Combustion and Flame 260 (2024) 113221

Contents lists available at ScienceDirect

Combustion and Flame

journal homepage: www.sciencedirect.com/journal/combustion-and-flame

An enhanced Sample-Partitioning Adaptive Reduced Chemistry method with a-priori error estimation

Pietro Pagani [a,b,*], Riccardo Malpica Galassi [b,c], Ruggero Amaduzzi [b,c], Alessandro Parente [b,c], Francesco Contino [a]

[a] Institute of Mechanics, Materials and Civil Engineering (iMMC), Université Catholique de Louvain (UCLouvain), Place du Levant, 2, 1348 Louvain-la-Neuve, Belgium
[b] Université Libre de Bruxelles, École Polytechnique de Bruxelles, Aero-Thermo-Mechanics Laboratory, Brussels, Belgium
[c] Université Libre de Bruxelles and Vrije Universiteit Brussel, Brussels Institute for Thermal-Fluid Systems and Clean Energy (BRITE), Brussels, Belgium

ARTICLE INFO

Keywords:
Reduced order modeling
Adaptive chemistry
Clustering
Data-driven method
Mild combustion

ABSTRACT

Reactor-based approaches for handling the Turbulence–Chemistry-Interactions closure have the advantage of embedding finite-rate chemistry in the combustion model of RANS and LES simulations, which might be crucial for the solution accuracy when complex combustion regimes are investigated. However, the numerical solution of the chemical ODEs is burdened with stiffness and increased dimensionality, especially when large detailed mechanisms are required. To this end, the Sample-Partitioning Adaptive Reduced Chemistry (SPARC) methodology couples adaptive chemistry and machine learning to speed-up the chemistry integration in reactive flows simulations. It consists in building a library of skeletal mechanisms, associated to clusters of similar thermo-chemical states identified in a training dataset, and then, at run-time, each computational cell is assigned to a specific cluster, whose skeletal mechanism is retrieved and employed for the time integration. Such workflow builds on four interacting *blocks*, i.e., training dataset generation, clustering, mechanism simplification, and classification, and its success tightly depends on the quality of each block, which generally results from a combination of theoretical, methodological, and computational choices. In this paper, we explore the effects of the mechanism simplification strategy on the SPARC performance and we develop an ad-hoc procedure that automatically identifies the cluster-wise optimal reduction parameters, delivering a higher global reduction and therefore a larger computational speed-up compared to a standard approach, along with an explicit a-priori control on accuracy. We implement and test this novel procedure on a RANS simulation of the Adelaide Jet-in-hot-coflow (AJHC) burner, and we attain a ∼2x CPU time improvement with respect to the simulation obtained with a 36-species detailed mechanism.

Novelty and significance This work makes a contribution towards the acceleration of chemistry integration in reactive flows simulations. More specifically, the Sample-Partitioning Adaptive Reduced Chemistry method, which couples adaptive chemistry and machine learning, is enhanced with automatic target species definition and a-priori error estimation. The novelties lie in the utilization of the computational singular perturbation (CSP) reduction algorithm, which provides means for automatically identifying an adaptive set of target species, and in the definition of a novel strategy for assessing the performance of the reduced mechanisms in the pre-processing phase, with the goal of estimating a measure of accuracy of the upcoming CFD simulation.

1. Introduction

Several solutions have been proposed in the past years to fulfill the demand for heat and power in the industry while complying with today's emissions regulation. Among them, the Moderate or Intense Low-oxygen Dilution (MILD) combustion showed to be promising due to a high combustion efficiency combined with low pollutant emissions [1]. The robust design of practical combustion devices operating in MILD regime will require the availability of reliable and affordable CFD tools.

The typically large scale of realistic devices, combined with the need for outer-loop applications such as optimizers, which iterate the

* Corresponding author at: Institute of Mechanics, Materials and Civil Engineering (iMMC), Université Catholique de Louvain (UCLouvain), Place du Levant, 2, 1348 Louvain-la-Neuve, Belgium.
E-mail address: pietro.pagani@uclouvain.be (P. Pagani).

https://doi.org/10.1016/j.combustflame.2023.113221
Received 23 May 2023; Received in revised form 27 November 2023; Accepted 27 November 2023
Available online 8 December 2023

design variables over multiple model evaluations, confines the range of applicability of numerical approaches to Reynolds-averaged Navier–Stokes (RANS) simulations, with Large Eddy Simulations (LES) possibly employed in a later stage of development. Due to the highly non-linear relationship between species concentrations and their formation/consumption rate, the chemical source terms cannot be directly computed from the resolved thermo-chemical variables, thus turbulence-chemistry interaction (TCI) closures are needed in RANS and LES approaches.

However, modeling TCI closures in reacting flows which exhibit the MILD combustion regime brings in additional difficulties on top of the already demanding modeling requirements of traditional regimes, because of the strong coupling between turbulent mixing and chemical kinetics. Examples of flamelet-progress variables approaches [2] that proved to work well under MILD conditions are present in literature, e.g. [3–6], in which the flamelets include the effects of strong dilution by products of the fuel/oxidizer mixture, e.g., by introducing an additional scalar to account for variations in the mixture composition. However, the use of reactor-based approaches for handling the TCI closure demonstrated to be more appropriate in the context of MILD combustion [7–9]. Nonetheless, a large number of stiff ordinary differential equations (ODEs) needs to be solved in each computational cell, making finite-rate models expensive, especially for realistic-sized combustion systems. In fact, the amount of operations required to integrate the chemical ODEs is proportional to the size and stiffness of the chosen oxidation mechanism [10].

Fundamental studies [11] demonstrated that the truly active chemical dynamics, which is a function of the local thermo-chemical state, involves a limited number of chemical pathways. Hence, a detailed kinetic mechanism, which is usually comprehensive, i.e., it is built to be applicable over the whole thermo-chemical space in many operating conditions, is locally oversized. This consideration opens many opportunities to tackle the computational cost burden of detailed kinetics. Several approaches have been proposed in the last decades, including methods for offline (pre-process) mechanism reduction [12,13] to generate beforehand a skeletal mechanism tailored to the specific operating condition under investigation, and on-the-fly reduction [14,15], to adaptively generate a locally reduced model at run-time. Nevertheless, if on the one hand a single skeletal mechanism produced beforehand still remains locally oversized, on the other hand the computational overhead needed for the on-the-fly reduction can be significant, especially for large mechanisms, since the reduction algorithms involve complex analyses of the non-linear chemical source terms. The combination of chemistry reduction and tabulation was therefore pursued in recent works [10,16–18], offering outstanding speed-ups. Li et al. [19] obtained a speed-up factor of more than 10 using a tabulated dynamic adaptive chemistry (TDAC) method [20], which couples in-situ adaptive tabulation (ISAT) [21] and dynamic adaptive chemistry (DAC) [14].

Another effective approach, called Sample-Partitioning Adaptive Reduced Chemistry (SPARC), was recently proposed to couple adaptive chemistry and machine learning [22,23]. This method consists in building a library of reduced mechanisms in a pre-processing step, i.e., before carrying out the CFD simulation, using a training dataset. This training dataset has to be designed, e.g., using ideal reactors or lower-fidelity simulations, in such a way to adequately cover the composition space (species mass fractions and temperature) expected to be visited during the CFD simulation, but without its computational overhead. The training dataset is thereupon pre-partitioned in a prescribed number of clusters, ensuring that each individual cluster is sufficiently homogeneous from the kinetic point of view, i.e., that a minimal reduced mechanism is able to capture the chemical dynamics emanating from all the samples belonging to the cluster [22]. At simulation run-time, before performing the chemical integration step, each cell of the CFD mesh is classified using an unsupervised-learning algorithm in order to identify the cluster to which the cell belongs, and to assign the corresponding reduced mechanism [22]. This workflow brings the advantages of the dynamic reduction techniques, i.e., the mechanism size is tailored to the local conditions, and simultaneously saves the cost associated with the on-the-fly adaptive reduction, which can be prohibitive when large kinetic mechanisms are used. A significant reduction in computational cost was achieved using SPARC, with simulations showing speed-up factors for the chemical step of ≈4 for laminar flames with a large detailed mechanism [22].

The performance of SPARC critically depends on the clusters homogeneity[1] and the mechanism reduction quality, i.e., the compromise between accuracy and size reduction. In this work, we focus on the latter. In the original SPARC workflow [22], the mechanism reduction is achieved with DRGEP [12], selecting a uniform reduction tolerance across the clusters, which controls the number of retained species and accuracy, and supplying a list of user-specified target species, namely fuel and oxidizer, around which the mechanism accuracy is tailored.

In this work, we show that this choice is sub-optimal, that is, a higher overall reduction is achievable with a proper selection of cluster-wise tolerances and target species. To this end, we employ the computational singular perturbation (CSP) reduction algorithm [13], which provides means for automatically identifying an adaptive set of target species. In addition, we propose a novel strategy for assessing a-priori the performance of the reduced mechanisms, with the goal of estimating a measure of accuracy of the upcoming CFD simulation. This assessment, which exploits the ideal reactors employed for the training dataset generation, is used to identify the optimal cluster-wise reduction trade-off, avoiding an expensive trial-and-error approach with CFD simulations.

We detail this enhanced SPARC workflow (called eSPARC in the present work) which accounts for automatic targets definition and error estimation, and we show its impact on the RANS simulation of the Adelaide Jet in Hot Coflow (AJHC) burner [24], both in terms of speed-up and accuracy improvements.

