CA22151: Cyber-Physical systems and digital twins for the decarbonisation of energy-intensive industries

Report on low-dimensional manifold identification and topology optimisation for constructing reduced-order models

Authors: Grégoire Corlùy (ULB), Anh Khoa Doan (TU Delft), Biagio Cassese (CNR-STEMS), Giancarlo Sorrentino (CNR-STEMS), Alessandro Parente (ULB)

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the European Union

CYPHER
cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY
Funded by the European Union

Table of Contents

1. Introduction

Analyzing combustion systems through experiments or simulations is expensive in cost and time. This makes it unaffordable to explore many operating conditions to iterate the design of a combustion system. For what concerns the simulations, the high computational cost comes from the high-dimensional features and complexity of the combustion system [1]. It is typically defined by tens to hundreds of species and contains hundreds to thousands of reactions to solve. Furthermore, the fluid dynamics and the chemistry are interconnected and happening at a large range of timescales making this a stiff problem to solve. The principal challenges associated with the modeling of turbulent combustion arise from the involvement of several chemical species, the presence of a wide range of spatial and time scales, and the intricate coupling among multiple physical sub-processes [2] . In hydrocarbon combustion, the number of chemical species can vary significantly, ranging from approximately 50 to over 7000, depending on the specific fuel considered. Nevertheless, for many common fuels, detailed chemical kinetic mechanisms are available that typically comprise between 150 and 250 species. The numerical resolution of the governing conservation equations requires an accurate representation of all relevant spatial and temporal scales. The characteristic length scales extend from the macroscopic dimensions of the combustion apparatus to the smallest dissipative scales, such as the Kolmogorov length scale of turbulence, or, in certain combustion regimes, to even finer scales within the reaction zones. Similarly, the pertinent time scales range from the overall residence time of the flow to the minimum of the Kolmogorov time scale and the fastest chemical time scales. Given the broad disparity across these spatial and temporal ranges, it is widely acknowledged that direct numerical simulation (DNS) of practical combustion devices will remain computationally prohibitive for the foreseeable future. Consequently, any viable modeling framework must necessarily represent the small-scale processes through statistical or averaged approaches rather than by direct numerical resolution. Over recent decades, several methodologies has been developed for the modeling of turbulent combustion. Most frameworks address the challenges posed by fine spatial–temporal scales and large chemical mechanisms through either RANS or LES closures in combination with reduced descriptions of the chemistry. The principal point of divergence lies in how the coupling between chemical reaction and molecular diffusion is represented. Broadly, prevailing strategies fall into two classes: flamelet-type and PDF-like approaches. The steady flamelet model is the canonical example of flamelet-type formulations. Its defining features are: - Strong assumptions regarding reaction–diffusion coupling, leading to the hypothesis that species mass fractions evolve on a very low-dimensional manifold (typically 2D or 3D) in the composition space. - Determination of manifold properties obtained through laminar-flame computations performed a-priori. - Tabulation of the manifold quantities required during the turbulent computation (feasible only for very low-dimensional manifolds). On the other hand, transported PDF methods exemplify PDF-like approaches, characterized by: - No imposed restriction that compositions lie on a low-dimensional manifold - A computational representation in which fluid composition is carried by species mass fractions associated with a large ensemble of particles or other stochastic elements. - An exact treatment of chemical source terms without additional modeling assumptions. Compared with flamelet-type models, PDF-like methods benefit from avoiding low-dimensional manifold constraints and from an exact treatment of reaction kinetics; however, they must deal with modeling reactive mixing and with the substantial computational expense of treating detailed chemistry within the turbulent calculation itself rather than in the pre-processing stage. Although they are high-dimensional, combustion system can intrinsically be characterized by a few controlling variables defining a

low-dimensional manifold which can allow to accelerate simulations [3]–[5]. Employing this property, it is possible to speed up the simulations by running in this reduced space, called a reduced-order model (ROM) [6]. In this way, instead of solving the equations for every species of the system, only the equations for the reduced variables have to be solved. After simulation in reduced space, regression can be performed to retrieve the original full-dimensional thermo-chemical pattern. Moreover, projecting the high-dimensional data onto the reduced space makes it possible to visualize it and understand the hidden patterns.

Experts usually define the low-dimensional manifold using thermochemical variables based on their knowledge of the system [7], [8]. But recently, data-driven techniques have emerged to parametrize automatically the low-dimensional manifold [9], [10]. These techniques help to find improved manifold topologies representing better the data in the reduced space which in turn also improve the accuracy of the ROMs.

2. Purpose of low-dimensional manifold for ROM

Combustion is governed by the conservation of the total mass and the mass of each component, the conservation of momentum and energy. Especially the conservation of mass for each component that is of interest here. For each species, a partial derivative equation (PDE) has to be solved as given in Equation 2.

$$\frac{\partial(\rho Y_i)}{\partial t} = -\nabla \cdot (\rho \mathrm{u} Y_i) + \nabla \cdot (\rho D_{i,m} \nabla Y_i) + \dot{\omega}_i. \tag{1}$$

with Y_i and $\dot{\omega}_i$ respectively the mass fraction and source term of species i, ρ the density, $D_{i,m}$ the diffusion coefficient. The first term of the right hand side corresponds to the convection term, the second term to the diffusion term and the third term to the source term coming from the chemical mechanism.

The computational burden of combustion simulations comes from this transport equation as this PDE has to be solved for every species [1]. A combustion process typically includes tens to hundreds of species and as a result this step becomes costly to compute.

Taking advantage from the fact that the combustion process can be described by a few variables [3]–[5], the number of equations to solve can be drastically reduced. A few manifold variables can be defined as a linear combination of the species' mass fraction and are commonly called progress variables (PV). Thus, the PVs are defined as $Y_{\mathrm{PV}} = \sum_{i=0}^{n} w_i * Y_i$ with w_i the weight assigned to species i. Using the linearity property of partial derivatives and plugging the definition of the PVs in Equation 2 gives the following equation:

$$\frac{\partial(\rho Y_{\mathrm{PV}})}{\partial t} = -\nabla \cdot (\rho \mathrm{u} Y_{\mathrm{PV}}) + \nabla \cdot (\rho D \nabla Y_{\mathrm{PV}}) + \dot{\omega}_{\mathrm{PV}}. \tag{2}$$

assuming the same diffusion coefficient for every species.

As a result, instead of solving tens to hundreds of PDEs for every species, only two or three need to be solved, *i.e.*, the number of progress variables summarizing the combustion, and hence speeding up the simulations.

