence of approximations with different resolutions at each subsequent level, l [28]. Let y be a new point in the full space. The α-th coordinate of the corresponding state in the reduced space u is evaluated in a multi-scale fashion as follows: $\psi_\alpha \approx s_\alpha^{(0)} + s_\alpha^{(1)} + s_\alpha^{(2)} + ...$, with:

$$\begin{aligned} s_\alpha^{(0)} &= \sum_{i=1}^{M} k^{(0)}\left(y_i, y\right)\psi_{i,\alpha} \qquad \text{for level } \; l = 0 \\ s_\alpha^{(l)} &= \sum_{i=1}^{M} k^{(l)}\left(y_i, y\right) d_{i,\alpha}^{(l)} \qquad \text{otherwise} \end{aligned} \tag{22}$$

and the differences:

$$\begin{aligned} d_\alpha^{(1)} &= \psi_\alpha - s_\alpha^{(0)} \qquad \text{for level } \; l = 1 \\ d_\alpha^{(l)} &= \psi_\alpha - \sum_{i=0}^{l-1} s_\alpha^{(i)} \quad \text{otherwise} \end{aligned} \tag{23}$$

are updated at each level, l. The functions, k^l, in Equation (22) are:

$$\begin{aligned} k^{(l)}\left(y_i, y\right) &= q_l^{-1} w^{(l)}\left(y_i, y\right) \\ q_l &= \sum_j w^{(l)}\left(y_j, y\right) \\ w^{(l)}\left(y_i, y\right) &= \exp\left[-\left\|y_i' - y'\right\|^2 / \sigma_l\right] \end{aligned} \tag{24}$$

In Equation (24), a Gaussian kernel is chosen, where the parameter $\sigma_l = \sigma_0 / 2^l$ decreases with the level, l, σ_0 is the fixed coarsest scale, while y_i' and y' denote the rescaled states ($y' = \mathbf{R}y$, $y_i' = \mathbf{R}y_i$). A maximum admissible error can be set *a priori*, and the values, $s_\alpha^{(l)}$, are only computed up to the finest level, where: $\left\|\psi_\alpha - \sum_k s_\alpha^{(k)}\right\| < \text{err}$. The use of Laplacian Pyramids for constructing a lifting operator, $y_\alpha \approx s_\alpha^{(0)} + s_\alpha^{(1)} + s_\alpha^{(2)} + ...$, is straightforward and only requires the substitution of $\psi_{i,\alpha}$ with $y_{i,\alpha}$ in Equations (22) and (23), while Euclidean distances in the reduced space are adopted for the kernel in Equation (24). Based on the resemblance of Equation (22) with the Nyström extension Equation (15), it follows that:

$$\begin{aligned} \frac{\partial s_\alpha^{(0)}}{\partial y_\beta} &= \left(\sum_{i=1}^{M} w^{(0)}\left(y_i, y\right)\right)^{-2} \sum_{i,j=1}^{M} w^{(0)}\left(y_j, y\right) \frac{\partial w^{(0)}(y_i, y)}{\partial y_\beta}\left[\psi_{i,\alpha} - \psi_{j,\alpha}\right] \\ \frac{\partial s_\alpha^{(1)}}{\partial y_\beta} &= \left(\sum_{i=1}^{M} w^{(1)}\left(y_i, y\right)\right)^{-2} \sum_{i,j=1}^{M} w^{(1)}\left(y_j, y\right) \frac{\partial w^{(1)}(y_i, y)}{\partial y_\beta}\left[\psi_{i,\alpha} - \psi_{j,\alpha}\right] - \frac{\partial s_\alpha^{(0)}}{\partial y_\beta} \\ \frac{\partial s_\alpha^{(2)}}{\partial y_\beta} &= \left(\sum_{i=1}^{M} w^{(2)}\left(y_i, y\right)\right)^{-2} \sum_{i,j=1}^{M} w^{(2)}\left(y_j, y\right) \frac{\partial w^{(2)}(y_i, y)}{\partial y_\beta}\left[\psi_{i,\alpha} - \psi_{j,\alpha}\right] - \frac{\partial s_\alpha^{0}}{\partial y_\beta} - \frac{\partial s_\alpha^{1}}{\partial y_\beta} \\ &\vdots \end{aligned} \tag{25}$$

with:

$$\frac{\partial w^{(l)}\left(y_i, y\right)}{\partial y_\beta} = 2\sigma_l^{-2} r_{\beta\beta}^2 w^{(l)}\left(y_i, y\right)\left(y_{i,\beta} - y_\beta\right) \tag{26}$$

and the Jacobian at the right-hand side of Equation (13) is given by:

$$\frac{\partial \psi_\alpha}{\partial y_\beta} = \sum_l \frac{\partial s_\alpha^{(l)}}{\partial y_\beta} \tag{27}$$

Similarly to RBF, LPcan be applied to a subset of the sample points, where, in the above procedure, only nn nearest neighbors of the state, y (u), are considered for restriction (lifting).

A brief explanatory illustration of the use of Laplacian Pyramids for interpolating a multi-scale function at four different levels of accuracy is given in Figure 5; in Figure 6, the same scheme provides an extension of the function $f(\vartheta) = \cos(3\vartheta)$, defined on the circle in $\Re^2$ given by $X^2 + Y^2 = 1$ with $\vartheta = \arctan(Y/X)$.

Figure 5. Illustrating Laplacian Pyramids for a multiscale target function (see the text). The sample dataset is formed by 2000 points evenly distributed in the interval $[0, 10\pi]$. (**Top**) Laplacian Pyramids used as an interpolation procedure at levels 2, 5, 8 and 11 with $\sigma_0 = 30$; (**Bottom**) The difference between the true function values and the Laplacian Pyramids estimates.

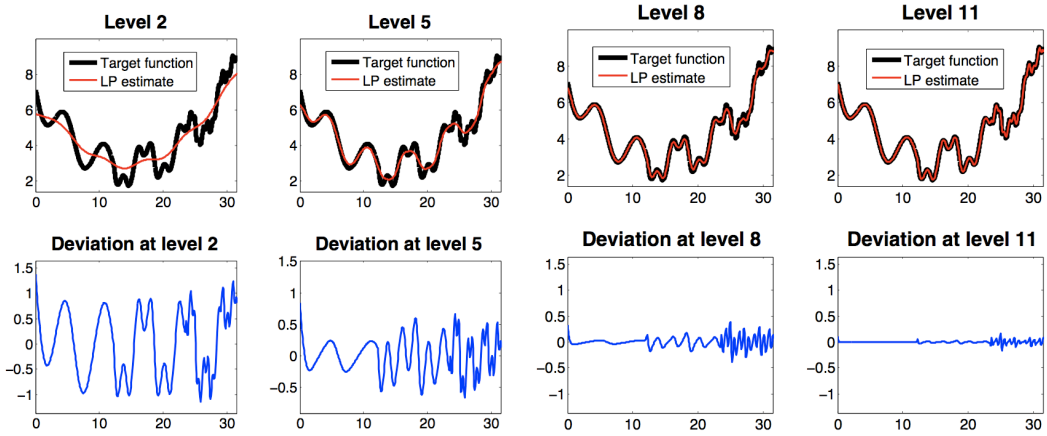
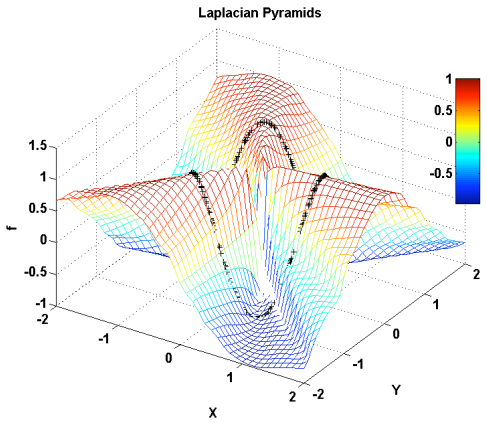


Figure 6. The function $f = \cos(3\vartheta)$, with $\vartheta = \arctan(Y/X)$, is extended to the plane (X, Y) by Laplacian Pyramids (with the finest level $l = 10$ and $\sigma_0 = 10$). The sample set, $\{\mathbf{y}_1, ..., \mathbf{y}_M\}$, is given by $M = 350$ points randomly selected on the unit circle black symbols).