2. Methodology

The Sample-Partitioning Adaptive Reduced Chemistry (SPARC) methodology [22,23] consists of three steps in the pre-processing stage, followed by the multi-dimensional CFD simulation. Here, we present an enhanced eSPARC workflow, which includes automatic targets definition and a novel error estimation step before the CFD simulation:

1. Dataset generation: solutions of representative archetypal flames, e.g., 1D laminar counterflow flames in diluted conditions for the AJHC case, with detailed chemistry;
2. Clustering of the composition space based on Vector Quantization Principal Component Analysis (VQPCA) [25];
3. For each cluster, automatic target species identification and generation of a family of skeletal mechanisms of decreasing size, using CSP [13];
4. Error estimation on archetypal flames, using the cluster-wise reduced kinetic mechanisms, and selection of the optimal mechanism for each cluster;
5. CFD simulation with on-the-fly classification and mechanisms retrieval.

2.1. Dataset generation

To obtain a meaningful and accurate clustering of the flame regions, the training dataset must properly cover the composition space which

[1] In fact, a non-homogeneous cluster departs from the condition of localness by encompassing a multitude of chemical dynamics which require a larger kinetic scheme to be described.

is expected to be visited during the simulation [22,26]. Ideal reactors, spanning different operating conditions, are fair candidates because of their low computational cost. In this way, access to the turbulent combustion simulation is not required. Detailed kinetics has to be employed in this step. For the application under scrutiny, we employ unsteady 1D flamelets [27] to generate the dataset. To account for the three-stream nature of the JHC burner, the dataset is generated by means of unsteady laminar flamelets [27] in diluted conditions, accounting for pure fuel at the fuel inlet and a mixture of coflow/air at the oxidizer inlet. At the oxidizer side, the degree of mixing between coflow and air is varied from pure co-flow to a level of dilution with air for which flamelets have no ignition. Moreover, in light of the self-ignition propensity of the MILD combustion regime [1], we selected the unsteady model in order to capture flame features unavailable in steady conditions. The effect of the training dataset structure on the clustering and classification quality has already been explored in other works [22,26].

2.2. Clustering with VQPCA

PCA is a statistical technique that allows a large number of interdependent variables to be reduced to a smaller number of uncorrelated variables, while preserving as much of the original data variance as possible [28,29].

In case of strongly non-linear systems, such as reactive flows, a large number of PCs may be needed to describe the system correctly. To overcome this limitation, a locally linear approach called Local Principal Component Analysis (LPCA) [25] may be employed, where, after partitioning the data-space into different clusters, PCA is performed within each cluster to find a reduced representation. The local PCs better represent features of a specific cluster by capturing the local variance instead of the global one [30].

Data clustering, prior to applying local PCA, can be performed with any unsupervised-learning algorithm of choice. We adopt a clustering technique abased on the Vector Quantization PCA (VQPCA) algorithm, which assigns each observation of the dataset to the cluster for which its PCA reconstruction error is minimal [29].

The parameters needed to perform VQPCA are the number of considered PCs (n_{PCs}) and the number of clusters (k). To determine the optimal pair of parameters, the Davies–Bouldin (DB) index [31] is computed. The DB index is the ratio between the scatter within each cluster and the distance between clusters and it measures how good the partition is performed. In fact, even though the reconstruction error $\epsilon_{GRSE,k}$ is monotonically decreasing with respect to n_{PCs}, the dimensionality adopted for the reconstruction has to be chosen in order to reproduce only the important features of the data, so that the samples can be effectively divided in groups. The optimum pair of k and n_{PCs} is chosen as the one which leads to the minimum of DB index [31].

2.3. Mechanism reduction

We perform chemistry reduction using `pyCSP` [32]. The CSP method provides theoretical and numerical tools, such as the fast/slow importance indices, for analyzing the spectral behavior of stiff dynamical systems [33,34], therefore offering well-founded means for the removal of unnecessary chemical reactions without compromising the intrinsic structure of the dynamics.

The user is only required to supply a representative database of chemical states and the original detailed mechanism. By using a series of sequentially applied thresholds on the importance indexes, a range of skeletal mechanisms of varying fidelity and size can be produced.

More specifically, the importance indexes $(I)^i_{j_{\mathrm{slow/fast}}}$ allow us to identify the relative contribution that a given reaction j has in the production/destruction of a species i in either the fast or slow subspace.[2] The indexes normalization is necessary to compare reactions contributing to distinct components of the source term, e.g., temperature, major species, and minor species, which might differ substantially in magnitude.

With the selection of a threshold $0 < \tau < 1$, a set of reactions whose $(I)^i_{j_{\mathrm{fast/slow}}} > \tau$ is retained in the skeletal mechanism as the ones mostly contributing the dynamics of the ith species of interest, i.e. a set of target species, whose dynamics is desired to be accurate in the skeletal representation. Once a set of important reactions is identified based on the importance indices to the target species, a new set of important species is defined as the one containing the species involved in the important reactions. Then, the importance indices are computed for this new set of species. The algorithm loops until this set of important species remains unchanged.

Since the CSP decomposition and the reaction elimination procedure are local – the algorithm works on a single thermo-chemical state at a time – a set of important species/reactions is identified in each state belonging to a database, and a global set of retained species/reactions is built as the union of the local sets of important species/reactions. A progressive increase of the threshold τ allows us to generate a family of reduced mechanisms of decreasing sizes, each associated with a certain – presumably decreasing – degree of accuracy in the replication of the target species chemical dynamics. Hence, the user controls the threshold to trade accuracy for mechanism size reduction, i.e., lower computational cost.

On this regard, it is worth noticing that the indexes normalization, while being crucial on a local (per-state) basis, carries a side-effect when analyzing a database of states. In fact, relative contributions are not comparable across states. More specifically, the importance index of one reaction in one state is not comparable to its analogue in another state, being the importance a relative measure. The consequence is that the effect of a uniform τ value across all the states of a database might be sub-optimal: for a given accuracy requirement on major observables, states carrying complex dynamical features would need a lower τ value (higher number of retained species) than states exhibiting simpler, or less relevant (to the targets) dynamics [13]. However, the specification of a non-uniform τ over a database of states is challenging, since an a-priori distinction between states carrying essential or less-essential dynamical features is not straightforward.[3]

On the other hand, a clustered database lends itself to the specification of cluster-wise τ values, alleviating the aforementioned issue. In fact, this approach introduces a higher degree of flexibility in the mechanism library construction, allowing (very) different degrees of reduction across the clusters. Moreover, since the clusters are built following a feature-extraction method, we expect a high degree of homogeneity in the dynamics emanating from the states belonging to each cluster, making the use of a uniform τ inside the cluster less penalizing.

In addition to a cluster-wise threshold, we introduce an automated target species selection. The target species set acts as a seed for the reduced mechanism, which will be tailored to meet accuracy requirements on these specific species. In traditional mechanism simplification works (e.g. [12,35–37]), this set of target species was prescribed to include the major species, such as fuel(s), oxidizer(s) and the major products, plus other species that are of interest for specific observables, e.g., OH and/or HCO for ignition. This choice is generally left to the user, despite being crucial: a blind choice may lead to

[2] The separate examination of the fast and slow subspace is pursued because the reaction contributions to the dynamics of one species may have differences of several orders of magnitude in one subspace compared to the other, typically large and balanced in the fast subspace, and smaller but unbalanced in the slow subspace.

[3] Even a definition of essential is not straightforward.

either unnecessarily oversized or inaccurate reduced mechanisms. It was demonstrated in [13] that an optimum, local set of target species exists that leads the algorithm to include a minimal number of species that ensures the replication of global flame observables, such as ignition delay time. This is even more true in a cluster of states identified unsupervisedly. In fact, the dynamics of a single cluster of points may entirely not involve fuel, oxidizer and/or products, or it may involve only a subset of them. In the context of SPARC, this feature is anticipated to be useful in minimizing the cluster-wiser number of retained species, if coupled with a proper cluster partitioning. To this end, we employ the Tangential Stretching Rate (TSR) criterion included in pyCSP which automatically identifies the locally most relevant system's state variables. More specifically, the TSR selects the most energetic kinetic eigen-modes. The most relevant species are identified as the ones participating in the reactions mostly contributing the TSR-selected modes.[4] Therefore, the species involved in the driving dynamics are automatically selected as the local set of target species [13], without the need for any user intervention.

2.4. Error estimation of reduced kinetics

For each cluster of data, a family of reduced kinetics is generated by varying the CSP threshold τ. Nevertheless, only one mechanism per cluster must be eventually selected to build the library and proceed with the CFD simulation. A possible rationale for the selection process would be to prescribe accuracy and/or size requirements. However, the heuristic nature of reduction algorithms precludes to drive the process with measures of quality of the reduced mechanisms: an accuracy assessment can only be performed a-posteriori, by using the reduced mechanisms in prototype flame problems, and comparing against the detailed solution of the same problems. Therefore, the mechanism that meets size and accuracy requirements is necessarily picked among a pool of candidates undergoing empirical benchmarks.

In the SPARC literature [22], the accuracy assessment of the reduced mechanisms library was only performed a-posteriori at the CFD level, that was, by comparing the results of the adaptive CFD simulation against a reference detailed one. However, this is impractical in an operative perspective, since many potential libraries of reduced mechanisms can be created, with $n_{clusters} \times n_\tau$ combinations, and/or a reference simulation may not be feasible/available.

In the present work, we add a novel step in the SPARC methodology for estimating the accuracy of each cluster-wise reduced mechanism in the pre-process phase, i.e., before the CFD simulation. Such evaluation enables an informed choice of the optimal cluster-wise reduction thresholds to be selected. In other words, given an accuracy requirement, this strategy delivers a minimally sized reduced mechanism per each cluster.