The concept of a low-dimensional manifold lies at the heart of reduced-order modeling because it captures the essential dynamics of complex reacting systems within a compact and computationally tractable subspace. In combustion, despite the apparent high dimensionality of the chemical composition space, the evolution of thermo-chemical states is often

strongly correlated and constrained by a limited number of slow, dominant modes. These slow modes govern the overall reaction progress, while fast transients rapidly relax toward a low-dimensional attractor. This property justifies the approximation of the full chemical state by a small set of controlling variables defining an intrinsic manifold in the composition space. Identifying such a manifold enables an efficient mapping between the reduced variables and the full thermo-chemical state. In the context of ROMs, this mapping replaces the explicit integration of detailed kinetics by precomputed or learned relationships, drastically decreasing computational cost while preserving accuracy for the relevant operating conditions. From a physical standpoint, the manifold encodes the coupled effects of chemistry and transport, representing the quasi-equilibrium surface that the system trajectories follow during combustion. From a computational standpoint, it provides a consistent reduced coordinate system for tabulation, interpolation, and machine-learning-based regression. The choice of the manifold variables and their topology directly determine the performance of the ROM. A well-chosen manifold ensures that key quantities of interest - such as temperature, intermediate species, and pollutant precursors - can be reconstructed accurately and uniquely from the reduced coordinates. Conversely, an ill-posed or non-unique mapping leads to ambiguity and loss of predictive capability. Consequently, significant research efforts have focused on both physics-informed and data-driven strategies for manifold identification and optimization. Recent developments in machine learning have made it possible to discover non-linear manifolds that better capture complex reaction–diffusion interactions while maintaining smoothness and monotonicity—properties essential for numerical stability and regression. In summary, the purpose of introducing a low-dimensional manifold is not merely to simplify the governing equations but to provide a physically meaningful and computationally efficient representation of the reacting flow. The manifold acts as the cornerstone upon which ROMs are constructed, defining the reduced coordinate space in which chemistry and transport can be accurately and rapidly evaluated. The following section introduces the flamelet-like models, which formalize this principle by precomputing manifold relationships under simplified laminar configurations and using them to close the governing equations in turbulent simulations.

2.1. Flamelet-like models

Among the most important manifold-based techniques, the flamelet-like models or tabulated chemistry methods [2] play a key role in including the effects of detailed chemistry in the modeling of turbulent reactive flows [11].

Tabulated chemistry aims at expressing the thermochemical variables in a reduced chemical state space prior to a CFD computation. The set of species mass fractions involved in detailed mechanisms is replaced by a reduced set of coordinates $(\psi_1, \psi_2, ..., \psi_n)$, where n is the number of dimensions of the thermochemical database. Tabulated chemistry is efficient in comparison with detailed chemistry if $n << \mathrm{N}_s^d$. This method relies on the observation that the chemical trajectories accessed during the combustion process are confined to a subspace of low dimension n called manifold. Thermo-chemical properties φ such as species mass fractions, chemical reaction rates, temperature are then approximated by a n-variables function $\mathcal{F}^{tab}$ in the reduced state of coordinates $(\psi_1, \psi_2, ..., \psi_n)$:

$$\varphi \approx \mathcal{F}^{tab}\left(\psi_1, ..., \psi_n\right). \tag{3}$$

Such manifold does not have an analytical expression but is defined in a discrete form through a chemical look-up table evaluated with detailed chemistry, explaining the terminology *tabulated chemistry*.

There exists different variants like the flamelet generated manifold (FGM) [6], flamelet-progress variable (FPV) [12], steady flamelet model [13], reaction-diffusion manifold (REDIM) [14], [15], flame prolongation of the intrinsic low-dimensional manifold (FPI) [16]. They all rely on projecting the state space onto a lower-dimensional manifold to replace the detailed combustion chemistry mechanism and therefore speeding up the simulation. As an example for FGM, the manifold is generated with multiple 1D flamelets at varying conditions as shown in Figure 1 [17]. Given the manifold, the chemistry is obtained through tabulation [18], [19]. The progress variable source term is therefore obtained by linear interpolation given the control variables of the table. The two major drawbacks of this technique are the memory usage and the inference time. To be accurate, a large table is required. Furthermore, retrieving a value from this large table is slow, which is not optimal when performing simulations. Hence, neural networks [20], [21] and other machine learning techniques [22] have become popular in the last decade to replace tabulation solving both the memory and inference speed problems.

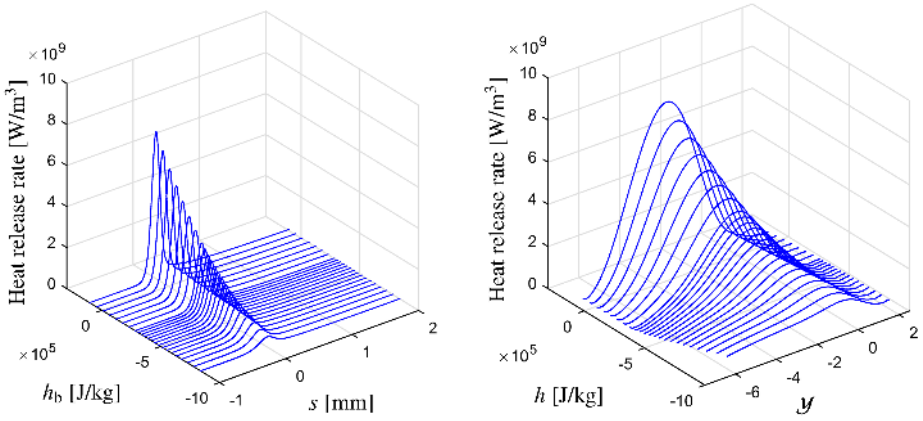

Figure 1: Example of a flamelet generated manifold (FGM) [17].

In this context, the use of single flamelet archetype is efficient to capture the chemical structure of well-identified flames (such as purely premixed or diffusion flames) but introduces bias in the prediction of the chemical structure of more complex situation such as stratified flames. Both premixed and non-premixed flamelet ingredients must be combined in the same look-up table. A solution is to solve the projection of the full set of mass conservation species balance equations into a restricted subset of the composition space or to use 1-D partially-premixed flamelets to generate a chemical look-up table. Another issue encountered when applying tabulated chemistry in practical combustion system simulation is the simultaneous treatment of complex phenomena such as heat losses, multiple fuel inlets, or dilution by hot gases. In the case of diluted combustion, it was showed that at least five coordinates are required to simulate the impact of fresh gas dilution by recirculating products. Good agreements have been observed between the numerical prediction and measurements in a non-adiabatic diluted combustor, but the generation of this chemical look-up table becomes very sophisticated and CPU expensive, requiring the computation of 100,000 flamelets. Tabulated chemistry encounters a limitation in the empirical selection of manifold coordinates, particularly when a multitude of physical parameters influence the chemical flame structure. Automatized dimension reduction techniques such as local PCA (Principal Component Analysis) is then helpful to automate the definition of tabulated chemistry dimensions. When the dimensionality of the look-up table increases, the tabulation process itself can also pose challenges. It becomes nontrivial to tabulate data that lack a structured organization in the phase space. A viable

solution is to take advantage of supervised non-linear regression algorithms, such as Artificial Neural Networks (ANNs) or Gaussian regression processes. Recent research has also focused on overcoming the traditional limitations of large lookup tables, which scale poorly with dimensionality. These "neural flamelet" or "hybrid tabulation" approaches have enabled coupling between data-driven surrogates and physics-based manifolds, allowing flexible extensions to multi-fuel, multi-regime, or heat-loss-affected combustion. Furthermore, research has highlighted the importance of manifold topology assessment for ensuring that the reduced coordinates provide a unique and smooth mapping to the full thermochemical state. Metrics based on conditional variance, gradient smoothness, or cost-function evaluation have been developed to quantify manifold quality before implementation in CFD solvers. This step is crucial for maintaining model robustness, as poor manifold topology can lead to non-unique regression and instability in reduced-order simulations. Looking forward, the integration of flamelet-like models with uncertainty quantification and active learning represents a promising direction. In such frameworks, the manifold is iteratively refined using high-fidelity data (e.g., DNS or experiments) in regions of high prediction uncertainty. This creates a self-correcting loop between data acquisition and manifold construction. These advances illustrate how the flamelet paradigm is evolving from static tabulation to adaptive, data-enriched reduced-order modeling, ensuring that the low-dimensional manifolds remain accurate, interpretable, and computationally efficient even in complex turbulent reacting flows. Finally, flamelet-based manifolds serve as an essential bridge between high-fidelity chemistry and reduced-order representations. By defining a physically interpretable low-dimensional subspace, they provide an effective starting point for data-driven ROM development and for hybrid methods that combine pre-tabulated physics with online learning or system identification. As such, modern flamelet-like models no longer function solely as static chemistry surrogates, but as adaptable, data-enriched frameworks for real-time and multiscale combustion simulations.