3.2.5. Geometric Harmonics

This is an alternative multi-scale scheme for extending functions only available at M scattered locations, inspired by the Nyström method, and making use of a kernel, w [29]. Let $\mathbf{W}$ be the symmetric $(M \times M)$ matrix, whose generic element reads as:

$$\mathbf{W}(i,j) = \exp\left[-d_{ij}^2/\varepsilon_0\right], \quad i,j = 1,...,M \tag{28}$$

with $\{\phi_{\alpha=1,...,M}\}$ being its full set of orthonormal eigenvectors sorted according to descending eigenvalues $\{\lambda_{\alpha=1,...,M}\}$. For $\delta > 0$, let us define the set of indices $S_\delta = \{\alpha \text{ such that } \lambda_\alpha \geq \delta\lambda_0\}$. The extension of a function, f, defined only at some sample points in $Z \subset \bar{Z}$ to an arbitrary new point in $\bar{Z}$ is accomplished by the following projection step (depending on the purpose, $\bar{Z}$ can be either the ambient space, y, or the reduced one, u):

$$f \to \mathrm{P}_\delta f = \sum_{\alpha \in S_\delta} \langle f, \phi_\alpha \rangle \, \phi_\alpha \tag{29}$$

and the subsequent extension of $\mathrm{P}_\delta f$:

$$\mathbf{E}f = \sum_{\alpha \in S_\delta} \langle f, \phi_\alpha \rangle \, \Psi_\alpha \tag{30}$$

where $\langle \bullet, \bullet \rangle$ denotes the inner product, while Ψ_j reads:

$$\Psi_\alpha = \lambda_\alpha^{-1} \sum_{i=1}^{M} w(y_i, y) \, \phi_{i,\alpha}, \quad \alpha \in S_\delta \tag{31}$$

where $w(y_i, y)$ is computed as indicated in Equation (15). The above is only the first step of a multi-scale scheme, where the function, f, is initially projected at a coarse scale with a large value of the parameter, ε_0, in Equation (28). Afterwards, the residual $f - \mathrm{P}_\delta f$ in the initial coarse projection is projected at a finer scale, ε_1, and so forth at even finer scales $\varepsilon_{l>1}$. A typical approach is to fix ε_0 and then project with $\varepsilon_l = 2^{1-l}\varepsilon_0$ at each subsequent step l till a norm of the residual $f - \mathrm{P}_\delta f$ remains larger than a fixed admissible error. Clearly, both our restriction and lifting operators can be based on Geometric Harmonics.

Similarly to RBF and LP, Geometric Harmonics (GH) can be applied to a subset of the sample points, where, in the above procedure, only nn nearest neighbors of the state, y (or u), are considered for restriction (or lifting).

Figure 7 provides an illustrative multi-scale example, where the Geometric Harmonics approach is used for interpolation purposes for the same multiscale function used in Figure 5. As expected, in the region with low-frequency components, a few steps are sufficient for accurately describing the true function, whereas more iterations are required in the high frequency domain.

We also illustrate the use of Geometric Harmonics in extending the function $f(\vartheta) = \cos(3\vartheta)$, defined on the circle in $\Re^2$, given by $X^2 + Y^2 = 1$ with $\vartheta = \arctan(Y/X)$ in Figure 8.

Figure 7. Geometric Harmonics on a multiscale target function (see the text). The sample dataset is formed by 2000 points evenly distributed in the interval $[0, 10\pi]$. (**Top**) The Geometric Harmonics (GH) scheme is used as an interpolation procedure with $\varepsilon_0 = 3$; (**Bottom**) The difference between the true function values and GH estimates. From left to right: Results corresponding to one, two and eight steps are reported.

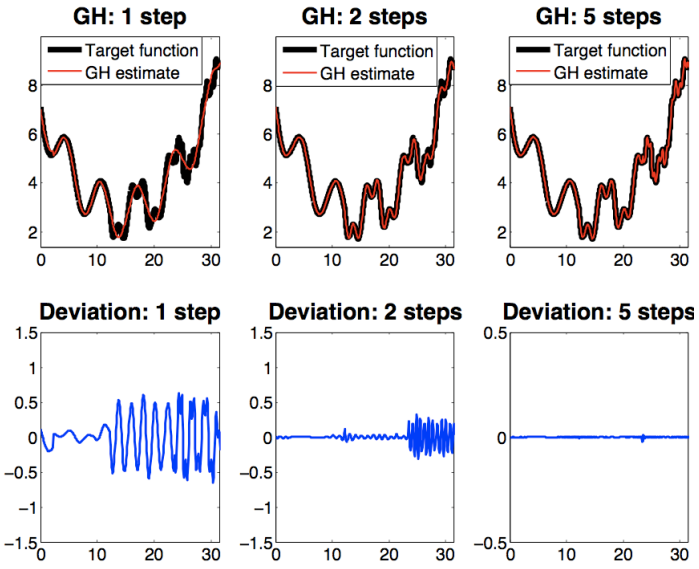
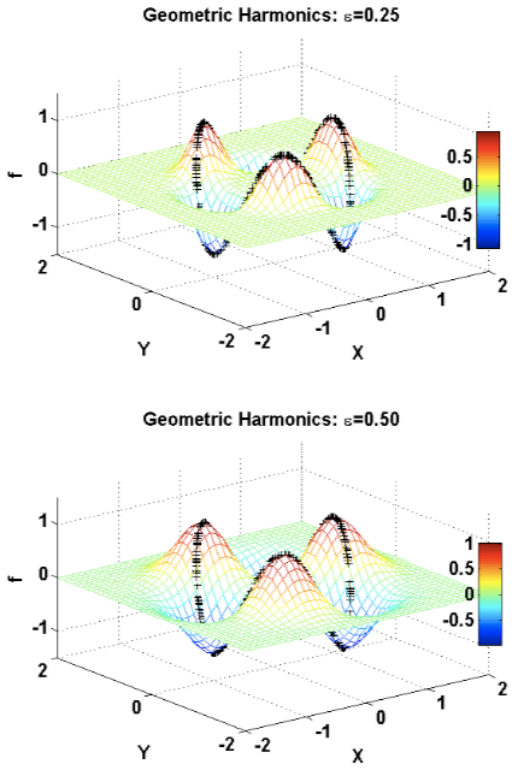


Figure 8. The function $f = \cos(3\vartheta)$, with $\vartheta = \arctan(Y/X)$, is extended to the plane, (X, Y), by Geometric Harmonics using $\varepsilon_0 = 0.25$ (**top**) and $\varepsilon_0 = 0.5$ (**bottom**). The sample set, $\{\mathbf{y}_1, ..., \mathbf{y}_M\}$, is given by $M = 350$ points randomly selected on the unit circle (black symbols).