To this end, it is necessary to identify an archetypal flame problem that is a low-dimensional/low-cost representative of the physics underlying the target CFD problem, and which covers the composition space of the clustered dataset. A natural choice is to employ the same model problem used to generate the dataset, i.e., the unsteady 1D flamelets described in Section 2.1, with a reduced mechanism in place of the detailed one. The comparison between detailed and reduced solutions of the model problem allows us to estimate error measures and therefore guide the reduced mechanisms selection.

However, this assessment poses a number of difficulties arising from the fact that, in general, the composition space of one cluster does not necessarily coincide with the composition space populated by the solution of the model problem, e.g., it may cover a portion of it, with the associated reduced mechanism being naturally inaccurate outside. To overcome this issue, and to evaluate the accuracy of each mechanism on its actual cluster conditions, we couple the VQPCA classifier with the model problem solver (the unsteady 1D flamelet solver in the present case). This way, for each cluster k we run a validation model problem in which the reduced mechanism under scrutiny is selected to supply kinetics in place of the detailed one only when VQPCA classifies a state as belonging to cluster k. The VQPCA-flamelet solver was implemented in Python using CVODE [39] to integrate the unsteady 1D flamelet equations.[5] The VQPCA implementation was taken from PCAfold [40].

In summary, for each cluster k, we follow the algorithmic procedure described here below and sketched in Fig. 1. In cluster k, we generate a family of reduced mechanisms $\mathcal{M}^k(\tau)$ for a given number n_τ of CSP thresholds τ. We run the model problem solver, in the same conditions employed for generating the full initial database. The different states are classified by VQPCA. If one point belongs to cluster k, only the species retained in $\mathcal{M}^k(\tau)$ are advanced in time during the chemical step, otherwise the detailed mechanism is used. Hence, for the considered cluster k, we obtain a solution of the model problem for each mechanism $\mathcal{M}^k(\tau)$, i.e, for each threshold τ.

The error assessment is performed by associating to each target species θ a relative error $\mathcal{E}^k_\theta(\tau)$ computed between the reduced (Y^τ_θ) and detailed (Y^{dtl}_θ) space/time profiles for species θ when mechanism $\mathcal{M}^k(\tau)$ is prescribed. Such $\mathcal{E}^k_\theta(\tau)$ is computed as the maximum relative error among the grid points p which are classified as belonging to cluster k:

$$\mathcal{E}^k_\theta(\tau) = \max_{p \in k} \left(\frac{\left| Y^\tau_{\theta,p} - Y^{dtl}_{\theta,p} \right|}{Y^{dtl}_{\theta,p}} \right) \qquad \theta = 1, \dots, N^k_{targets} \qquad k = 1, \dots, N_{clusters}. \tag{1}$$

In this calculation species with mass fraction lower than 10^{-8} are considered traces and not taken into account for the evaluation . The overall measure of error for mechanism $\mathcal{M}^k(\tau)$, indicated as $\mathcal{E}^k(\tau)$, is selected as the worst error figure over the targets:

$$\mathcal{E}^k(\tau) = \max_{\theta \in N^k_{targets}} \mathcal{E}^k_\theta(\tau). \tag{2}$$

The optimal mechanism in each cluster is therefore chosen as the smallest mechanism satisfying an error requirement ϵ_r over all the target species, so that the final library is defined as:

$$\{\mathcal{M}^k(\tau_{opt} : \mathcal{E}^k(\tau_{opt}) < \epsilon_r)\}_{k=1,\dots,N_{clusters}}. \tag{3}$$

We note that the validation loop over the threshold is able to identify the most compact mechanism over a family of mechanisms and for a given accuracy requirement regardless of the chosen chemistry reduction algorithm. In the present work, we employ the CSP method because it offers a straightforward local automatic target species identification, which we consider as important as the threshold selection criterion, bringing a non-negligible advantage over other methods.

Lastly, we remark that this a-priori error analysis does not fully replace the a-posteriori CFD error evaluation. Indeed, the procedure's reliability is as much higher as the archetypal problem (e.g., the 1D flamelet) resembles the physics captured by the CFD. Therefore, a user should be cautious about the a-priori error evaluation if the complexity of the CFD simulation transcends the limitations of the chosen simplified model problem. Nonetheless, with respect to the original SPARC methodology, the analysis offers a substantial improvement, delivering upfront cluster-wise information on the validity of the reduced mechanisms, and therefore leading to more accurate and less computationally expensive CFD simulation without the need to run several benchmark multidimensional RANS. Such improvements in terms of speed up and accuracy are detailed in Section 3.

[4] The interested reader is referred to [13,33,38].

[5] Once a state is classified as belonging to a cluster, the corresponding reduced kinetics is supplied by zeroing out the chemical source terms of the removed species.

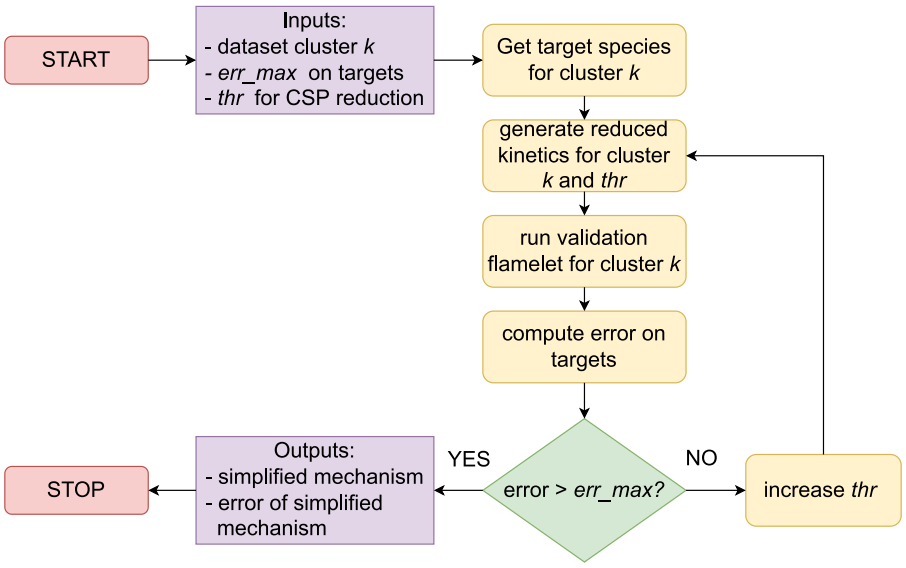


Fig. 1. Reduction scheme: The algorithm loops to find the optimal CSP threshold.

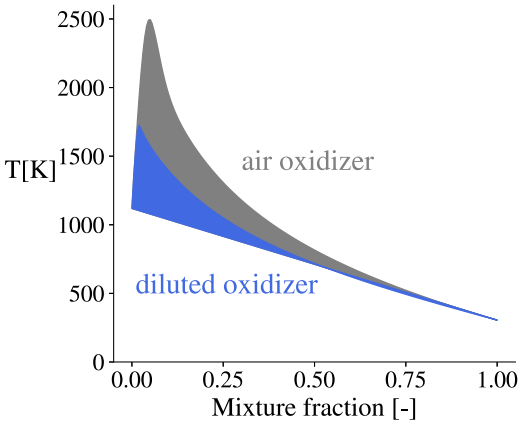


Fig. 2. Temperatures experienced by two unsteady flamelets in the mixture fraction space: diluted condition with 85% coflow in the oxidizer (blue) and pure air (21% Oxygen) in the oxidizer (gray). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

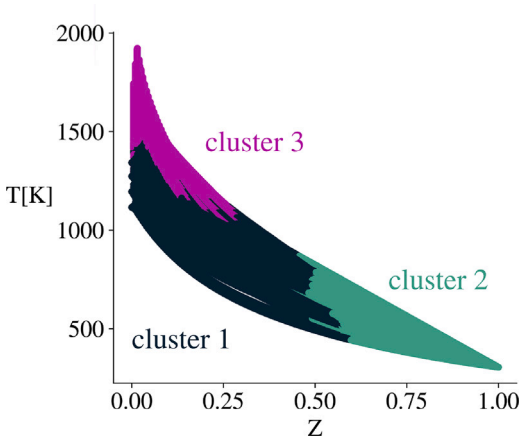


Fig. 3. Partitioning solution of VQPCA algorithm in the mixture fraction space for the considered dataset. VQPCA is able to identify groups of points close in the thermochemical space.