3. Low-dimensional manifold identification

To parametrize the manifold, physics-based variables were originally used relying on expert knowledge. Nowadays, data-driven approaches have gained increased attention to find the optimal control variables of the manifold. The two categories are described in the following sections.

3.1. Physics-based variables

Based on the characteristics of the combustion process under consideration, experts selected specific physics-based variables to define the manifold. The most commonly used variables are listed below, though the list is not exhaustive. Among them, the progress variable (PV) [6]–[8], [17], [23]–[27] is one of the most widely employed, which indicates the progress of the combustion. As explained in Section 2, the progress variable can be either one species' mass fraction or a linear combination of species' mass fractions. For example, the main products and reactants of the combustion are typically used. For hydrogen combustion, hydrogen itself [28] or a combination of hydrogen, water and oxygen [23], [24] is used. In case, one is interested to capture NO_x, nitric oxide (NO) can be used as a second progress variable or combined with other species [17]. Moreover, the progress variable can be associated with the mixture fraction [7], [25], [29]. The mixture fraction [30] indicates the local ratio between fuel and oxidizer and is therefore complementary of the progress variable. Another typical controlling variable of the manifold is the enthalpy, typically used for non-adiabatic systems [6], [8], [25],

[27]. Finally, the dissipation rate is yet another possible variable [29].

3.2. Data-driven approaches

With the increasing computational power and amount of data in the last decades, data-driven approaches have gained prominence [31], [32]. Compared to physics-based approaches, the manifold topology identification is automated and relies on data. Most common methods are principal components analysis (PCA) and the autoencoder discussed in the following paragraphs.

3.2.1. PCA/POD

Principal components analysis (PCA) [33] or proper orthogonal decomposition (POD) [34] is a common technique to reduce the dimension of the data. The technique finds the main directions of variation (principal components) by solving an eigenvalue decomposition of the covariance matrix. To obtain a low-dimensional manifold, the first two or three principal components are selected.

This technique is appealing owing to the limited computational cost compared to other data-driven approaches introduced later. Moreover, the technique does not require a lot of hyperparameter tuning. Only the preprocessing of the input data is needed. But, the main drawback of this technique is that the linear projection is optimal for a linear reconstruction of the original data. Therefore, the projection is not optimal in case a non-linear regression technique can be used to reconstruct the original data. To circumvent this drawback, local PCA variants [35] have been introduced. A typical method is to cluster the data and apply PCA locally on each cluster [36], [37]. Another method to non-linearize the PCA technique is kernel PCA [38]. A kernel function projects the data in a higher dimension and PCA is then applied in that higher dimensional space. However, this variant makes the projection non-linear which is not ideal in case it is used for a ROM as the transformation of the transport equations is not straightforward anymore.

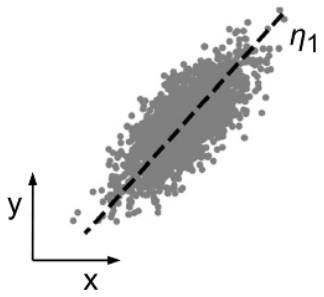

Figure 2: Example of PCA applied on a 2-dimensional dataset with the first principal component indicated [10].

3.2.2. Autoencoders

The autoencoder [39] is a type of neural network architecture that naturally learns in an unsupervised way a low-dimensional representation of the data. An example of this architecture is shown in Figure 3 [10]. The neural network is composed of an encoder, a decoder and a bottleneck in-between. The encoder projects the input onto a reduced space, corresponding to the bottleneck, and the decoder reconstructs the original input data given the reduced

representation. Both parts can be either linear or non-linear [40]. Therefore, through back-propagation during training, the weights of the neural network are optimized to learn the optimal low-dimensional representation of the input data to reconstruct the output variables as accurately as possible.

Compared to PCA, the autoencoder can be seen as a more general method to find a low-dimensional manifold as non-linear projection and reconstruction are possible. Moreover, the neural network is not restricted to reconstructing the input. An arbitrary choice of output variables is possible as opposed to PCA [10], [40].

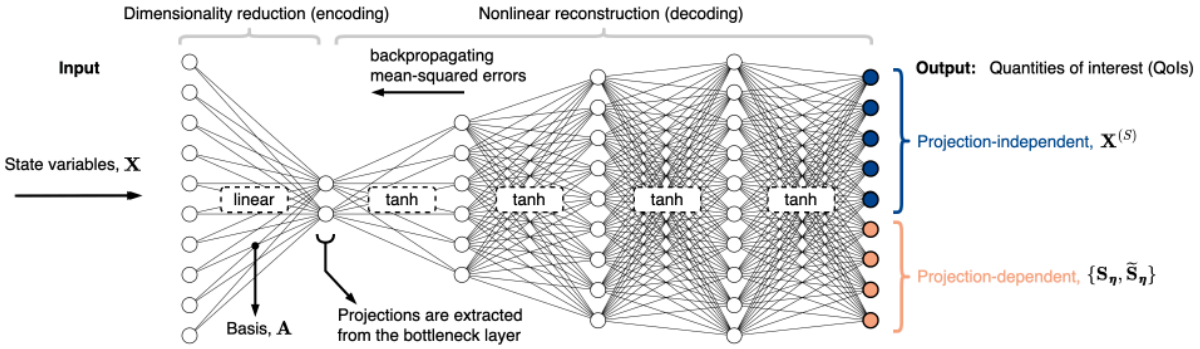

Figure 3: Example of an encoder-decoder architecture for a combustion application. [10] The complete state space is provided as input and the encoder-decoder learns to find a low-dimensional manifold that can optimally reconstruct the important quantities at the ouptut.

3.2.3. Other techniques

Other data-driven approaches for manifold identification are t-distributed stochastic neighbor embedding (t-SNE) [41], isomap [42], uniform manifold approximation and projection (UMAP) [43], locally linear embedding (LLE) [44], diffusion maps [45], [46]... These techniques are less popular for the field of reacting flows due to their non-linear character.

4. Topology optimisation

This section discusses different state-of-the-art methods to get improved manifold topologies. The discussed methods are the progress variable optimization, the optimal estimator, methods to improve PCA projections and the autoencoder.

4.1. Progress variable optimization

Since the progress variable was originally defined from expert knowledge, it constitutes a good candidate for optimization. As Ihme et al. [47] stated, the progress variable should respect some principles. First of all, it should be monotonically increasing to have a one-on-one correspondence with the state space variables. Otherwise, regression becomes challenging. Moreover, the variables should vary smoothly over the progress variable [48] also to facilitate regression afterwards. The other principles that Ihme et al. stated were that the PV should easily be used in a transport equation, that the species included in the PV should have similar timescales and that all manifold parameters should be independent of each other.