4. Application to an Illustrative Example: Homogeneous Combustion

We employ our proposed approach described in Section 3 above to search for a two-dimensional reduced system describing the combustion of a mixture of hydrogen and air at stoichiometric proportions under fixed total enthalpy ($H = 300(kJ/kg)$) and pressure ($P = 1[bar]$). We assume that the detailed chemical kinetics is dictated by the Li *et al.* mechanism [30], where 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 involved in the reaction. As shown in Figure 9, the manifold is described by 3810 points and parameterized with respect to the two diffusion map variables, ψ_2 and ψ_3. This two-dimensionality is also consistent with the two slowest stable eigenvalues of the linearization in the neighborhood of the equilibrium point, that are followed by a sizeable gap before the third, much faster stable eigenvalue; this indicates a separation of time scales and suggest a two-dimensional stable submanifold close to the equilibrium.

Figure 9. Homogeneous reactive mixture of hydrogen and air at stoichiometric proportions with fixed enthalpy ($H = 300[kJ/kg]$) and pressure ($P = 1[bar]$). Two dimensional Diffusion Map (DMAP) parameterization of 3810 points in terms of the two nontrivial leading eigenvectors, ψ_2 and ψ_3, of the Markov matrix, $\mathbf{K}$. Colors represent mass fractions, while the black filled circle and the black diamond represent the fresh mixture condition and equilibrium state, respectively.

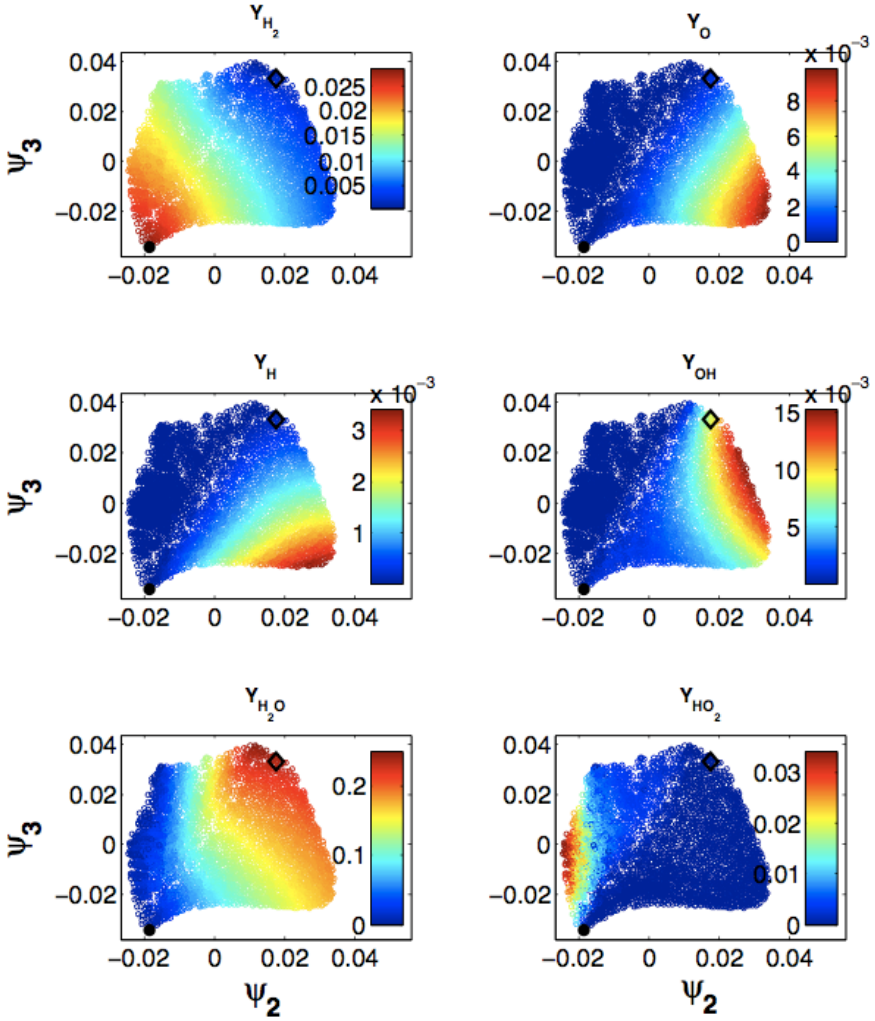
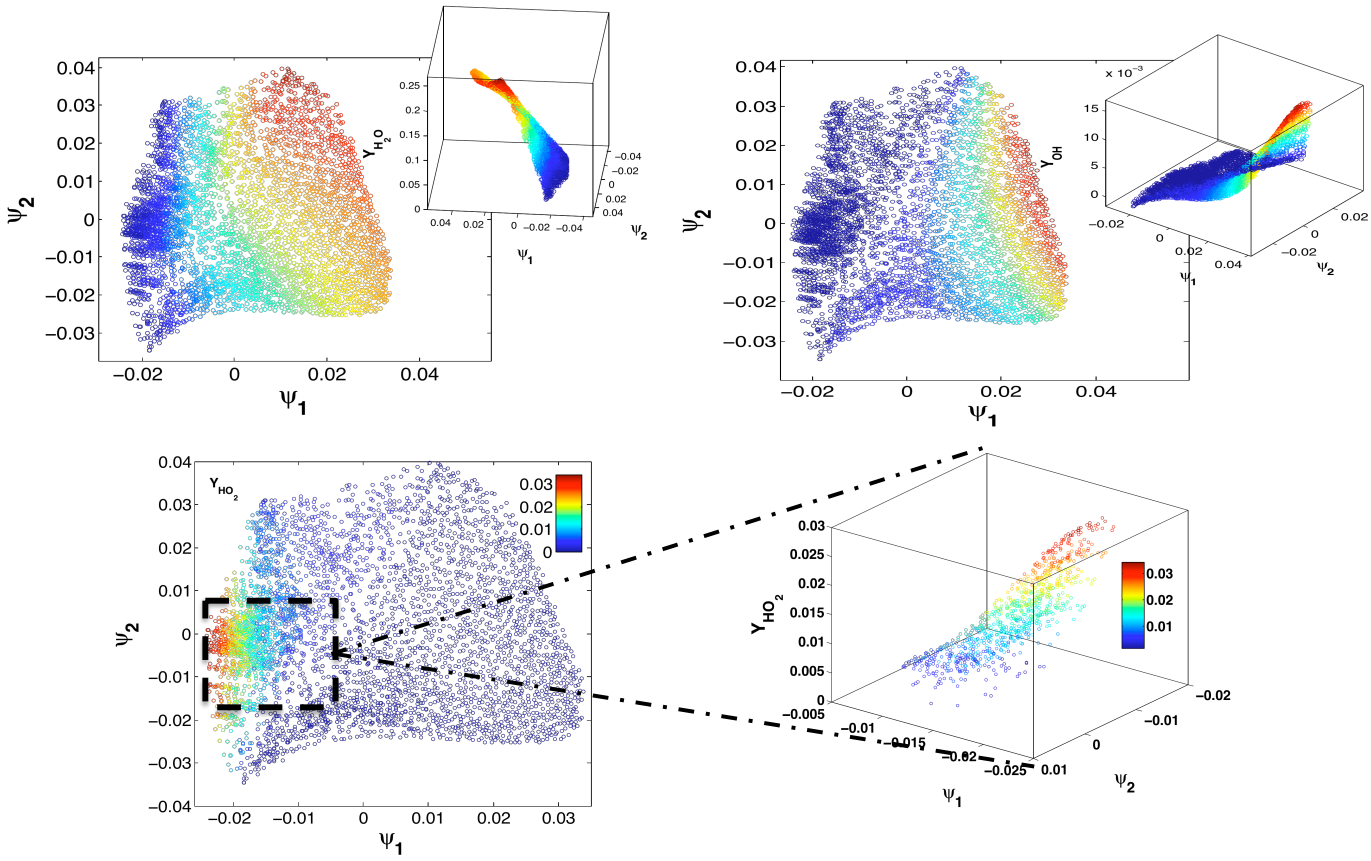


Figure 10. (**Top**) Projection of the manifold onto the $\psi_2 - \psi_3$ plane, where colors refer to the mass fraction of H_2O (**left-hand side**) and OH (**right-hand side**). From the three dimensional views reported in the insets, we can infer that a two-dimensional reduced description is suitable for most of the domain. (**Bottom**) Projection of the manifold onto the $\psi_2 - \psi_3$ plane, where colors refer to the mass fraction of HO_2 (**left-hand side**). A three-dimensional view (**right-hand side**) highlights that in the zoomed region, a higher dimensional description is needed.