2.5. CFD error measures

We assess the CFD accuracy by measuring the difference on QoIs, i.e., species and temperature, with respect to the detailed simulation. For each QoIs, the CFD error E_Θ is computed as a relative l^2 error norm over all the grid points N_{points} of the computational mesh[6]:

$$E_\Theta = \frac{\left(\sum_{p=1}^{N_{points}} |Y_{\Theta,p} - Y_{\Theta,p}^{dtl}|^2\right)^{\frac{1}{2}}}{\left(\sum_p (Y_{\Theta,p}^{dtl})^2\right)^{\frac{1}{2}}} \qquad \Theta = 1, \dots, N_{QoI}. \tag{4}$$

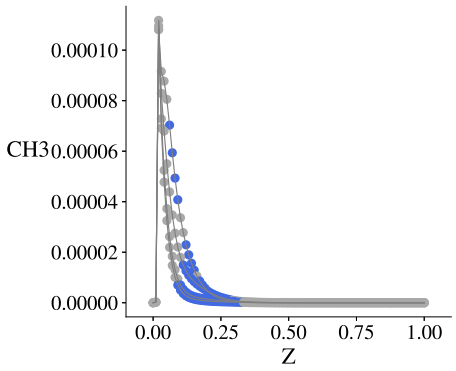
Therefore, the overall error for one CFD simulation E_{CFD} is given by the arithmetic mean of each error on QoIs:

$$E_{CFD} = \frac{\sum^{N_{QoI}} E_\Theta}{N_{QoI}} \qquad \Theta = 1, \dots, N_{QoI} \tag{5}$$

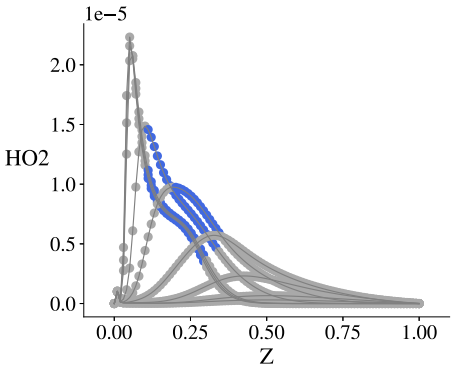
[6] Such kind of error metrics may be affected by the mesh topology and grading, since the grid points may not be evenly distributed in space. To avoid ambiguities and have a mesh-independent measure of the error, we compute CFD error metrics on a uniform mesh.

2.6. Validation case and numerical setup

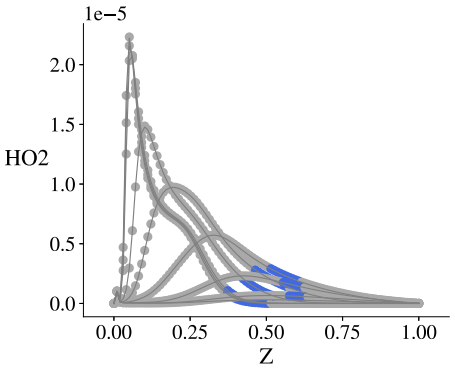
The eSPARC workflow is tested on the $Re = 10^4$, 3% O_2 coflow case of the Adelaide Jet in Hot Coflow (AJHC) burner investigated experimentally by Dally et al. [24]. The AJHC burner is comprised of three streams: an insulated and cooled central jet issuing a equi-molar mixture of CH_4 and H_2, a coaxial coflow of hot combustion products and an outer standard air flow surrounding the coflow. The simulation domain extends 175 mm axially downstream the burner outlet and 80 mm towards the radial direction. The entire domain is discretized with a structured axisymmetric 2D mesh. The unsteady Reynolds Averaged Navier–Stokes (uRANS) simulations were carried out using the `FiReSMOKE` code [41], a CFD solver based on `OpenFOAM`. The turbulence model employed for the uRANS simulations is the standard $k-\epsilon$ model, while the detailed kinetic mechanism is the 36-species `GRI3.0` without the NOx submechanism [42]. The combustion model employed to handle chemistry-turbulence interactions is the Partially-Stirred Reactor (PaSR) model [43]: the same model and code were previously validated by Li et al. [9], and therefore validation on the experimental data will not be reported in this work.



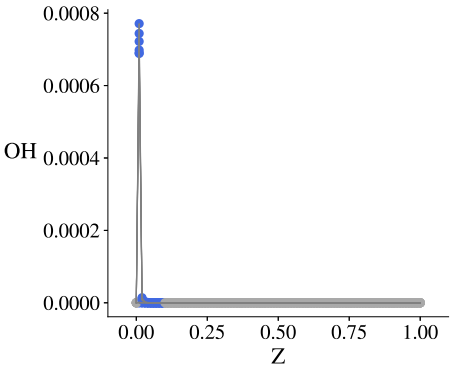
(a) Evaluation points on CH3 profile for cluster 0.



(b) Evaluation points on HO2 profile for cluster 1.



(c) Evaluation points on HO2 profile for cluster 3.



(d) Evaluation points on OH profile for cluster 5.

Fig. 4. Select target species profiles at different time instants in 1D flamelets. Blue symbols correspond to states associated to cluster 0, 1, 3 and 5 respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Training dataset

As described in Section 2.1, we generate the training dataset by means of unsteady laminar flamelets in diluted conditions. At the oxidizer inlet, we use three different levels of mixing between co-flow and air: 99% of co-flow, which is close to pure co-flow conditions; 70% of co-flow, which is the lower limit below which the flamelet is too diluted and there is no ignition, and 85% of co-flow, taken as an intermediate level between the two bounds. To give a pictorial idea of the effect of the co-flow dilution, Fig. 2 shows a comparison between temperatures covered by one diluted flamelet with 85% coflow as oxidizer and one with pure air as oxidizer. The scalar dissipation rate of the diluted flamelets (χ) has been varied from 10^{-3} to 10^{3} s^{-1}, i.e., from no-ignition to blow-off conditions. The final dataset consists of approximately 17 000 state points.

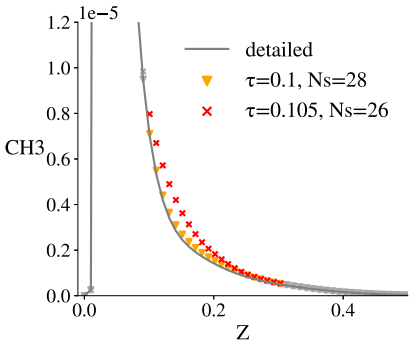
3.2. Clustering

The partition of the composition space is performed with the VQPCA classification methodology described in Section 2.2. Fig. 3 shows a partitioning solution of the temperature field using this classifier when 3 clusters are considered. We observe that the three groups of points correspond to different thermochemical regions where we expect different dominating phenomena, such as steady-state heat release (cluster 3), pure mixing (cluster 2) and transient ignition (cluster 1). However, as explained in Section 2.2, the optimum pair of number of clusters k and number of PCS n_{PCs} is chosen as the one that leads to the minimum DB index [31]. In this work, this criterion leads to a partition into $k = 8$ clusters, with $n_{PCs} = 1$, with a corresponding DB index of 1.52. These different 8 regions, due to the high dataset dimensionality, are non-trivial to visualize in a 2D plot as the one in Fig. 3, therefore we do not show them to avoid confusion.

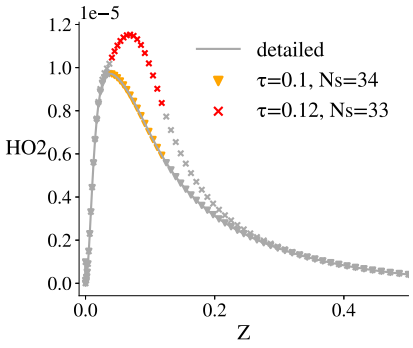
3.3. Mechanism reduction and error estimation

For each of the aforementioned clusters of data, a family of reduced mechanisms of decreasing size is generated by varying the CSP threshold τ. We thereupon apply the novel error estimation approach, summarized in Fig. 1, which selects the smallest reduced mechanism in each cluster, given an accuracy requirement. The target species for each thermochemical state are automatically identified by the TSR metric included in pyCSP. The species whose profiles are used for mechanism validation are those which appear in at least 10% of the observations of a selected cluster. In this way, outlier species due to a non perfectly homogeneous classification are not considered for the evaluation. The target species are collected and shown in Table 1 for each cluster. As expected, each cluster is characterized by a different chemical dynamics, leading to the identification of different sets of target species. Once the targets are identified, the error estimation loop is performed for each cluster.

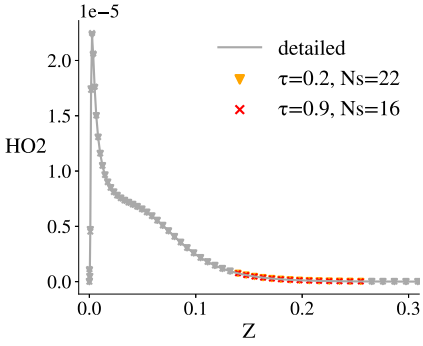
To clarify how the error estimation algorithm works, we firstly show the how the VQPCA clustering operates on the unsteady flamelets.



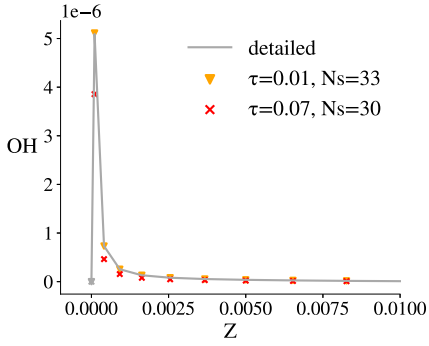
(a) CH3 profile for different size of reduced mechanisms for cluster 0 at $t = 0.1s$.



(b) HO2 profile for different size of reduced mechanisms for cluster 1 at $t = 0.03s$.



(c) HO2 profile for different size of reduced mechanisms for cluster 3 at $t = 0.001s$.



(d) OH profile for different size of reduced mechanisms for cluster 5 at $t = 0.1s$.

Fig. 5. Select target species profiles obtained with different reduced mechanisms in clusters 0, 1, 3, and 5 respectively (colored symbols). Gray symbols corresponds to states integrated in time using detailed kinetics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Targets obtained through TSR-CSP algorithm for each cluster.