Given these constraints and the fact that the PV is a linear combination of the species' mass

fractions, researchers have tried to solve this optimization problem. Ihme et al. [47] was the first one trying to optimize the progress variable with an unconstrained optimization problem. Subsequently, Niu et al. [48] tried based on similar criteria to solve a constrained optimization problem. Vasavan et al. [49], [50] optimized the PV using a multi objective optimization (MOOP) method with a trade-off between monotonicity and reconstruction accuracy of the state space. More recently, Tang et al. [51] and Rahnama et al. [52] employed genetic algorithms to find the optimized PV with the constraint of being monotonic. Zdybał et al. [53] optimized the progress variable using an encoder-decoder with the goal to reconstruct accurately some user-defined quantities from the original state space. This method is further discussed in a following section about autoencoders.

Alternatively, instead of solving an optimization problem, a sensitivity analysis can be performed to find the most suitable PV from a set of predefined PVs [54], [55].

4.2. Optimal estimator

The concept of optimal estimator [25], [56]–[58] is a method to identify the best set of input parameters based on the conditional mean. The method considers two sources of errors: irreducible error and functional error. The first one corresponds to the error inherently present when defining a model with a set of input parameters. The latter corresponds to the error introduced by the functional form of the model, being an approximation of the optimal estimator. Using histograms, neural networks or other regression techniques, they approximate the irreducible error reducing the functional error to a minimum. Comparing then the irreducible error employing different sets of input parameters, it can be attempted to identify the best set of parameters defining the low-dimensional manifold for a given application.

4.3. PCA

To improve the manifold obtained with PCA, several preprocessing steps can be applied [59]. One important preprocessing step is the centering and scaling of the input data prior to applying PCA. There is no universal rule on the optimal centering and scaling. The most common ones, like auto, pareto, VAST, range, level, max, $\langle 0, 1 \rangle$ and $\langle -1, 1 \rangle$, have to be tested [60] on the specific case.

To automate the pre-processing, the cost function [61] can be used by comparing the manifolds obtained from differently pre-processed PCAs. Moreover, using this cost function, the input variables can be automatically selected [60] further improving the manifold representation.

Another preprocessing step is the outlier removal [59] since they can significantly affect the principal components. Finally, to address issues related to data imbalance, kernel density weighting [62] can be applied to ensure a more uniform sampling of the data space.

4.4. Autoencoder

Autoencoders have recently been used for manifold topology optimization for combustion problems. Zdybał et al. [10], [32] introduced an encoder-decoder architecture with a single-layer linear encoder and a non-linear decoder, depicted in Figure 3. The encoder optimizes the linear combination of input species' mass fractions to find the optimal progress variables, driven by the quantities of interests at the output layer. They also introduced projection-dependent variables at the output, such as the progress variable source terms, whose definition

change throughout the optimization. Later, Zdybał et al. [53] applied a modified version of the encoder-decoder architecture on different datasets of ammonia/hydrogen mixtures. In this architecture, fixed mixture fraction is directly used as one of the manifold parameters and only the progress variable is being optimized. This shows that the architecture can be a hybrid combination of a physics-informed and data-driven approach. They also showed an example where the stoichiometric enthalpy defect is used as third manifold parameter for non-adiabatic systems. This illustrates the versatility of the encoder-decoder architecture with tunable input and output variables. Similarly, Perry et al. [63] used the encoder-decoder architecture to co-optimize the manifold parameters and the subfilter closure model for LES. At the input, they used the internal energy and density directly as fixed manifold parameters. Next to these variables, they optimized two progress variables using species' mass fractions at the input. Additionally, the subfilter variances of the progress variables were included at input of the decoder to optimize the subfilter closure model.

For hydrogen flamelets, Armstrong et al. [64] optimized one progress variable alongside mixture fraction or two progress variables using the encoder-decoder architecture. For the decoder, they used the Partition of Unity Network (POUnet) [65] instead of a traditional fully-connected neural network. This special type of network learns a set of local regression models, *e.g.* RBF networks, and offers a high prediction accuracy [66]. Castellanos et al. [67] used a variant of the autoencoder to construct a reduced space better suited for reduced-order modeling. They applied a time-lag autoencoder where there is a time shift between the input and output. Scherding et al. [68] showed a successful implementation of the autoencoder for a hypersonic flow in chemical non-equilibrium to reduce the number of transport equations. A final application of the autoencoder is given by Rubini et al. [69] introducing the framework "ChemZIP" to perform ROM simulations of turbomachinery reactors. In this work, temperature and pressure are used next to two optimized progress variables to define the low-dimensional manifold.

5. Tools to assess the topology of low-dimensional manifolds

To assess *a priori* the quality of a manifold topology, we identify two families of metrics. The first one focuses on the structural correspondence between the low-dimensional and high-dimensional manifolds [70]–[77]. These methods assess whether local and global structure in the original space are preserved in the low-dimensional manifold. However, good performance on such metrics does not necessarily guarantee the success of the resulting reduced-order model. More recently, a second family of approaches has been introduced to assess the manifold quality. Instead of focusing on structural correspondence, these methods evaluate how well user-defined quantities of interest (QoIs) are represented in the low-dimensional manifold [61], [78]. They measure how each QoI varies in the manifold at different scales. Thus, the method penalizes manifold defects such as steep gradients and non-uniqueness, which make regression difficult or even impossible, while rewarding properties like smoother gradients and enlarged manifolds. In contrast to the first family, these metrics provide an indication of how easily the QoIs can be regressed from the low-dimensional manifold. Therefore, a good performance on these metrics suggest the potential for a successful reduced-order model.

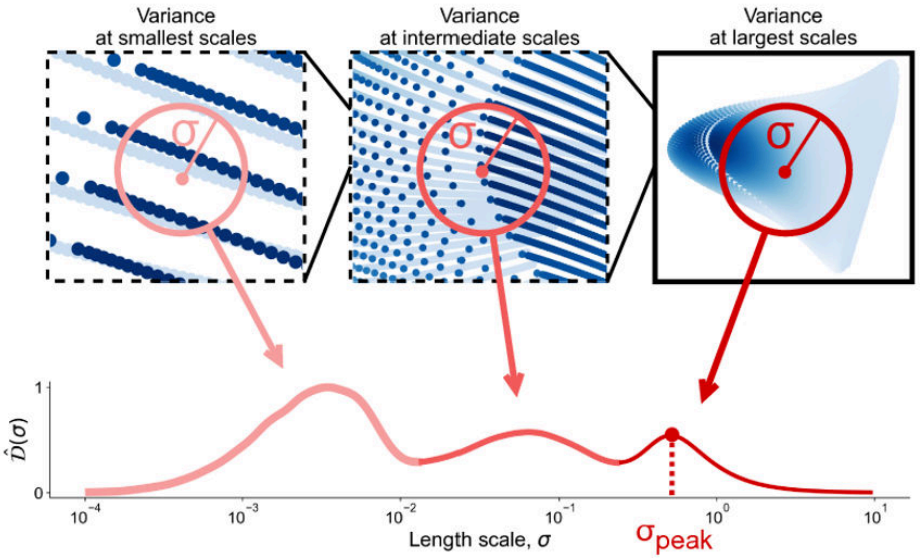

Figure 4: Illustration of how the manifold topology is assessed with the tool of Armstrong et al. [78]. The variance of the QoI is measured at different scales [61] giving an evolution of $\hat{D}$ at different length scales σ.

6. Successful ROMs employing a low-dimensional manifold

This section presents applications of the previously introduced manifold topology optimization techniques to reduced-order models (ROMs) that operate on low-dimensional manifolds.