It is worth stressing that this two-dimensionality does not persist over the entire phase space.Judging from the sample density in the Diffusion Map space, the considered cloud of points clearly lies on a manifold with different dimensions in different regions. As expected, indeed, low temperature regions (e.g., $T < 1000$ [K]) require a larger number of reduced variables ($m > 2$) to be correctly described (see Figures 10 and 11) [4]. Therefore, in the example below, we only utilize the portion of the manifold with high temperature (say $T > 1200$ [K]). Coping with manifolds with varying dimension is beyond the scope of this paper and should be addressed in forthcoming publications. We mention, however, that attempts of automatically detecting variations of the manifold dimension in the framework of Diffusion Maps have been also recently reported in [23].

Figure 11. (**Top**) A sample detailed transient solution is shown in the plane, $\psi_2 - \psi_3$. Restriction is done by the Nyström method, while colors refer to the temperature (Kelvin) of the gas mixture; (**Bottom**) Time evolution of the absolute deviation between detailed and reduced solution trajectories (in the reduced space) $\left\|\psi^{\mathrm{red}} - \psi^{\mathrm{det}}\right\|$. Numbers in the legend correspond to the first six methods in Table 1.

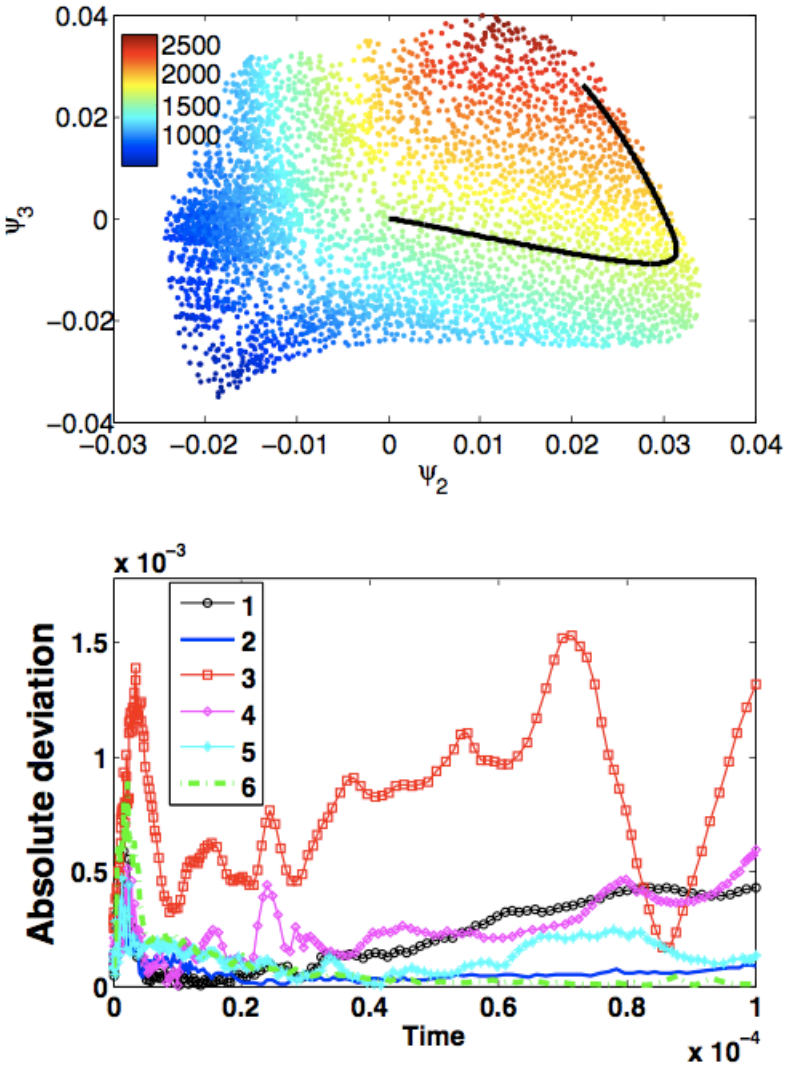


Table 1. Comparison of reduced and detailed solution trajectories (with initial condition $u = [0,0]$ and $0 \leq t \leq \bar{t} = 1 \times 10^{-4}[s]$) corresponding to several schemes implementing lifting and restriction operators (see the text). $\|\delta\psi\|$ indicates the mean deviation between the reduced and detailed solution trajectory (in the reduced space): $\|\delta\psi\| = \bar{t}^{-1}\int_0^{\bar{t}} \left\|\psi^{\text{det}} - \psi^{\text{red}}\right\| dt$, with ψ^{det}, ψ^{red} and $\|\bullet\|$ denoting the restricted detailed solution, the reduced solution and the Euclidean norm, respectively. Similarly, $|\delta y_\alpha|$ is the mean deviation for species α (in the ambient space): $|\delta y_\alpha| = \bar{t}^{-1}\int_0^{\bar{t}} \left|y_\alpha^{\text{det}} - y_\alpha^{\text{red}}\right| dt$.

Method	$\|\delta\psi\|$	$\|\delta y_1\|$	$\|\delta y_3\|$	$\|\delta y_4\|$	$\|\delta y_5\|$	$\|\delta y_6\|$	$\|\delta y_7\|$	$\|\delta y_8\|$	$\|\delta y_9\|$
1	2.28×10^{-4}	2.07×10^{-5}	8.52×10^{-6}	3.58×10^{-5}	4.39×10^{-5}	1.80×10^{-4}	3.16×10^{-4}	2.48×10^{-6}	1.15×10^{-5}
2	5.66×10^{-5}	4.09×10^{-6}	1.31×10^{-6}	5.53×10^{-6}	1.03×10^{-5}	3.78×10^{-5}	8.59×10^{-5}	2.18×10^{-6}	9.65×10^{-6}
3	8.11×10^{-4}	6.90×10^{-5}	2.58×10^{-5}	1.04×10^{-4}	1.62×10^{-4}	6.86×10^{-4}	8.00×10^{-4}	2.33×10^{-6}	9.98×10^{-6}
4	2.64×10^{-4}	2.83×10^{-5}	9.44×10^{-6}	5.37×10^{-5}	1.35×10^{-4}	2.88×10^{-4}	2.71×10^{-4}	1.86×10^{-6}	7.64×10^{-6}
5	1.27×10^{-4}	1.26×10^{-5}	4.23×10^{-6}	2.16×10^{-5}	4.29×10^{-5}	1.17×10^{-4}	1.47×10^{-4}	2.08×10^{-6}	8.63×10^{-6}
6	7.31×10^{-5}	8.38×10^{-6}	2.30×10^{-6}	9.15×10^{-6}	1.78×10^{-5}	6.76×10^{-5}	9.46×10^{-5}	5.70×10^{-6}	2.68×10^{-5}
7	7.39×10^{-4}	7.25×10^{-5}	2.86×10^{-5}	1.24×10^{-4}	1.97×10^{-4}	5.96×10^{-4}	9.94×10^{-4}	1.87×10^{-6}	7.93×10^{-6}
8	8.81×10^{-4}	5.90×10^{-5}	4.95×10^{-5}	1.74×10^{-4}	1.14×10^{-4}	4.99×10^{-4}	6.36×10^{-4}	6.27×10^{-6}	2.83×10^{-5}
9	0.0058	3.83×10^{-4}	2.45×10^{-4}	0.00107	9.66×10^{-4}	0.0034	0.0061	6.50×10^{-6}	3.75×10^{-5}
10	0.0140	0.00126	7.24×10^{-4}	0.00283	0.00184	0.0123	0.0162	2.27×10^{-5}	9.89×10^{-5}
11	8.08×10^{-4}	9.03×10^{-5}	4.10×10^{-5}	1.63×10^{-4}	1.75×10^{-4}	8.33×10^{-4}	0.00116	2.79×10^{-6}	1.19×10^{-5}
12	0.0237	0.00331	0.00103	0.00424	0.00453	0.030	0.0331	1.11×10^{-4}	5.45×10^{-4}