Cluster #	Targets for reduction
0	H2 - H - CH4 - CH3 - C2H3
1	H2 - H - HO2 - CH3 - CH4
2	H2 - H - O2 - HO2 - CH3 - C2H6
3	O2 - HO2 - H2O2 - CH3 - CH4
4	O2 - HO2 - H2O2 - CH3 - CH4
5	H2 - H - OH - CH3 - CH4 - CO - CO2
6	H2 - H - CH3 - CH4 - C2H5 - C2H6
7	H2 - H - O2 - HO2 - CH3 - CH4

Fig. 4 shows the CH_3, HO_2, and OH profiles at select times instants: the blue symbols identify the thermochemical states associated to clusters 0, 1, 3 and 5, respectively, by the classifier, that is, where the reduced model is actually employed in place of the detailed mechanism. The gray symbols represent the thermochemical states associated to the other clusters, which are therefore solved with the detailed mechanism. Such flamelets are advanced in time, with local adaptive chemistry, for each one of the differently-sized reduced mechanisms, and the different solutions are reported in Fig. 5, where each color is associated to a given reduced mechanism (we only show 2 representative reduced mechanisms for clarity).

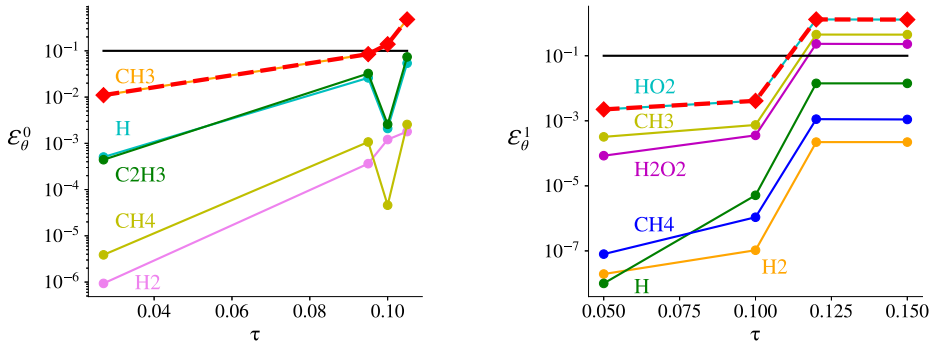
In these plots, we observe both the effect of reducing the chemical mechanism and the effect of using it in different local conditions of the flamelets. In particular, Fig. 5(b) shows that for cluster 1, a slightly reduced mechanism (32 species) already exhibits significant discrepancies in the HO_2 predictions with respect to the detailed solution. On the other hand, according to Fig. 5(c), HO_2 has a smaller deviation from the detailed solution as far as cluster 3 is concerned, even using strongly reduced kinetics (16 species), because cluster 3 corresponds to a region where fewer chemical pathways are active.

Errors on target species are computed between the reduced and detailed profiles according to Eq. (1), for each clusters and are shown in Fig. 6 as functions of the CSP threshold. For each threshold, the maximum error among all targets is considered for the error assessment of that cluster (red dashed line).

Lastly, Fig. 7 collects such estimated errors for each cluster $\mathcal{E}^k(\tau)$ as a function of the CSP threshold. It is evident how the error behavior largely varies between clusters; therefore, this plot highlights the benefit of an a-priori error assessment.

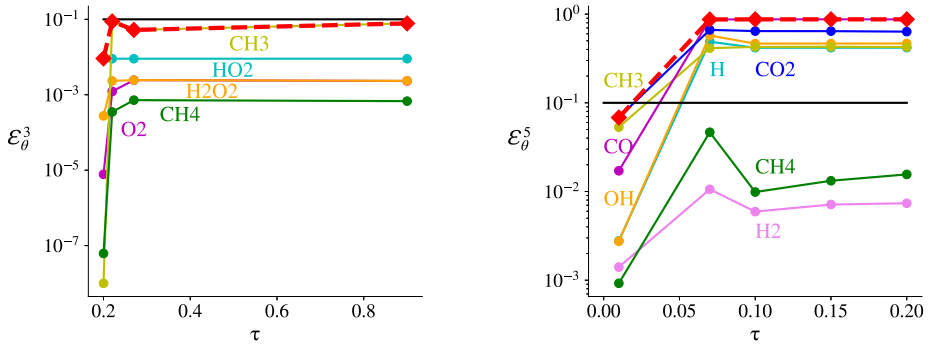
In fact, the effects of reduction vary significantly from cluster to cluster, and so does the error. This way, by fixing an accuracy requirement, e.g. 10^{-1}, we are able to select an optimal threshold τ_{opt} for each cluster, that is, the one which leads to the smallest mechanism satisfying the given accuracy requirement. Table 2a reports the number of retained species and the corresponding error estimate for the 8 clusters with the optimal threshold.

To further highlight the benefit of this methodology, we additionally build reduced mechanism libraries at fixed CSP threshold across the clusters, that is, neglecting any a-priori error assessments.



(a) Error estimation on targets for cluster 0.

(b) Error estimation on targets for cluster 1.



(c) Error estimation on targets for cluster 3.

(d) Error estimation on targets for cluster 5.

Fig. 6. Errors on target species computed between the reduced and detailed profiles according to Eq. (1), for clusters 0, 1, 3, and 5.

We will thereupon compare CFD speed-up and accuracy using fixed-threshold and optimal-thresholds libraries. We select a range of thresholds $\tau_{fixed} = [0.05, 0.1, 0.15, 0.2, 0.25, 0.3]$. We expect high overall accuracy, but also larger mechanisms and higher CPU time when using the lower thresholds, and smaller mechanisms, lower CPU time, and lower accuracy when using the higher ones.

The crucial difference between fixed-threshold and optimal-thresholds libraries, as it will be clear in the following, is the accuracy unbalance among the clusters, which is inevitable in the fixed-threshold case. Note that, even a single inaccurate cluster may jeopardize the entire CFD simulation. By virtue of the a-priori campaign, we are also capable of giving error estimates for the fixed-threshold libraries: when the a-priori error exceeds the maximum admissible value, such mechanisms are expected to increase inaccuracies in the CFD simulation. In Table 2b it may be noticed that even in the less reduced – and presumably more accurate – $\tau_{fixed} = 0.05$ configuration, the estimated error for cluster 5 exceeds the error requirement of 10^{-1}, therefore the CFD simulation performed with this library is expected to suffer as well from accuracy issues.

4. CFD results

In this section, we perform and discuss a comparison between the six mechanism libraries obtained with fixed CSP threshold and the optimal library obtained with the novel strategy, in terms of CFD speed-up and accuracy with respect to a detailed RANS simulation of the AJHC burner. The performance of the VQPCA classifier in multidimensional simulations is firstly tested. Fig. 8 shows the CFD predictions for the considered case, obtained with the optimal kinetics library in Table 2a. More specifically, Figs. 8(a), 8(b) and 8(c), show the temperature, OH

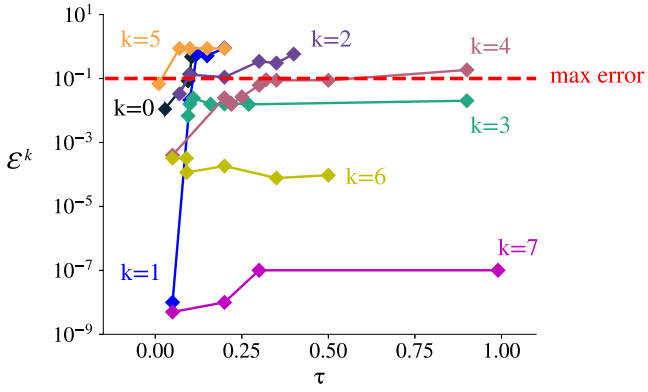
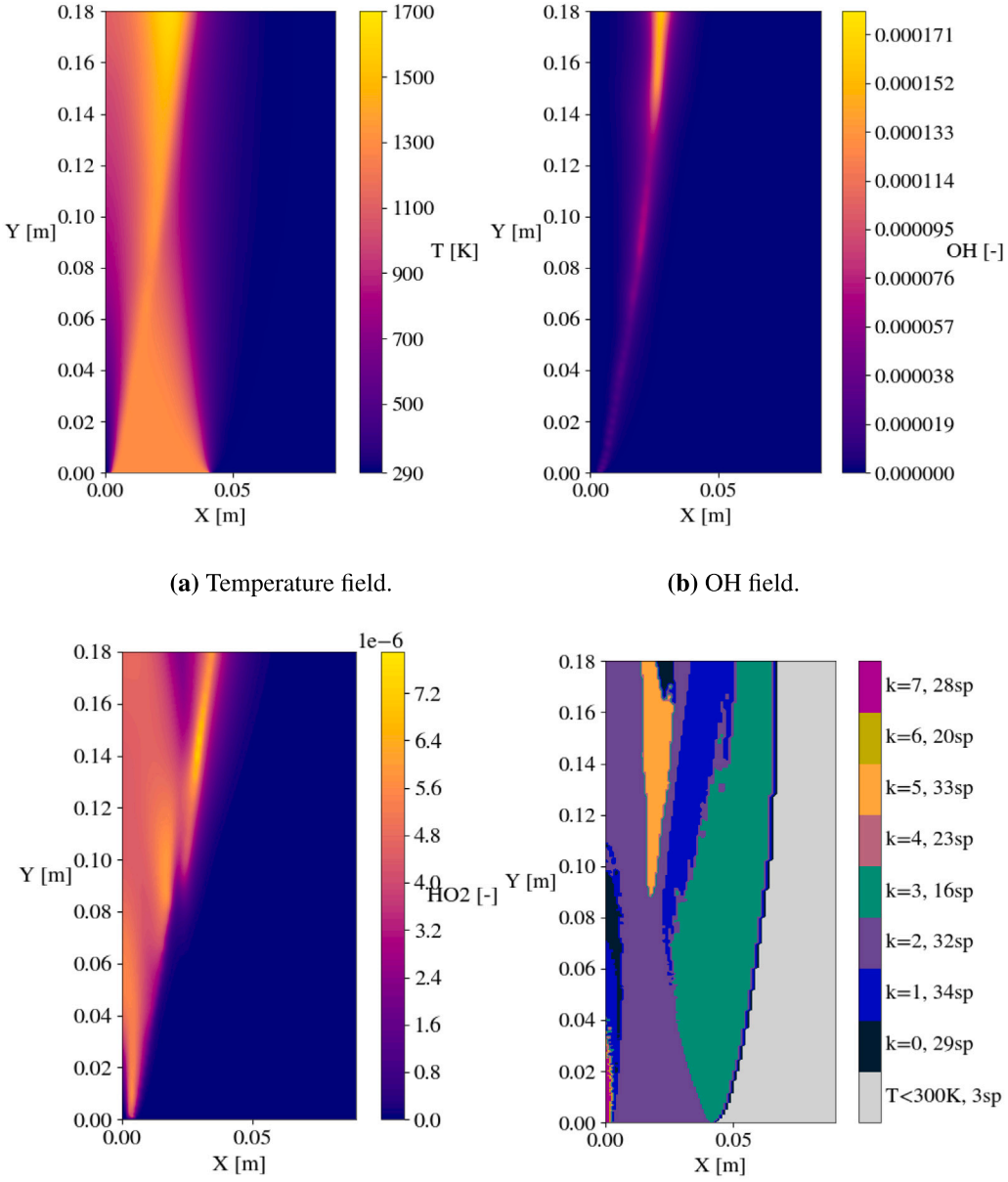