6.1. PCA/POD

Sutherland et Parente [9], [36] introduced the concept of using PCA for combustion modeling. Since then, many researchers have used PCA to perform ROM simulations on combustion cases. As explained previously, PCA offers an optimal linear dimension reduction. However, non-linear regression techniques are now used to reconstruct the original state space increasing the accuracy of the ROM [37], [79]. PCA can be combined with any regression technique like linear regression [80], Kringing [81]–[85], neural networks [80], [85]–[87], polynomial regression [85], support vector regression [80], [85], k-neirest neighbors (kNN) [85], gaussian process regression (GPR) [37], [80], multivariate adaptive regression spline (MARS) [88], [89]... With this, PCA has been applied successfully as ROM for many applications, *e.g.* LES [37], [88], flame-wall interactions [87], 1D flame [81], [90], [91], 2D flame [81], [90], 3D DNS [92], plasma [93], oil reservoir simulations [94]...

6.2. Autencoder

There are a few examples of successful ROMs employing a manifold obtained from an autoencoder. In section 4.4, cases of manifold optimization with the autoencoder were discussed. Here, they are further explained on how they applied the optimized manifold on a ROM. Zdybał et al. [53] optimized the progress variable using the encoder-decoder architecture. Then, they performed a ROM simulation for one trajectory of an autoignition case. A separate neural network was trained to predict the PV source term. They found out that the optimized progress variable performed better than traditional PVs defined by experts owing to the reduced non-uniqueness. Similarly, Castellanos et al. [67] applied the time-lag au-

toencoder on a hydrogen autoignition case which performed better than PCA. Armstrong et al. [64] tested the optimized progress variables on 1D and 2D DNS simulations and showed higher accuracy with the POUnet than traditional physics-based models. Rubini et al. [69] performed the simulation of a turbo-reactor using an encoder-decoder. AN additional neural network was used to predict the progress variable and temperature source terms. Finally, Scherding et al. [68] used the encoder-decoder to optimize the progress variables for a hypersonic flat-plate boundary layer and shock-wave boundary layer interaction with chemical non-equilibrium. Next, they used random forest to cluster the data and a radial basis function network to predict back the original state space. They reduced with 70% the CPU time while keeping a high accuracy compared to the true solution.

6.3. Other manifold techniques

Here, examples of ROMs are presented using less common techniques of manifold identification. Franz et al. [95] used isomap to construct the manifold for a ROM simulation of a transonic flow over an airfoil. They compared it with traditional POD methods and the projection-based ROM using the isomap was shown to be more accurate. Similarly, Zheng et al. [96] used local linear embedding also applied on a transonic flow. Bykov et al. performed ROM simulations based on the reaction-diffusion manifold (REDIM) for 1D and 2D flames showing that the REDIM could be extended to more complicated cases. Yu et al. developed also REDIM models. In this work, the model was tested on steady and transient counterflow flames with a focus on the NO_x prediction.

7. Conclusion

The use of low-dimensional manifolds to perform ROM simulations of reacting flows has shown promising results in the past decades. Originally, manifolds were defined based on expert knowledge. Nowadays, with the advance of data-driven techniques, PCA and autoencoders have become popular tools to define and optimize manifold topologies. PCA has the advantage of being computationally inexpensive and requiring minimal tuning. However, the technique is not optimal for non-linear reconstruction of the original state space. On the other hand, autoencoders offer a high flexibility in architecture, but are computationally demanding and require careful hyperparameter optimization.

In recent years, many applications of both methods can be found to speed up simulations of combustion systems. A path forward in the development of ROMs employing a low-dimensional manifold would be the development of a general framework performing the dimension reduction and the ROM simulation simultaneously.

Acknowledgments

This deliverable is based upon work from COST Action CA22151 – CYPHER, supported by COST (European Cooperation in Science and Technology). COST is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation. For more information, visit `www.cost.eu`.

References

[1] P. Trisjono and H. Pitsch, "Systematic Analysis Strategies for the Development of Combustion Models from DNS: A Review," en, *Flow, Turbulence and Combustion*, vol. 95, no. 2, pp. 231–259, Oct. 2015, ISSN: 1573-1987.

[2] S. B. Pope, "Small scales, many species and the manifold challenges of turbulent combustion," *Proceedings of the Combustion Institute*, vol. 34, no. 1, pp. 1–31, Jan. 2013, ISSN: 1540-7489.

[3] U. Maas and S. B. Pope, "Implementation of simplified chemical kinetics based on intrinsic low-dimensional manifolds," *Symposium (International) on Combustion*, Twenty-Fourth Symposium on Combustion, vol. 24, no. 1, pp. 103–112, Jan. 1992, ISSN: 0082-0784.

[4] U. Maas and S. B. Pope, "Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space," *Combustion and Flame*, vol. 88, no. 3, pp. 239–264, Mar. 1992, ISSN: 0010-2180.

[5] U. Maas and D. Thévenin, "Correlation analysis of direct numerical simulation data of turbulent non-premixed flames," *Symposium (International) on Combustion*, Twenty-Seventh Sysposium (International) on Combustion Volume One, vol. 27, no. 1, pp. 1183–1189, Jan. 1998, ISSN: 0082-0784.

[6] J. V. OIJEN and L. D. GOEY, "Modelling of premixed laminar flames using flamelet-generated manifolds," *Combustion Science and Technology*, vol. 161, no. 1, pp. 113–137, Dec. 2000.

[7] P.-D. Nguyen, L. Vervisch, V. Subramanian, and P. Domingo, "Multidimensional flamelet-generated manifolds for partially premixed combustion," *Combustion and Flame*, vol. 157, no. 1, pp. 43–61, Jan. 2010, ISSN: 0010-2180.

[8] Y. Luo, M. Steinhausen, D. Kaddar, C. Hasse, and F. Ferraro, "Assessment of flamelet manifolds for turbulent flame-wall interactions in large-eddy simulations," *Combustion and Flame*, vol. 255, p. 112 923, Sep. 2023, ISSN: 0010-2180.

[9] J. C. Sutherland and A. Parente, "Combustion modeling using principal component analysis," *Proceedings of the Combustion Institute*, vol. 32, no. 1, pp. 1563–1570, Jan. 2009, ISSN: 1540-7489.

[10] K. Zdybał, A. Parente, and J. C. Sutherland, "Improving reduced-order models through nonlinear decoding of projection-dependent outputs," en, *Patterns*, vol. 4, no. 11, p. 100 859, Nov. 2023, ISSN: 26663899.

[11] G. Sorrentino, G. B. Ariemma, F. Ferraro, and B. Fiorina, "Including detailed chemistry features in the modeling of emerging low-temperature reactive flows: A review on the application to diluted and mild combustion systems," *Applications in Energy and Combustion Science*, vol. 20, p. 100 291, 2024.

[12] M. Ihme and H. Pitsch, "Modeling of radiation and nitric oxide formation in turbulent nonpremixed flames using a flamelet/progress variable formulation," *Physics of Fluids*, vol. 20, no. 5, p. 055 110, May 2008, ISSN: 1070-6631.

[13] N. Peters, "Laminar diffusion flamelet models in non-premixed turbulent combustion," *Progress in Energy and Combustion Science*, vol. 10, no. 3, pp. 319–339, Jan. 1984, ISSN: 0360-1285.

[14] C. Yu, P. Shrotriya, X. Li, and U. Maas, "Reduced modeling of the NOx formation based on the reaction-diffusion manifolds method for counterflow diffusion flames," *Proceedings of the Combustion Institute*, vol. 39, no. 2, pp. 1587–1596, Jan. 2023, ISSN: 1540-7489.