We discretized the reduced space by a 60×60 uniform Cartesian grid with $-0.025 < \psi_2 < 0.035$ and $-0.035 < \psi_3 < 0.04$. At every grid node, the values of the right-hand side of Equations (13) or (14) are computed according to several interpolation schemes chosen form the ones described above in Section 3, and stored in tables for later use. In particular, tables were created using the following methods:

1. The lifting operator consists of radial basis function interpolation with $p = 3$ performed over 50 nearest neighbors of an arbitrary point in the reduced space, u. Restriction is done by radial basis function interpolation with $p = 3$ performed over 50 nearest neighbors of an arbitrary point in the ambient space, y. The reduced dynamical system is expressed in the form of Equation (13).
2. The lifting operator consists of radial basis function interpolation with $p = 3$ performed over 50 nearest neighbors of an arbitrary point in the reduced space, u. Restriction is done by the Nyström method. The reduced dynamical system is expressed in the form of Equation (13).
3. The lifting operator is based on Laplacian Pyramids up to a level $l = 20$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the reduced space, u. Restriction is based on the Laplacian Pyramids up to a level $l = 7$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of y. The reduced dynamical system is expressed in the form of Equation (13).
4. The lifting operator is based on Laplacian Pyramids up to a level $l = 20$ with $\sigma_0 = 0.5$. Restriction is done by the Nyström method. The reduced dynamical system is expressed in the form of Equation (13).
5. The lifting operator is based on Geometric Harmonics locally performed over 15 nearest neighbors of an arbitrary point in the reduced space, u. Refinements are performed until the Euclidean norm of the residual is larger than 5×10^{-4}. Restriction is done by the Nyström method. The reduced dynamical system is expressed in the form of Equation (13).
6. The lifting operator is based on Kriging performed over eight nearest neighbors of an arbitrary point in the reduced space, u (DACE package [27], with a second order polynomial regression

model, a Gaussian correlation model and parameter $\theta = 10^{-3}$). Restriction is done by the Nyström method. The reduced dynamical system is expressed in the form of Equation (13).

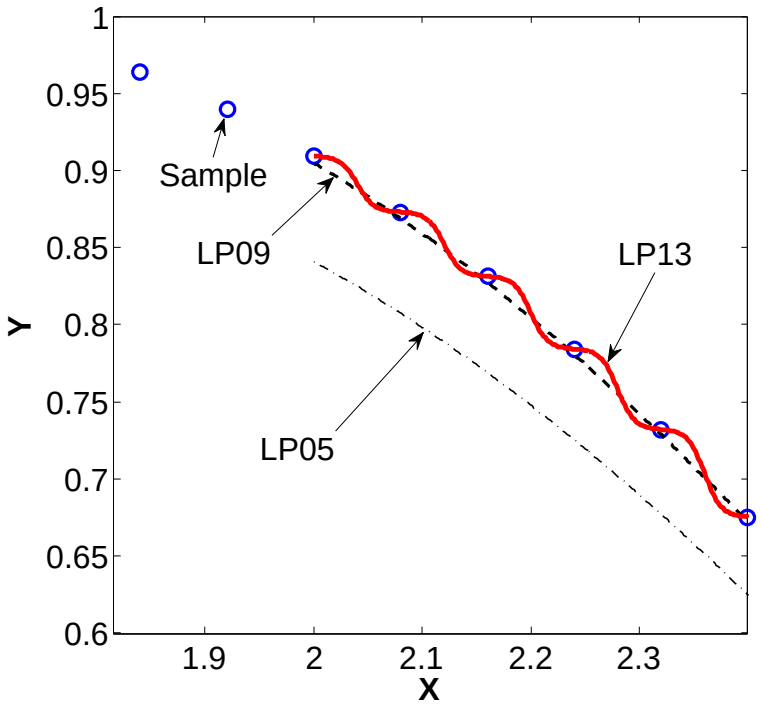
7. The lifting operator is based on Geometric Harmonics locally performed over 10 nearest neighbors of an arbitrary point in the reduced space, u. Refinements are performed until the Euclidean norm of the residual is larger than 10^{-3}. Restriction is done using the Nyström method. The reduced dynamical system is expressed in the form of Equation (13).
8. The lifting operator is based on Kriging performed over eight nearest neighbors of an arbitrary point in the reduced space, u (DACE package [27], with a second order polynomial regression model, a Gaussian correlation model and parameter $\theta = 10^{-3}$). Restriction is done by the Nyström method. The reduced dynamical system is expressed in the form of Equation (14).
9. The lifting operator is based on Kriging performed globally over all samples (package [27], with a second order polynomial regression model, a Gaussian correlation model and parameter $\theta = 13$). Restriction is done by the Nyström method. The reduced dynamical system is expressed in the form of Equation (14).
10. The lifting operator is based on the Laplacian Pyramids up to a level $l = 20$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the reduced space, u. Restriction is based on the Laplacian Pyramids up to a level $l = 3$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the ambient space, y. The reduced dynamical system is expressed in the form of Equation (13).
11. The lifting operator is based on the Laplacian Pyramids up to a level $l = 20$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the reduced space, u. Restriction is based on the Laplacian Pyramids up to a level $l = 9$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the ambient space, y. The reduced dynamical system is expressed in the form of Equation (13).
12. The lifting operator is based on the Laplacian Pyramids up to a level $l = 20$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the reduced space, u. Restriction is based on Laplacian Pyramids up to a level $l = 12$ with $\sigma_0 = 0.5$ over 80 nearest neighbors of an arbitrary point in the ambient space, y. The reduced dynamical system is expressed in the form of Equation (13).

Each of the above tables was utilized for providing systems Equations (13) and (14) with a closure, where the rates of reduced variables are efficiently retrieved via bi-variate interpolation in the Diffusion Map space. In Figure 11, a sample trajectory (starting from $u = [0, 0]$) is reported in the top part, while the Euclidean norm of the absolute deviation between the reduced and detailed solution (in the $\psi_2 - \psi_3$ plane) is reported in the lower part of the figure as a function of time. A more detailed comparison is reported in the Table 1. In our (not optimized) implementation, all trajectories are computed by the Matlab's solver, **ode45**, with the reduced system showing a speedup of roughly four times compared to the detailed one.

In terms of accuracy, we found that the best performances are achieved combining a local lifting operator (e.g., interpolation/extension over nearest neighbors) with the Nyström method for restriction. For instance, we notice that a proper combination between radial basis function interpolation (for lifting) and Nyström extension may offer excellent accuracy (in terms of deviation errors $\|\delta\psi\|$ and $|\delta y_i|$), as shown in Table 1 for the solution trajectory in Figure 11. Clearly, radial basis functions are simpler to

implement and require less computational resources compared to other approaches, such as Kriging and Geometric Harmonics. We should stress, though, that the latter techniques present similar performances and are certainly to be preferred in cases where (unlike Figures 9 and 11) samples are not uniformly distributed (*i.e.*, sample clustering). Moreover, we observe that approaches based on Laplacian Pyramids (for restriction) present poorer performances even with large values of l. An explanation for this is a possible inaccurate estimate of the derivatives at the right-hand side of the reduced dynamical system, which we attempt to illustrate through the caricature in Figure 12.