Fig. 7. Cluster-wise a-priori error estimation versus CSP threshold.

and HO_2 fields respectively, while Fig. 8(d) shows the cluster index for each computational cell.[7] This image helps in visualizing how the algorithm classifies the different regions of the flame as they belong to the prescribed groups: the progressive fuel consumption is captured by clusters 7, 0 and 1; the interactions between air and coflow and

[7] A 3-species cluster appears as well (colored in gray). This is not a proper cluster, but rather a region of the field where the chemical activity is absent, identified by a threshold on the temperature value of 290 K, and in which the chemical source term is not evaluated (in both the detailed and the SPARC simulations).



(a) Temperature field. **(b)** OH field.

(c) HO2 field. **(d)** Number of active species cluster-wise.

Fig. 8. CFD RANS result for AJHC case using eSPARC methodology with optimal CSP threshold cluster-wise. Fig. 8(d) shows how the VQPCA classifier maps the flame before every chemistry integration, local number of active species is also indicated for each cluster. Note that only 3 species are solved for cells with T lower than 300 K, which is a feature provided by `FiReSMOKE` solver.

Table 2
Threshold, active species and error within clusters for two considered configurations.

(a) Reduction τ_{opt}				**(b)** Reduction τ fixed			
k	τ_{opt}	N_s	err	k	τ	N_s	err
0	0.095	29	8×10^{-2}	0	0.05	33	2×10^{-2}
1	0.1	34	2×10^{-2}	1	0.05	35	8×10^{-9}
2	0.07	32	3×10^{-2}	2	0.05	33	3×10^{-2}
3	0.9	16	2×10^{-2}	3	0.05	32	6×10^{-3}
4	0.3	23	5×10^{-2}	4	0.05	35	4×10^{-4}
5	0.01	33	6×10^{-2}	5	0.05	32	8×10^{-1}
6	0.5	20	8×10^{-2}	6	0.05	30	3×10^{-4}
7	0.99	28	9×10^{-8}	7	0.05	34	5×10^{-9}

between fuel and co-flow are captured by cluster 3 and clusters 2 respectively; the main reaction zone, where OH is produced and heat is released, is contained in cluster 5.

A measure of accuracy for each RANS simulations is given by comparing the errors on species computed with respect to the detailed simulation, according to Eq (4). Fig. 9 shows the errors on species E_Θ, for the 6 fixed-threshold simulations, normalized by the error obtained with eSPARC ($E_{\Theta_{eSPARC}}$), as functions of the average number of species solved for in the CFD.

Being this normalized error almost always greater than 1 (except for O_2 for $\tau = 0.05$), we observe that the cluster-wise choice of the reduction threshold leads to an error on species which is the lowest in all the species for each configuration. Moreover, considering the overall CFD error E_{CFD} computed according to Eq. (5) and shown in Fig. 10(a), we notice that this error is lowest than those obtained in all the 6 fixed-threshold cases despite the average number of solved species (~17.5)

Table 3
CFD performance summary. Note that the number of species solved in the detailed simulation is less than the number of species in the detailed mechanism, since only 3 active species are assigned for cells with T lower than 300 K.

τ	Detailed	0.05	0.1	0.15	0.2	0.25	0.3	τ_{opt}
N species	28.6	26.1	22.9	22.2	20.3	18.6	17.9	20.5
E_{CFD}	0	3.5×10^{-2}	5.7×10^{-1}	6.3×10^{-2}	1.0×10^{-1}	6.3×10^{-1}	1.5×10^{0}	9.7×10^{-3}
Speed-up	1	1.08	1.29	1.38	1.56	1.57	1.69	1.97

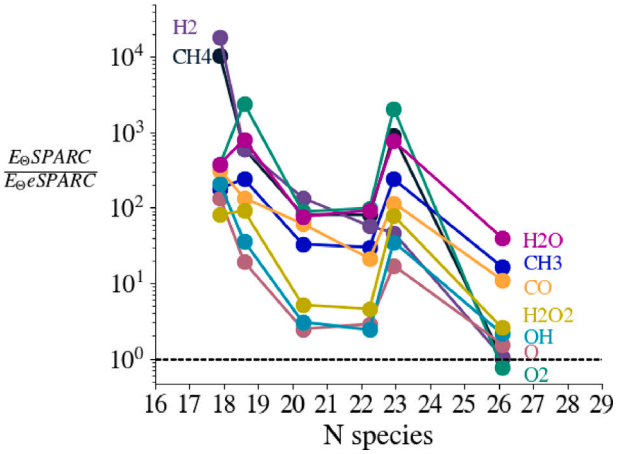
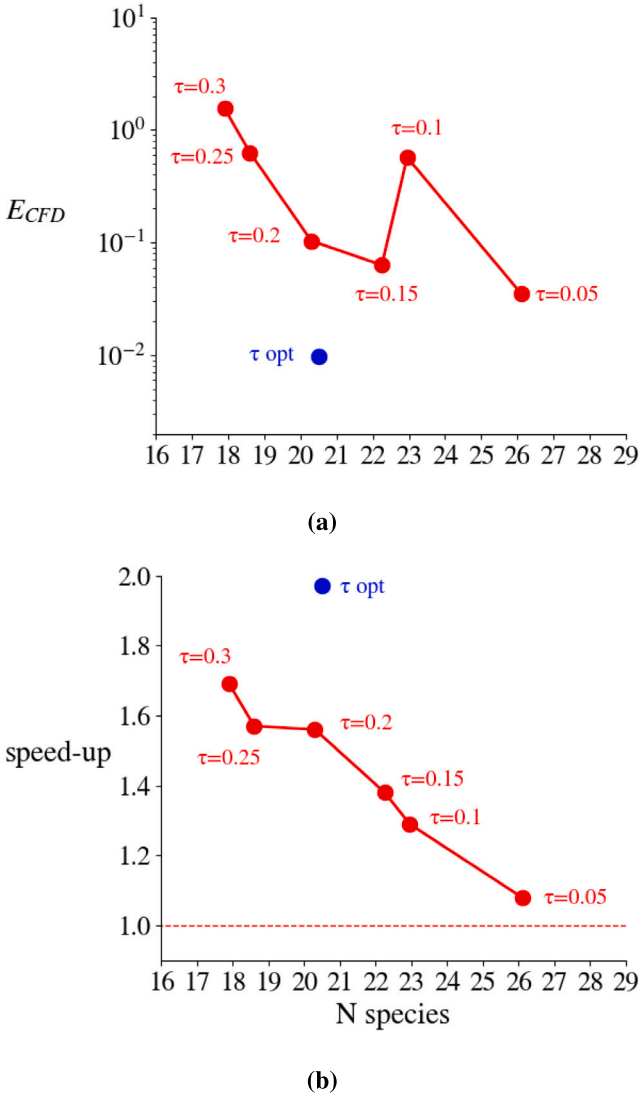


Fig. 9. A-posteriori CFD error of fixed-threshold SPARC ($E_{\Theta_{SPARC}}$), normalized with the one of the optimal-threshold eSPARC ($E_{\Theta_{eSPARC}}$), for different thresholds, and corresponding number of averaged solved species.



(a)

(b)

Fig. 10. A-posteriori CFD error measured according to Eq. (4) and computational speed up as a function of mean number of species solved for the seven considered configurations. The optimal threshold choice cluster-wise, offered by eSPARC (blue), brings to the largest CPU time saving and also the best accuracy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

being comparable to that of $\tau = 0.2$. In other words, to obtain such an accuracy with a fixed-threshold approach, more than 22.5 species would have been required, on average. This demonstrates that the error estimation step of the eSPARC workflow allowed us to generate an optimal kinetics library consisting of locally minimal mechanisms for the given accuracy requirement.