[15] V. Bykov and U. Maas, "Problem adapted reduced models based on Reaction–Diffusion Manifolds (REDIMs)," *Proceedings of the Combustion Institute*, vol. 32, no. 1, pp. 561–568, Jan. 2009, ISSN: 1540-7489.

[16] O. Gicquel, N. Darabiha, and D. Thévenin, "Liminar premixed hydrogen/air counterflow flame simulations using flame prolongation of ILDM with differential diffusion," *Proceedings of the Combustion Institute*, vol. 28, no. 2, pp. 1901–1908, Jan. 2000, ISSN: 1540-7489.

[17] J. A. van Oijen, A. Donini, R. J. M. Bastiaans, J. H. M. ten Thije Boonkkamp, and L. P. H. de Goey, "State-of-the-art in premixed combustion modeling using flamelet generated manifolds," *Progress in Energy and Combustion Science*, vol. 57, pp. 30–74, Nov. 2016, ISSN: 0360-1285.

[18] W. Ramaekers, "Development of flamelet generated manifolds for partially-premixed flame simulations," ISBN: 9789038627502, Phd Thesis 1 (Research TU/e / Graduation TU/e), Technische Universiteit Eindhoven, Eindhoven, 2011.

[19] A. Najafi-Yazdi, B. Cuenot, and L. Mongeau, "Systematic definition of progress variables and Intrinsically Low-Dimensional, Flamelet Generated Manifolds for chemistry tabulation," *Combustion and Flame*, vol. 159, no. 3, pp. 1197–1204, Mar. 2012, ISSN: 0010-2180.

[20] M. Ihme, C. Schmitt, and H. Pitsch, "Optimal artificial neural networks and tabulation methods for chemistry representation in LES of a bluff-body swirl-stabilized flame," *Proceedings of the Combustion Institute*, vol. 32, no. 1, pp. 1527–1535, Jan. 2009, ISSN: 1540-7489.

[21] Y. Zhang, S. Xu, S. Zhong, X.-S. Bai, H. Wang, and M. Yao, "Large eddy simulation of spray combustion using flamelet generated manifolds combined with artificial neural networks," *Energy and AI*, vol. 2, p. 100 021, Nov. 2020, ISSN: 2666-5468.

[22] K. Li, P. Rahnama, R. Novella, and B. Somers, "Combining flamelet-generated manifold and machine learning models in simulation of a non-premixed diffusion flame," *Energy and AI*, vol. 14, p. 100 266, Oct. 2023, ISSN: 2666-5468.

[23] X. Wen, L. Berger, L. Cai, A. Parente, and H. Pitsch, "Thermodiffusively unstable laminar hydrogen flame in a sufficiently large 3D computational domain – Part I: Characteristic patterns," en, *Combustion and Flame*, vol. 263, p. 113 278, May 2024, ISSN: 00102180.

[24] X. Wen, L. Berger, L. Cai, A. Parente, and H. Pitsch, "Thermodiffusively unstable laminar hydrogen flame in a sufficiently large 3D computational domain – Part II: NOx formation mechanism and flamelet modeling," en, *Combustion and Flame*, vol. 265, p. 113 497, Jul. 2024, ISSN: 00102180.

[25] H. Böttler, D. Kaddar, T. J. P. Karpowski, *et al.*, "Can flamelet manifolds capture the interactions of thermo-diffusive instabilities and turbulence in lean hydrogen flames?—An a-priori analysis," *International Journal of Hydrogen Energy*, vol. 56, pp. 1397–1407, Feb. 2024, ISSN: 0360-3199.

[26] J. A. van Oijen, R. J. M. Bastiaans, and L. P. H. de Goey, "Low-dimensional manifolds in direct numerical simulations of premixed turbulent flames," *Proceedings of the Combustion Institute*, vol. 31, no. 1, pp. 1377–1384, Jan. 2007, ISSN: 1540-7489.

[27] J. A. van Oijen, F. A. Lammers, and L. P. H. de Goey, "Modeling of complex premixed burner systems by using flamelet-generated manifolds," *Combustion and Flame*, vol. 127, no. 3, pp. 2124–2134, Nov. 2001, ISSN: 0010-2180.

[28] L. Berger, A. Attili, J. Wang, K. Maeda, and H. Pitsch, "Development of Large-Eddy Simulation Combustion Models for Thermodiffusive Instabilities in Turbulent Hydrogen Flames," en, *Proceeding of the Summer Program, Center for Turbulence Research, Stanford University*, p. 247, 2022.

[29] C. E. Lacey, A. G. Novoselov, and M. E. Mueller, "In-Situ Adaptive Manifolds: Enabling computationally efficient simulations of complex turbulent reacting flows," *Proceedings of the Combustion Institute*, vol. 38, no. 2, pp. 2673–2680, Jan. 2021, ISSN: 1540-7489.

[30] R. W. Bilger, S. H. Stårner, and R. J. Kee, "On reduced mechanisms for methane air combustion in nonpremixed flames," *Combustion and Flame*, vol. 80, no. 2, pp. 135–149, May 1990, ISSN: 0010-2180.

[31] R. Vinuesa and S. L. Brunton, "Enhancing computational fluid dynamics with machine learning," en, *Nature Computational Science*, vol. 2, no. 6, pp. 358–366, Jun. 2022, Publisher: Nature Publishing Group, ISSN: 2662-8457.

[32] K. Zdybał, M. R. Malik, A. Coussement, J. C. Sutherland, and A. Parente, "Reduced-order modeling of reacting flows using data-driven approaches," in *Machine Learning and Its Application to Reacting Flows: ML and Combustion*, Springer International Publishing Cham, 2023, pp. 245–278.

[33] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, no. 6, pp. 417–441, 1933, Place: US Publisher: Warwick & York, ISSN: 1939-2176.

[34] A. Chatterjee, "An introduction to the proper orthogonal decomposition," *Current Science*, vol. 78, no. 7, pp. 808–817, 2000, Publisher: Temporary Publisher, ISSN: 0011-3891.

[35] N. Kambhatla and T. K. Leen, "Dimension Reduction by Local Principal Component Analysis," *Neural Computation*, vol. 9, no. 7, pp. 1493–1516, Jul. 1997, ISSN: 0899-7667.

[36] A. Parente, J. C. Sutherland, L. Tognotti, and P. J. Smith, "Identification of low-dimensional manifolds in turbulent flames," *Proceedings of the Combustion Institute*, vol. 32, no. 1, pp. 1579–1586, Jan. 2009, ISSN: 1540-7489.

[37] M. R. Malik, B. J. Isaac, A. Coussement, P. J. Smith, and A. Parente, "Principal component analysis coupled with nonlinear regression for chemistry reduction," *Combustion and Flame*, vol. 187, pp. 30–41, Jan. 2018, ISSN: 0010-2180.

[38] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," en, in *Artificial Neural Networks — ICANN'97*, W. Gerstner, A. Germond, M. Hasler, and J.-D. Nicoud, Eds., Berlin, Heidelberg: Springer, 1997, pp. 583–588, ISBN: 978-3-540-69620-9.

[39] I. D. Mienye and T. G. Swart, "Deep Autoencoder Neural Networks: A Comprehensive Review and New Perspectives," en, *Archives of Computational Methods in Engineering*, Mar. 2025, ISSN: 1886-1784.

[40] P. Baldi, "Autoencoders, Unsupervised Learning, and Deep Architectures," en, in *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, ISSN: 1938-7228, JMLR Workshop and Conference Proceedings, Jun. 2012, pp. 37–49.