Figure 12. Illustrating a possible pathology. Samples (circles) are uniformly chosen in X, with $Y = \sin(X)$. Laplacian Pyramids are adopted for interpolation between samples with $\sigma_0 = 10$. Estimated values with the finest level $l = 5$, $l = 9$ and $l = 13$ are denoted by LP05, LP09 and LP13, respectively. At the latter level, the estimates *of derivatives* are no longer accurate.



We finally find that solutions to system Equation (14) typically lead to larger errors compared to those obtained solving Equation (13).

For completeness, in Figure 13, we report the time series of the Diffusion Maps' variables as obtained by Methods 2 and 3 in Table 1, as well as the restriction of the corresponding detailed solution. Moreover, in Figures 14 and 15, a comparison of the time series in the detailed space is reported as obtained by reconstruction of the states in $\Re^9$ from the reduced solutions in Figure 13.

Figure 13. Time evolution of the two diffusion map variables along the solution trajectory of Figure 11 as obtained by Method 2 (**top**) and by Method 3 (**bottom**) (see Table 1). The initial condition in the Diffusion Maps' space $[0, 0]$ is first lifted into $\Re^9$ and then relaxed towards the equilibrium point by the detailed kinetics Equation (8) using the readily available Matlab solver, **ode45**. The latter time series is afterwards restricted to the Diffusion Maps' space and reported with a continuous line. Symbols denote the corresponding solution directly obtained in the reduced space by solving system Equation (13) by the same Matlab solver, **ode45**.

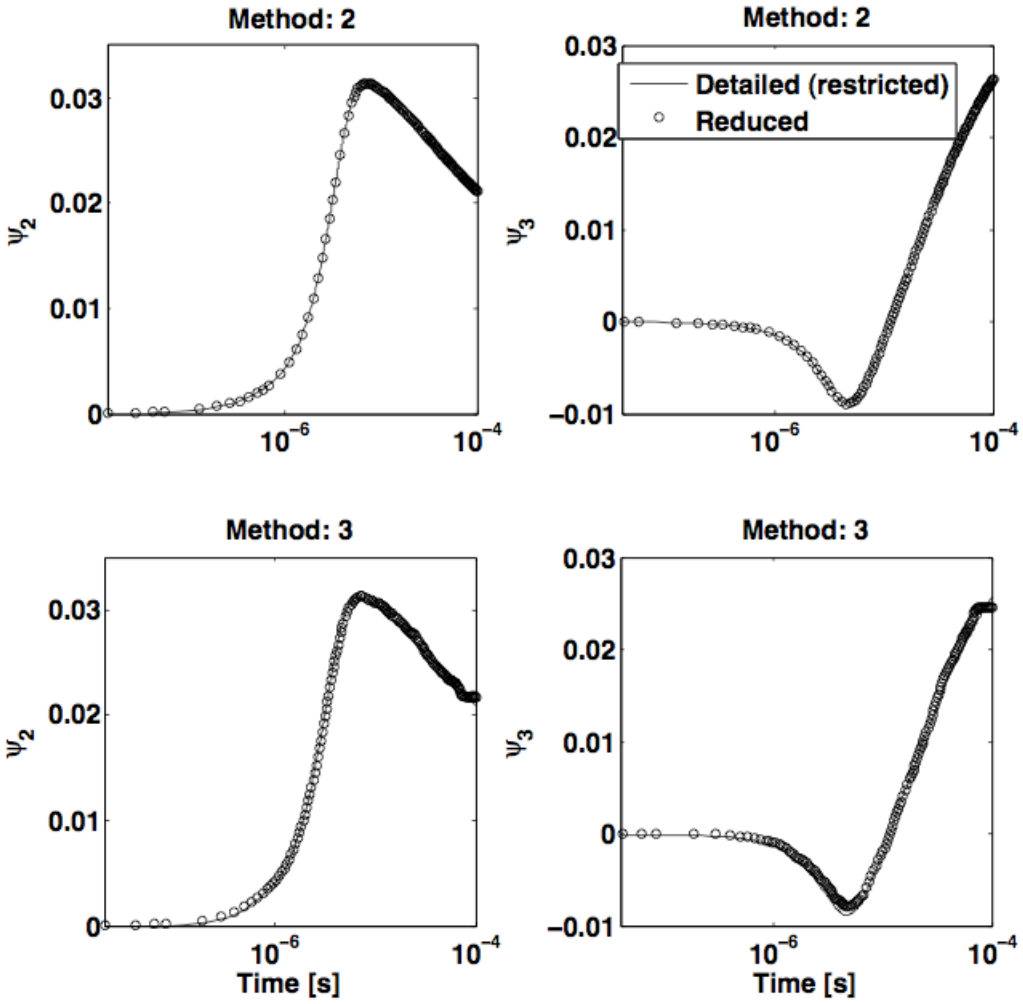


Figure 14. The initial condition in the Diffusion Maps' space $[0, 0]$ is first lifted into $\Re^9$ and then relaxed towards the equilibrium point by the detailed kinetics Equation (8) using the readily available Matlab solver, **ode45** (continuous line). Symbols report the corresponding time series as obtained by lifting the reduced solution at the top of Figure 13 (*i.e.*, Method 2).