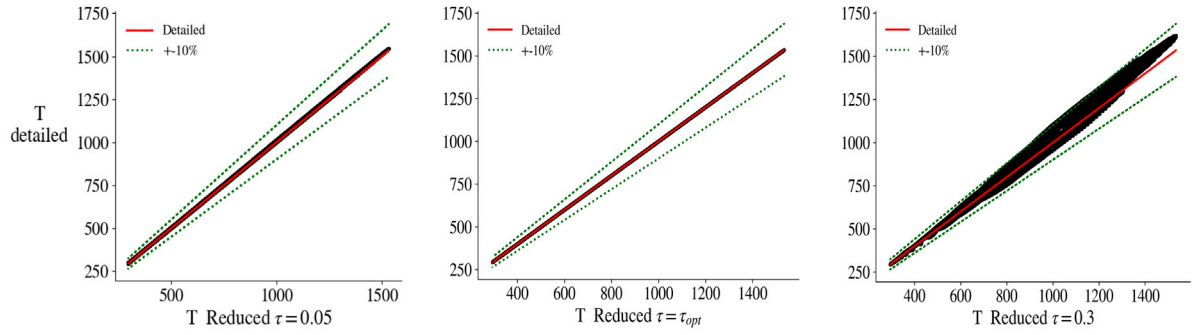
The CFD error E_{CFD} is also shown in Table 3 for the seven configurations, along with the corresponding average (over the cells) number of species, and computational speed-up (also shown in Fig. 10(b)) of the chemical step. The numbers confirm that the τ_{opt} configuration is more accurate than the $\tau = 0.05$ one, but with an average number of solved species which is closer to the $\tau = 0.2$ case. Moreover, from Fig. 10(b) we notice that, even if the mean number of species of case τ_{opt} (17.55) is close to the one of $\tau = 0.2$ (17.45), the correspondent chemistry speed-up is even larger then the one for $\tau = 0.3$ configuration. We conjecture that, even if the mean number of species is approximately the same, the species are better distributed between clusters and this reduces the time required for integrating chemistry.

Lastly, a visual inspection of the parity plots with respect to simulations with uniform threshold reveals an evident disparity in terms of accuracy. Parity plots of temperature and, CH_3, CO and OH concentrations, are shown in Figs. 11(c), 12(c), 12(f), 12(i), in Figs. 11(a), 12(a), 12(d), 12(g) and in Figs. 11(b), 12(b), 12(e), 12(h) for $\tau = 0.3$, $\tau = 0.05$ and $\tau = \tau_{opt}$ respectively. The improvement in accuracy given by the τ_{opt} library is visibly evident, while, according to Table 3, the simulation is the least computational expensive for solving chemistry.

5. Conclusion

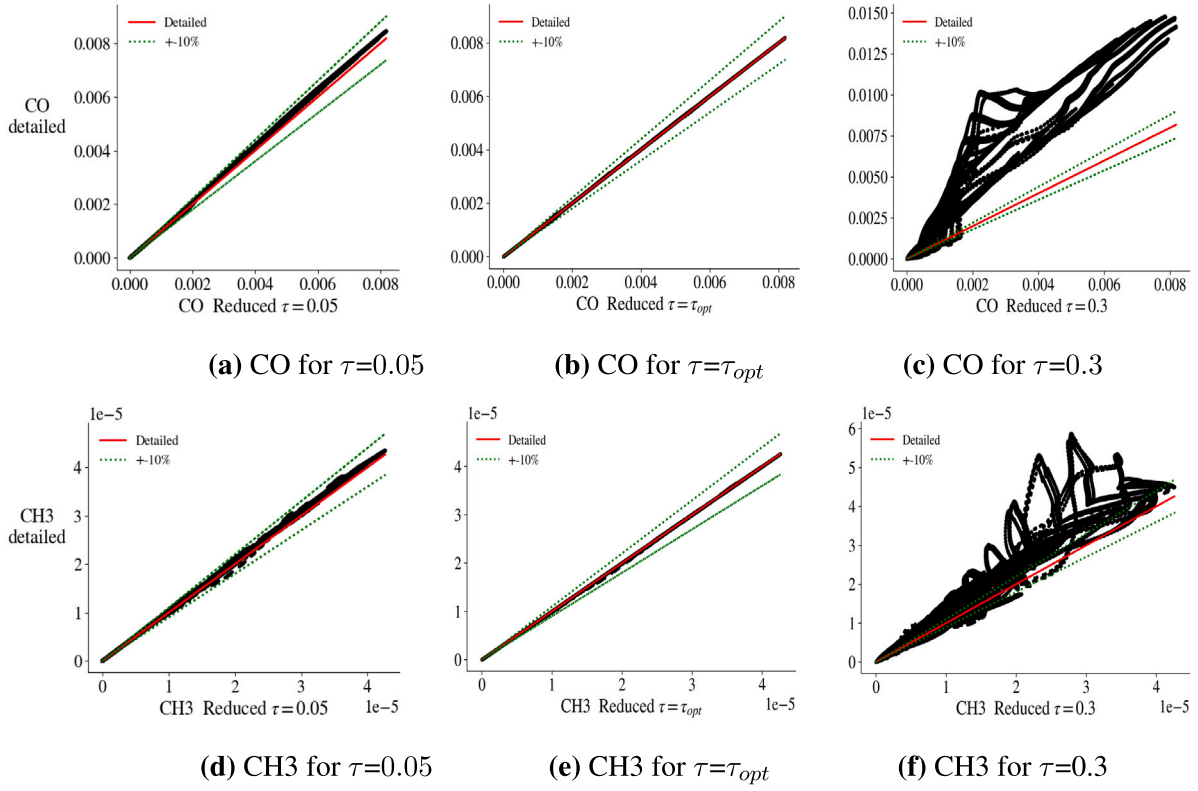
We proposed an enhanced Sample-Partitioning Adaptive Reduced Chemistry (eSPARC) workflow, in which the mechanism reduction step is enriched with automatic target species definition and a-priori error estimation.

In the original SPARC method, the reduction is performed using DRGEP, and the same target species, as well as the same reduction thresholds, are kept across all the clusters. In this work, we showed that this choice is sub-optimal, that is, a higher global reduction is achievable with a proper selection of cluster-wise thresholds and target species. To this aim, the reduction in eSPARC is performed with the CSP method, augmented with TSR analysis, which automatically identifies the locally most relevant species, to be employed as target species, instead of user-defined species based on a-priori knowledge. Moreover, a novel error estimation step is added in the pre-processing part, which implements a strategy for determining the accuracy of the reduced mechanisms on 1D flames, i.e., before performing the multidimensional CFD. This leads to select cluster-wise mechanisms that comply with



(a) T for τ=0.05 **(b)** T for $\tau=\tau_{opt}$ **(c)** T for τ=0.3

Fig. 11. SPARC vs. eSPARC: Parity plots for temperature.



(a) CO for τ=0.05 **(b)** CO for $\tau=\tau_{opt}$ **(c)** CO for τ=0.3

(d) CH3 for τ=0.05 **(e)** CH3 for $\tau=\tau_{opt}$ **(f)** CH3 for τ=0.3

(g) OH for τ=0.05 **(h)** OH for $\tau=\tau_{opt}$ **(i)** OH for τ=0.3

Fig. 12. SPARC vs. eSPARC: Parity plots for radical species.

a given error requirement, optimizing the mechanism accuracy-size compromise.

The methodology was successfully tested on RANS simulation of the Adelaide Jet-in-hot-coflow (AJHC) burner using the GRI3.0 mechanism, showing remarkable improvements with respect to the original SPARC workflow, in terms of both speed-up and accuracy. We compared CFD speed-up and accuracy of reductions performed at fixed CSP threshold. Results confirm that for the τ_{opt} case, while the average number of species solved is close to the $\tau = 0.2$ case and the speed up in chemistry is higher than the $\tau = 0.3$ configuration, at the same time it is more accurate than the $\tau = 0.05$ case, where high overall accuracy was obtained at the price of larger mechanisms.

The new methodology proved to be a powerful tool because no prior knowledge on the reduction step is needed. The obtained speed-up was ∼2 with respect to the detailed simulation, but better performances are expected when using larger detailed mechanisms.

CRediT authorship contribution statement

Pietro Pagani: Design and implementation of the computational framework, Simulations, Analysis of the results, Writing – original draft. **Riccardo Malpica Galassi:** Conceived the study, Supervision, Analysis of the results, Writing – original draft. **Ruggero Amaduzzi:** Design of the computational framework, Preliminary numerical calculations. **Alessandro Parente:** Supervision. **Francesco Contino:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 714605, Marie Sklodowska-Curie Actions under Grant Agreements No. 861079 and 801505 and from the Walloon Region as part of a FRIA grant funding. All authors approved the final version of the manuscript.