[41] L. v. d. Maaten and G. Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008, ISSN: 1533-7928.

[42] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000, Publisher: American Association for the Advancement of Science.

[43] L. McInnes, J. Healy, and J. Melville, *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*, arXiv:1802.03426 [stat], Sep. 2020.

[44] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," EN, *Science*, Dec. 2000, Publisher: American Association for the Advancement of Science.

[45] R. R. Coifman, S. Lafon, A. B. Lee, *et al.*, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proceedings of the National Academy of Sciences*, vol. 102, no. 21, pp. 7426–7431, May 2005, Publisher: Proceedings of the National Academy of Sciences.

[46] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, Special Issue: Diffusion Maps and Wavelets, vol. 21, no. 1, pp. 5–30, Jul. 2006, ISSN: 1063-5203.

[47] M. Ihme, L. Shunn, and J. Zhang, "Regularization of reaction progress variable for application to flamelet-based combustion models," *Journal of Computational Physics*, vol. 231, no. 23, pp. 7715–7721, Oct. 2012, ISSN: 0021-9991.

[48] Y.-S. Niu, L. Vervisch, and P. D. Tao, "An optimization-based approach to detailed chemistry tabulation: Automated progress variable definition," *Combustion and Flame*, vol. 160, no. 4, pp. 776–785, Apr. 2013, ISSN: 0010-2180.

[49] A. Vasavan, "Numerical modeling of igniting non-premixed combustion systems using fgm," Ph.D. dissertation, Eindhoven University of Technology, 2021.

[50] A. Vasavan, P. de Goey, and J. van Oijen, "A novel method to automate FGM progress variable with application to igniting combustion systems," *Combustion Theory and Modelling*, vol. 24, no. 2, pp. 221–244, Mar. 2020, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/13647830.2019.1673902, ISSN: 1364-7830.

[51] T. Tang, Z. Wang, H. Li, *et al.*, "A method for optimizing reaction progress variable and its application," *Aerospace Science and Technology*, vol. 130, p. 107 888, Nov. 2022, ISSN: 1270-9638.

[52] P. Rahnama, A. Maghbouli, H. Bao, A. Vasavan, R. Novella, and B. Somers, "Generalizing progress variable definition in CFD simulation of combustion systems using tabulated chemistry models," *Applications in Energy and Combustion Science*, vol. 14, p. 100 132, Jun. 2023, ISSN: 2666-352X.

[53] K. Zdybał, J. C. Sutherland, and A. Parente, "Optimizing progress variables for ammonia/hydrogen combustion using encoding–decoding networks," *Combustion and Flame*, vol. 276, p. 114 152, Jun. 2025, ISSN: 0010-2180.

[54] H. Gupta, O. J. Teerling, and J. A. van Oijen, "Effect of progress variable definition on the mass burning rate of premixed laminar flames predicted by the Flamelet Generated Manifold method," *Combustion Theory and Modelling*, vol. 25, no. 4, pp. 631–645, Jun. 2021, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/13647830.2021.1926544, ISSN: 1364-7830.

[55] V. S. Wehrmann, N. Chakraborty, M. Klein, and J. Hasslberger, "Choice of reaction progress variable under preferential diffusion effects in turbulent syngas combustion based on detailed chemistry direct numerical simulations," en, *Scientific Reports*, vol. 14, no. 1, p. 14 861, Jun. 2024, Publisher: Nature Publishing Group, ISSN: 2045-2322.

[56] A. Moreau, O. Teytaud, and J. P. Bertoglio, "Optimal estimation for large-eddy simulation of turbulence and application to the analysis of subgrid models," *Physics of Fluids*, vol. 18, no. 10, p. 105 101, Oct. 2006, ISSN: 1070-6631.

[57] L. Berger, K. Kleinheinz, A. Attili, F. Bisetti, H. Pitsch, and M. E. Mueller, "Numerically accurate computational techniques for optimal estimator analyses of multi-parameter models," *Combustion Theory and Modelling*, vol. 22, no. 3, pp. 480–504, May 2018, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/13647830.2018.1424353, ISSN: 1364-7830.

[58] L. Berger, A. Attili, M. Gauding, and H. Pitsch, "LES combustion model for premixed turbulent hydrogen flames with thermodiffusive instabilities: A priori and a posteriori analysis," en, *Journal of Fluid Mechanics*, vol. 1003, A33, Jan. 2025, ISSN: 0022-1120, 1469-7645.

[59] A. Parente and J. C. Sutherland, "Principal component analysis of turbulent combustion data: Data pre-processing and manifold sensitivity," *Combustion and Flame*, vol. 160, no. 2, pp. 340–350, Feb. 2013, ISSN: 0010-2180.

[60] K. Zdybał, J. C. Sutherland, and A. Parente, "Manifold-informed state vector subset for reduced-order modeling," en, *Proceedings of the Combustion Institute*, vol. 39, no. 4, pp. 5145–5154, 2023, ISSN: 15407489.

[61] K. Zdybał, E. Armstrong, J. C. Sutherland, and A. Parente, "Cost function for low-dimensional manifold topology assessment," en, *Scientific Reports*, vol. 12, no. 1, p. 14 496, Aug. 2022, ISSN: 2045-2322.

[62] A. Coussement, O. Gicquel, and A. Parente, "Kernel density weighted principal component analysis of combustion processes," *Combustion and Flame*, vol. 159, no. 9, pp. 2844–2855, Sep. 2012, ISSN: 0010-2180.

[63] B. A. Perry, M. T. Henry de Frahan, and S. Yellapantula, "Co-optimized machine-learned manifold models for large eddy simulation of turbulent combustion," *Combustion and Flame*, vol. 244, p. 112 286, Oct. 2022, ISSN: 0010-2180.

[64] E. Armstrong and J. C. Sutherland, "Reduced-order modeling with reconstruction-informed projections," en, *Combustion and Flame*, vol. 259, p. 113 119, Jan. 2024, ISSN: 00102180.

[65] K. Lee, N. A. Trask, R. G. Patel, M. A. Gulian, and E. C. Cyr, *Partition of unity networks: Deep hp-approximation*, arXiv:2101.11256 [cs], Jan. 2021.

[66] E. Armstrong, M. A. Hansen, R. C. Knaus, N. A. Trask, J. C. Hewson, and J. C. Sutherland, "Accurate Compression of Tabulated Chemistry Models with Partition of Unity Networks," *Combustion Science and Technology*, vol. 196, no. 6, pp. 850–867, Apr. 2024, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00102202.2022.2102908, ISSN: 0010-2202.

[67] L. Castellanos, R. S. M. Freitas, A. Parente, and F. Contino, "Deep learning dynamical latencies for the analysis and reduction of combustion chemistry kinetics," *Physics of Fluids*, vol. 35, no. 10, p. 107 143, Oct. 2023, ISSN: 1070-6631.

[68] C. Scherding, G. Rigas, D. Sipp, P. J. Schmid, and T. Sayadi, "Data-driven framework for input/output lookup tables reduction: Application to hypersonic flows in chemical nonequilibrium," *Physical Review Fluids*, vol. 8, no. 2, p. 023 201, Feb. 2023, Publisher: American Physical Society.

[69] D. Rubini and B. Rosic, "ChemZIP: A data-driven accelerated modeling framework for complex aerothermochemical interactions in novel turbomachinery reactors," *Chemical Engineering Journal*, vol. 515, p. 163 240, Jul. 2025, ISSN: 1385-8947.