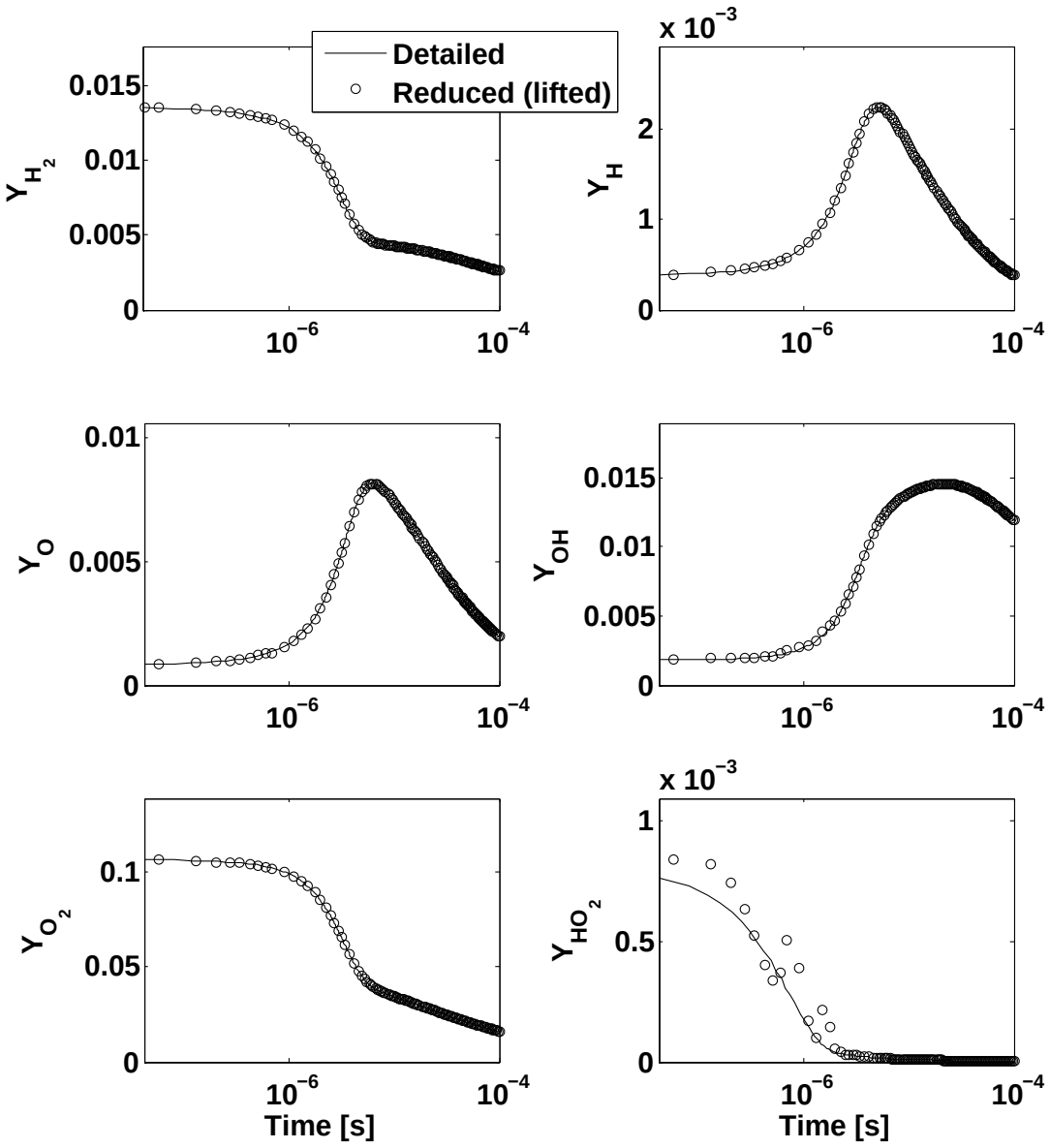
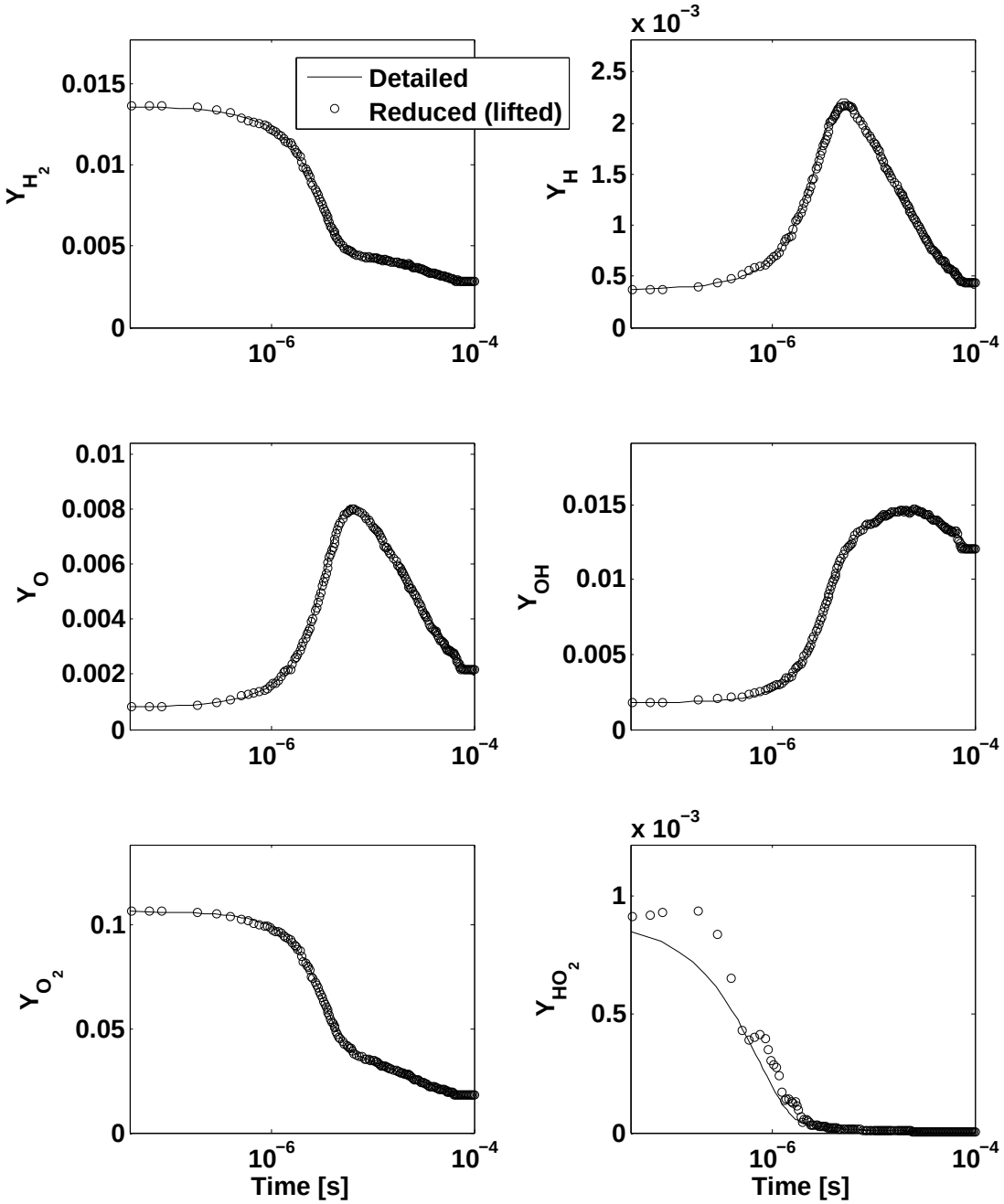


Figure 15. The initial condition in the Diffusion Maps' space $[0, 0]$ is first lifted into $\Re^9$ and then relaxed towards the equilibrium point by the detailed kinetics Equation (8) using the readily available Matlab solver, **ode45** (continuous line). Symbols report the corresponding time series as obtained by lifting the reduced solution at the bottom of Figure 13 (*i.e.*, method 3).



5. Conclusions

In this work, we showed that the Diffusion Map (DMAP) technique is a promising tool for extracting a global parameterization of low-dimensional manifolds arising in combustion problems. Based on the slow variables automatically identified by the process, a reduced dynamical system can be obtained and

solved. Both lifting and restriction operators (*i.e.*, mapping of any point in the region of interest of the reduced space into the full space and *vice versa*) lie at the heart of such an approach. To construct these operators, methods for extending empirical functions only known at scattered locations must be employed, and we have tested several.

For chemical kinetics governing a non-isothermal reactive gas mixture of hydrogen and air, a comparison is carried out on the basis of the deviation error between sample detailed solutions and the corresponding reduced ones in both the full and reduced spaces. Several combinations of interpolation schemes were implemented in the procedure restrictions/liftings, with the reduced rates, du/dt, pre-computed and stored in tables to be utilized at a later time for providing system Equation (13) with a closure. In the considered case, approaches based on a local lifting operator, *i.e.*, interpolation/extension over nearest neighbors, combined with the Nyström method (for restriction) have shown superior performances in terms of accuracy in recovering the (longer-time) transient dynamics of the detailed model.