References

[1] A. Cavaliere, M. de Joannon, Mild combustion, Prog. Energy Combust. Sci. 30 (2004) 329–366.
[2] J. Van Oijen, F. Lammers, L. De Goey, Modeling of complex premixed burner systems by using flamelet-generated manifolds, Combust. Flame 127 (2001) 2124–2134.
[3] F. Chitgarha, A. Mardani, Assessment of steady and unsteady flamelet models for MILD combustion modeling, Int. J. Hydrogen Energy 43 (32) (2018) 15551–15563, http://dx.doi.org/10.1016/j.ijhydene.2018.06.071.
[4] X. Huang, M.J. Tummers, E.H. van Veen, D.J. Roekaerts, Modelling of MILD combustion in a lab-scale furnace with an extended FGM model including turbulence–radiation interaction, Combust. Flame 237 (2022) 111634, http://dx.doi.org/10.1016/j.combustflame.2021.111634, URL https://www.sciencedirect.com/science/article/pii/S0010218021003771.
[5] M. Mayrhofer, M. Koller, P. Seemann, H. Bordbar, R. Prieler, C. Hochenauer, MILD combustion of hydrogen and air – An efficient modelling approach in CFD validated by experimental data, Int. J. Hydrogen Energy 47 (9) (2022) 6349–6364, http://dx.doi.org/10.1016/j.ijhydene.2021.11.236.
[6] M. Ihme, Y.C. See, LES flamelet modeling of a three-stream MILD combustor: Analysis of flame sensitivity to scalar inflow conditions, Proc. Combust. Inst. 33 (1) (2011) 1309–1317, http://dx.doi.org/10.1016/j.proci.2010.05.019.
[7] A. Parente, M.R. Malik, F. Contino, A. Cuoci, B.B. Dally, Extension of the eddy dissipation concept for turbulence/chemistry interactions to MILD combustion, Fuel 163 (2016) 98–111.
[8] M. Evans, C. Petre, P. Medwell, A. Parente, Generalisation of the eddy-dissipation concept for jet flames with low turbulence and low Damköhler number, Proc. Combust. Inst. 37 (2018).
[9] Z. Li, M. Ferrarotti, A. Cuoci, A. Parente, Finite-rate chemistry modelling of non-conventional combustion regimes using a partially-stirred reactor closure: Combustion model formulation and implementation details, Appl. Energy 225 (2018) 637–655.
[10] F. Contino, T. Lucchini, G. D'Errico, C. Duynslaegher, V. Dias, H. Jeanmart, Simulations of advanced combustion modes using detailed chemistry combined with tabulation and mechanism reduction techniques, SAE Int. J. Engines 5 (2012) 185–196.
[11] T. Turányi, A.S. Tomlin, Reduction of reaction mechanisms, in: Analysis of Kinetic Reaction Mechanisms, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 183–312.
[12] P. Pepiot-Desjardins, H. Pitsch, An efficient error-propagation-based reduction method for large chemical kinetic mechanisms, Combust. Flame 154 (2008) 67–81.
[13] R. Malpica Galassi, P.P. Ciottoli, S.M. Sarathy, H.G. Im, S. Paolucci, M. Valorani, Automated chemical kinetic mechanism simplification with minimal user expertise, Combust. Flame 197 (2018) 439–448.
[14] The use of dynamic adaptive chemistry in combustion simulation of gasoline surrogate fuels, Combust. Flame 156 (2009) 1493–1502.
[15] R. Malpica Galassi, P.P. Ciottoli, M. Valorani, H.G. Im, An adaptive time-integration scheme for stiff chemistry based on computational singular perturbation and artificial neural networks, J. Comput. Phys. 451 (2022) 110875.
[16] Z. Ren, Y. Liu, T. Lu, L. Lu, O.O. Oluwole, G.M. Goldin, The use of dynamic adaptive chemistry and tabulation in reactive flow simulations, Combust. Flame 161 (2014) 127–137.
[17] Y. Liang, S.B. Pope, P. Pepiot, A pre-partitioned adaptive chemistry methodology for the efficient implementation of combustion chemistry in particle PDF methods, Combust. Flame 162 (2015) 3236–3253.
[18] A. Newale, S. Pope, P. Pepiot, Computationally-efficient and accurate particle PDF simulations of turbulent combustion using coupled pre-partitioned adaptive chemistry and tabulation, Proc. Combust. Inst. 38 (2021) 2721–2729.
[19] Z. Li, M.T. Lewandowski, F. Contino, A. Parente, Assessment of on-the-fly chemistry reduction and tabulation approaches for the simulation of moderate or intense low-oxygen dilution combustion, Energy Fuels 32 (2018) 10121–10131.
[20] F. Contino, H. Jeanmart, T. Lucchini, G. D'Errico, Coupling of in situ adaptive tabulation and dynamic adaptive chemistry: An effective method for solving combustion in engine simulations, Proc. Combust. Inst. 33 (2) (2011) 3057–3064.
[21] S. Pope, Computationally efficient implementation of combustion chemistry using in situ adaptive tabulation, Combust. Theory Model. 1 (1997) 41–63.
[22] G. D'Alessio, A. Parente, A. Stagni, A. Cuoci, Adaptive chemistry via pre-partitioning of composition space and mechanism reduction, Combust. Flame 211 (2020) 68–82.
[23] G. D'Alessio, A. Cuoci, G. Aversano, M. Bracconi, A. Stagni, A. Parente, Impact of the partitioning method on multidimensional adaptive-chemistry simulations, Energies 13 (2020) 2567.
[24] B. Dally, A. Karpetis, R. Barlow, Structure of turbulent non-premixed jet flames in a diluted hot coflow, Proc. Combust. Inst. 29 (2002) 1147–1154.
[25] N. Kambhatla, T.K. Leen, Dimension reduction by local principal component analysis, Neural Comput. 9 (1997) 1493–1516.
[26] A.S. Newale, P. Sharma, S.B. Pope, P. Pepiot, A feasibility study on the use of low-dimensional simulations for database generation in adaptive chemistry approaches, Combust. Theory Model. (2022) 1239–1261.
[27] N. Peters, Laminar flamelet concepts in turbulent combustion, Symp. (Int.) Combust. 21 (1988) 1231–1250.
[28] I. Jolliffe, Principal component analysis, in: M. Lovric (Ed.), International Encyclopedia of Statistical Science, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 1094–1096.
[29] A. Parente, J. Sutherland, B. Dally, L. Tognotti, P. Smith, Investigation of the MILD combustion regime via principal component analysis, Proc. Combust. Inst. 33 (2011) 3333–3341.
[30] A. Parente, J. Sutherland, L. Tognotti, P. Smith, Identification of low-dimensional manifolds in turbulent flames, Proc. Combust. Inst. 32 (2009) 1579–1586.
[31] D.L. Davies, D.W. Bouldin, A cluster separation measure, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-1 (2) (1979) 224–227.
[32] R. Malpica Galassi, PyCSP: A python package for the analysis and simplification of chemically reacting systems based on computational singular perturbation, Comput. Phys. Comm. 276 (2022) 108364.
[33] M. Valorani, F. Creta, P.P. Ciottoli, R. Malpica Galassi, D.A. Goussis, H.N. Najm, S. Paolucci, H.G. Im, E.-A. Tingas, D.M. Manias, A. Parente, Z. Li, T. Grenga, in: H. Pitsch, A. Attili (Eds.), Data Analysis for Direct Numerical Simulations of Turbulent Combustion: From Equation-Based Analysis to Machine Learning, Springer International Publishing, Cham, 2020, pp. 43–64.
[34] M. Valorani, F. Creta, P.P. Ciottoli, R. Malpica Galassi, D.A. Goussis, H.N. Najm, S. Paolucci, H.G. Im, E.-A. Tingas, D.M. Manias, A. Parente, Z. Li, T. Grenga, in: H. Pitsch, A. Attili (Eds.), Data Analysis for Direct Numerical Simulations of Turbulent Combustion: From Equation-Based Analysis to Machine Learning, Springer International Publishing, Cham, 2020, pp. 65–88.
[35] M. Valorani, F. Creta, D.A. Goussis, J.C. Lee, H.N. Najm, An automatic procedure for the simplification of chemical kinetic mechanisms based on CSP, Combust. Flame 146 (1) (2006) 29–51, http://dx.doi.org/10.1016/j.combustflame.2006.03.011.

[36] T. Lu, C.K. Law, Linear time reduction of large kinetic mechanisms with directed relation graph: n-heptane and iso-octane, Combust. Flame 144 (1) (2006) 24–36, http://dx.doi.org/10.1016/j.combustflame.2005.02.015.

[37] T. Zeuch, G. Moréac, S.S. Ahmed, F. Mauss, A comprehensive skeletal mechanism for the oxidation of n-heptane generated by chemistry-guided reduction, Combust. Flame 155 (4) (2008) 651–674, http://dx.doi.org/10.1016/j.combustflame.2008.05.007.

[38] A.S. AlRamadan, R. Malpica Galassi, P.P. Ciottoli, M. Valorani, S.M. Sarathy, Multi-stage heat release in lean combustion: Insights from coupled tangential stretching rate (TSR) and computational singular perturbation (CSP) analysis, Combust. Flame 219 (2020) 242–257.

[39] S.D. Cohen, A.C. Hindmarsh, CVODE, a stiff/nonstiff ODE solver in C, Comput. Phys. 10 (2) (1996) 138–143.

[40] K. Zdybał, E. Armstrong, A. Parente, J.C. Sutherland, PCAfold: Python software to generate, analyze and improve PCA-derived low-dimensional manifolds, SoftwareX 12 (2020) 100630, http://dx.doi.org/10.1016/j.softx.2020.100630, URL http://www.sciencedirect.com/science/article/pii/S2352711020303435.

[41] FiReSMOKE, a collection of finite-rate chemistry solvers for combustion simulations for OpenFOAM (https://github.com/burn-research/FiReSMOKE).

[42] G.P. Smith, D.M. Golden, M. Frenklach, N.W. Moriarty, B. Eiteneer, M. Goldenberg, C.T. Bowman, R.K. Hanson, S. Song, W.C. Gardiner Jr., V.V. Lissianski, Z. Qin, 2009. http://www.me.berkeley.edu/gri-mech/.

[43] A. Péquin, S. Iavarone, R. Malpica Galassi, A. Parente, The partially stirred reactor model for combustion closure in large eddy simulations: Physical principles, sub-models for the cell reacting fraction, and open challenges, Phys. Fluids 34 (2022) 055122.