[70] J. Venna and S. Kaski, "Local multidimensional scaling," *Neural Networks*, Advances in Self Organising Maps - WSOM'05, vol. 19, no. 6, pp. 889–899, Jul. 2006, ISSN: 0893-6080.

[71] J. Lee and M. Verleysen, "Quality assessment of nonlinear dimensionality reduction based on K-ary neighborhoods," en, in *Proceedings of the Workshop on New Challenges for Feature Selection in Data Mining and Knowledge Discovery at ECML/PKDD 2008*, ISSN: 1938-7228, PMLR, Sep. 2008, pp. 21–35.

[72] J. A. Lee, M. Verleysen, *et al.*, "Rank-based quality assessment of nonlinear dimensionality reduction.," in *ESANN*, 2008, pp. 49–54.

[73] J. A. Lee and M. Verleysen, "Quality assessment of dimensionality reduction: Rank-based criteria," *Neurocomputing*, Advances in Machine Learning and Computational Intelligence, vol. 72, no. 7, pp. 1431–1443, Mar. 2009, ISSN: 0925-2312.

[74] L. Chen and A. Buja, "Local Multidimensional Scaling for Nonlinear Dimension Reduction, Graph Drawing, and Proximity Analysis," EN, *Journal of the American Statistical Association*, Mar. 2009, Publisher: Taylor & Francis.

[75] Y. Goldberg and Y. Ritov, "Local procrustes for manifold embedding: A measure of embedding quality and embedding algorithms," en, *Machine Learning*, vol. 77, no. 1, pp. 1–25, Oct. 2009, ISSN: 1573-0565.

[76] P. Zhang, Y. Ren, and B. Zhang, "A new embedding quality assessment method for manifold learning," *Neurocomputing*, vol. 97, pp. 251–266, Nov. 2012, ISSN: 0925-2312.

[77] S. Boral, M. Sarkar, and A. Ghosh, "MEQA: Manifold embedding quality assessment via anisotropic scaling and Kolmogorov-Smirnov test," *Pattern Recognition*, vol. 139, p. 109 447, Jul. 2023, ISSN: 0031-3203.

[78] E. Armstrong and J. C. Sutherland, "A technique for characterising feature size and quality of manifolds," *Combustion Theory and Modelling*, vol. 25, no. 4, pp. 646–668, Jun. 2021, Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/13647830.2021.1931715, ISSN: 1364-7830.

[79] P. Benner, S. Gugercin, and K. Willcox, "A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems," en, *SIAM Review*, Nov. 2015, Publisher: Society for Industrial and Applied Mathematics.

[80] B. J. Isaac, J. N. Thornock, J. Sutherland, P. J. Smith, and A. Parente, "Advanced regression methods for combustion modelling using principal components," *Combustion and Flame*, vol. 162, no. 6, pp. 2592–2601, Jun. 2015, ISSN: 0010-2180.

[81] G. Aversano, A. Bellemans, Z. Li, A. Coussement, O. Gicquel, and A. Parente, "Application of reduced-order models based on PCA & Kriging for the development of digital twins of reacting flow applications," *Computers & Chemical Engineering*, vol. 121, pp. 422–441, Feb. 2019, ISSN: 0098-1354.

[82] G. Aversano, G. D'Alessio, A. Coussement, F. Contino, and A. Parente, "Combination of polynomial chaos and Kriging for reduced-order model of reacting flow applications," *Results in Engineering*, vol. 10, p. 100 223, Jun. 2021, ISSN: 2590-1230.

[83] G. Aversano, M. Ferrarotti, and A. Parente, "Digital twin of a combustion furnace operating in flameless conditions: Reduced-order model development from CFD simulations," *Proceedings of the Combustion Institute*, vol. 38, no. 4, pp. 5373–5381, Jan. 2021, ISSN: 1540-7489.

[84] G. Aversano, J. C. Parra-Alvarez, B. J. Isaac, *et al.*, "PCA and Kriging for the efficient exploration of consistency regions in Uncertainty Quantification," *Proceedings of the Combustion Institute*, vol. 37, no. 4, pp. 4461–4469, Jan. 2019, ISSN: 1540-7489.

[85] C. (Ni, S. (Ding, J. (Li, X. (Chu, Z. (Ren, and X. (Wang, "Projection-based reduced order modeling of multi-species mixing and combustion," *Physics of Fluids*, vol. 36, no. 7, p. 077 168, Jul. 2024, ISSN: 1070-6631.

[86] H. Mirgolbabaei and T. Echekki, "A novel principal component analysis-based acceleration scheme for LES–ODT: An *a priori* study," *Combustion and Flame*, vol. 160, no. 5, pp. 898–908, May 2013, ISSN: 0010-2180.

[87] J. Bissantz, J. Karpowski, M. Steinhausen, *et al.*, "Application of dense neural networks for manifold-based modeling of flame-wall interactions," *Applications in Energy and Combustion Science*, vol. 13, p. 100 113, Mar. 2023, ISSN: 2666-352X.

[88] A. Biglari and J. C. Sutherland, "A filter-independent model identification technique for turbulent combustion modeling," *Combustion and Flame*, vol. 159, no. 5, pp. 1960–1970, May 2012, ISSN: 0010-2180.

[89] Y. Yang, S. B. Pope, and J. H. Chen, "Empirical low-dimensional manifolds in composition space," *Combustion and Flame*, vol. 160, no. 10, pp. 1967–1980, Oct. 2013, ISSN: 0010-2180.

[90] A. Coussement, B. J. Isaac, O. Gicquel, and A. Parente, "Assessment of different chemistry reduction methods based on principal component analysis: Comparison of the MG-PCA and score-PCA approaches," *Combustion and Flame*, vol. 168, pp. 83–97, Jun. 2016, ISSN: 0010-2180.

[91] M. R. Malik, P. Obando Vega, A. Coussement, and A. Parente, "Combustion modeling using Principal Component Analysis: A posteriori validation on Sandia flames D, E and F," *Proceedings of the Combustion Institute*, vol. 38, no. 2, pp. 2635–2643, Jan. 2021, ISSN: 1540-7489.

[92] M. R. Malik, R. Khamedov, F. E. Hernández Pérez, A. Coussement, A. Parente, and H. G. Im, "Dimensionality reduction and unsupervised classification for high-fidelity reacting flow simulations," *Proceedings of the Combustion Institute*, vol. 39, no. 4, pp. 5155–5163, Jan. 2023, ISSN: 1540-7489.

[93] A. Bellemans, T. Magin, A. Coussement, and A. Parente, "Reduced-order kinetic plasma models using principal component analysis: Model formulation and manifold sensitivity," *Physical Review Fluids*, vol. 2, no. 7, p. 073 201, Jul. 2017, Publisher: American Physical Society.

[94] S. Le Clainche, F. Varas, and J. M. Vega, "Accelerating oil reservoir simulations using POD on the fly," en, *International Journal for Numerical Methods in Engineering*, vol. 110, no. 1, pp. 79–100, 2017, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.5356, ISSN: 1097-0207.

[95] T. Franz, R. Zimmermann, S. Görtz, and N. Karcher, "Interpolation-based reduced-order modelling for steady transonic flows via manifold learning," EN, *International Journal of Computational Fluid Dynamics*, Mar. 2014, Publisher: Taylor & Francis, ISSN: 1061-8562.

[96] B. Zheng, W. Yao, and M. Xu, "Nonlinear Manifold Learning and Model Reduction for Transonic Flows," en, *AIAA Journal*, Sep. 2023, Publisher: American Institute of Aeronautics and Astronautics.