While the feasibility of the presented approach has been demonstrated here, a number of open issues remain. In particular, future studies should focus on computationally efficient implementations of the method without pre-tabulation, since handling tables at high dimensions (say $m > 4$) becomes computationally complex. Clearly, the approach can only start after the initial collection of a (rich enough) dataset from simulations. It is worth noting, however, that the reduced coordinates that result from data mining can also be helpful in designing subsequent data collection to *further extend* the manifold if necessary: Taylor-series-based extensions close to the boundaries (in a reduced-dimension Diffusion Map space) of the dataset collected can provide initial conditions for further data-collection simulations. The idea has been, in principle, proposed in the past [31] in the context of navigating effective free energy surfaces, and its incorporation in our Diffusion Map-based approach here is worth pursuing.In this paper the initially collected data came from a detailed set of kinetic ordinary differential equations, and the effective model for the reduced variables was also assumed to be a set of ODEs in these variables. The approach, however, can also, in principle, be used when the data do not need to come from a large set of ODEs, but, for example, from multiscale Stochastic Simulation Algorithm descriptions of chemical kinetic schemes (for such a recent application, see [32]). Moreover, as demonstrated also in the presented combustion example, the method should be able to cope with manifolds, whose dimensions possibly vary across distinct regions of the phase-space (see [23]); how to consistently express and solve reduced systems across manifolds with disparate dimensions remains out of reach of the present method, requiring further investigation.

Acknowledgments

E.C. acknowledges partial support of the U.S.-Italy Fulbright Commission and the Italian Ministry of Research (FIRBgrant RBFR10VZUG). I.G.K. and C.W.G. gratefully acknowledge partial support by the U.S. DOE. C.J.D. acknowledges support by the U.S. Department of Energy Computation Science Graduate Fellowship (grant number DE-FG02-97ER25308).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Maas, U.; Goussis, D. Model Reduction for Combustion Chemistry. In *Turbulent Combustion Modeling*; Echekki, T., Mastorakos, E., Eds.; Springer: New York, NY, USA, 2011; pp. 193–220.
2. Chiavazzo, E.; Karlin, I.V.; Gorban, A.N.; Boulouchos, K. Coupling of the model reduction technique with the lattice Boltzmann method for combustion simulations. *Combust Flame* **2010**, *157*, 1833–1849.
3. Ren, Z.; Pope, S.B.; Vladimirsky, A.; Guckenheimer, J.M. The invariant constrained equilibrium edge preimage curve method for the dimension reduction of chemical kinetics. *J. Chem. Phys.* **2006**, *124*, 114111.
4. Chiavazzo, E.; Karlin, I.V. Adaptive simplification of complex multiscale systems. *Phys. Rev. E* **2011**, *83*, 036706.
5. Chiavazzo, E. Approximation of slow and fast dynamics in multiscale dynamical systems by the linearized Relaxation Redistribution Method. *J. Comp. Phys.* **2012**, *231*, 1751–1765.
6. Pope, S.B. Computationally efficient implementation of combustion chemistry using *in situ* adaptive tabulation. *Combust Theory Model* **1997**, *1*, 41–63.
7. Coifman, R.R.; Lafon, S.; Lee, A.B.; Maggioni, M.; Nadler, B.; Warner, F.; Zucker, S.W. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7426–7431.
8. Coifman, R.R.; Lafon, S.; Lee, A.B.; Maggioni, M.; Nadler, B.; Warner, F.; Zucker, S.W. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Multiscale methods. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7432–7437.
9. Coifman, R.R.; Lafon, S. Diffusion maps. *Appl. Comput. Harmon. Anal.* **2006**, *21*, 5–30.
10. Jolliffe, I.T. *Principal Component Analysis*; Springer-Verlag: New York, NY, USA, 2002.
11. Jones, P.W.; Maggioni, M.; Schul, R. Manifold parametrizations by eigenfunctions of the laplacian and heat kernels. *Proc. Nat. Acad. Sci. USA* **2008**, *105*, 1803–1808.
12. Lafon, S. Diffusion Maps and Geometric Harmonics. Ph.D. Thesis, Yale University, New Haven, CT, USA, 2004.
13. Grassberger, P. On the Hausdorff dimension of fractal attractors. *J. Stat. Phys.* **1981**, *26*, 173–179.
14. Grassberger, P.; Procaccia, I. Measuring the strangeness of strange attractors. *Physica D* **1983**, *9*, 189–208.
15. Coifman, R.R.; Shkolnisky, Y.; Sigworth, F.J.; Singer, A. Graph laplacian tomography from unknown random projections. *IEEE Trans. Image Process.* **2008**, *17*, 1891–1899.
16. Maas, U.; Pope, S. Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space. *Combust Flame* **1992**, *88*, 239–264.
17. Kevrekidis, I.G.; Gear, C.W.; Hyman, J.M.; Kevrekidis, P.G.; Runborg, O.; Theodoropoulos, C. Equation-free, coarse-grained multiscale computation: Enabling mocroscopic simulators to perform system-level analysis. *Comm. Math. Sci.* **2003**, *1*, 715–762.

18. Kevrekidis, I.G.; Gear, C.W.; Hummer, G. Equation-free: The computer-aided analysis of complex multiscale systems. *AIChE J.* **2004**, *50*, 1346–1355.
19. Kroese, D.P.; Taimre, T; Botev, Z.I. *Handbook of Monte Carlo Methods*; Wiley: Hoboken, NJ, USA, 2011.
20. Kaufman, D.E.; Smith, R.L. Direction choice for accelerated convergence in hit-and-run sampling. *Oper. Res.* **1998**, *46*, 84–95.
21. Gear, C.W. Parameterization of Non-Linear Manifolds. Available online: http://www.princeton.edu/wgear/ (accessed on 23 September 2013).
22. Sonday, B.E.; Gear, C.W.; Singer, A.; Kevrekidis, I.G. Solving differential equations by model reduction on learned manifolds. Unpublished, 2011.
23. Rohrdanz, M.A.; Zheng, W.; Maggioni, M.; Clementi, C. Determination of reaction coordinates via locally scaled disffusion map. *J. Chem. Phys.* **2011**, *134*, 124116.
24. Nyström, E.J. Über die praktische Auflösung von linearen integralgleichungen mit anwendungen auf randwertaufgaben der potentialtheorie. *Commentationes Physico-Mathematicae* **1928**, *4*, 1–52 (in German).
25. Sonday, B.E. Systematic Model Reduction for Complex Systems through Data Mining and Dimensionality Reduction. Ph.D. Thesis, Princeton University, Princeton, NJ, USA, 2011.
26. Isaaks, E.H.; Srivastava, R.M. *An Introduction to Applied Geostatistics*; Oxford University Press: New York, NY, USA, 1989.
27. Lophaven, S.N.; Nielsen, H.B.; Søndergaard, J. DACE A Matlab Kriging Toolbox. In *Technical Report IMM-TR-2002-12*; Technical University of Denmark: Kongens Lyngby, Denmark, 2002; pp. 1–26.
28. Rabin, N.; Coifman, R.R. Heterogeneous Datasets Representation and Learning Using Diffusion Maps and Laplacian Pyramids. In Proceedings of the 12th SIAM International Conference on Data Mining, Anaheim, CA, USA, 26–28 August 2012; pp. 189–199.
29. Coifman, R.R.; Lafon, S. Geometric harmonics: A novel tool for multiscale out-of-sample extension of empirical functions. *Appl. Comput. Harmon. Anal.* **2006**, *21*, 31–52.
30. Li, J.; Zhao, Z.; Kazakov, A.; Dryer, F.L. An updated comprehensive kinetic model of hydrogen combustion. *Int. J. Chem. Kinet.* **2004**, *36*, 566–575.
31. Frewen, T.A.; Hummer, G.; Kevrekidis, I.G. Exploration of effective potential landscapes using coarse reverse integration. *J. Chem. Phys.* **2009**, *131*, 134104.
32. Dsilva, C.J.; Talmon, R.; Rabin, N.; Coifman, R.R.; Kevrekidis, I.G. Nonlinear intinsic variables and state reconstruction in multiscale simulations. *J. Chem. Phys.* **2013**, *139*, 184109.