A priori examination of reduced chemistry models derived from canonical stirred reactors using three-dimensional direct numerical simulation datasets

Pei Zhang*, Ramanan Sankaran†
Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Evatt R. Hawkes
School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney, NSW 2021, Australia

Data-driven approaches to construct reduced chemical kinetic models, that rely heavily on thermo-chemical datasets with full chemical kinetics, have been gaining popularity. Datasets from direct numerical simulations (DNS) under three-dimensional (3-D) realistic turbulent flow conditions are desirable but limited to carefully designed parametric conditions due to the computational cost. Constructing datasets from a large ensemble of zero-dimensional stirred reactors like perfectly stirred reactor (PSR) and partially stirred reactor (PaSR) is a computationally efficient solution to consider the turbulence-chemistry interactions and cover a broad range of parametric conditions. In this paper, we derive reduced chemistry models from solutions of a large number of PSR and PaSR reactors using autoencoder (AE) neural networks and principal component analysis (PCA), and conduct a priori examination of the reduced models in three temporally evolving 3-D DNS jet flames featuring local extinction and re-ignition. The results show that the reduced models derived from PaSR datasets, *i.e.*, AE-PaSR and PCA-PaSR, generally show significant improvement over the ones derived from PSR datasets. Among all the reduced models, AE-PaSR shows the best agreement with DNS results on the reconstruction accuracy and the representation of temporally evolving local extinction and re-ignition events.

I. Introduction

An accurate description of chemical kinetics is essential for combustion simulation and hence the design of next generation propulsion devices. Such a description with detailed chemical kinetic models for hydrocarbon fuels typically involves hundreds of species and an even larger number of elementary reactions. Inclusion of complex detailed chemistry models in multidimensional computational fluid dynamics (CFD) simulations for combustion engines strains the computational capability. In spite of the large number of species in the detailed chemical kinetic models, it is observed that there exist low dimensional manifolds (LDMs) in the high-dimensional thermo-chemical state space. Reduced order models (ROMs) for chemical kinetics have been driven by the need to reduce the computational cost and by taking advantage of the existence of LDMs.

Most of previous LDM reduced chemistry modeling work involves physical assumptions or simplifications of the problem and consequently the models developed are applicable to certain regime of problems, such as flamelet models [1] and conditional moment closure (CMC) [2] for non-premixed combustion and flamelet generated manifold (FGM) [3] and flamelet prolongation of intrinsic LDM (FPI) [4] for premixed combustion. Recently, data-driven LDM modeling approaches, such as principal component analysis (PCA) [5–11] and neural networks [12–16] have also been applied to derive reduced chemistry models, wherein LDM models are derived from sampling datasets and hence providing a data-based approach that is complementary to a physics-based model reduction. The datasets consist of thermo-chemical state variables from experiments or simulations with a detailed chemical kinetic model. A reduced

*Member AIAA
†Associate Fellow AIAA

number of variables, such as principal components in PCA or the code layer variables in the autoencoder (AE) neural networks, are defined as a function of the thermo-chemical state variables based on the datasets to describe the chemical system.

Datasets are critically important in data-driven methods for dimension reduction. For the reduced chemistry modeling, ideally datasets from high-fidelity three-dimensional (3-D) direct numerical simulation (DNS) with detailed chemistry are desired, wherein the thermo-chemical states come from turbulent flow conditions. The high expense of generating DNS datasets requires careful selection of the parametric conditions for such DNS. There exists an opportunity for supplementing the data from a large ensemble of canonical configurations, such as zero-dimensional (0-D) and one-dimensional (1-D) reactors that can efficiently cover a broader range of conditions. These canonical reactors can be calculated at a much smaller computational cost with detailed chemical kinetic models. From solutions of an unsteady perfectly stirred reactors (PSR), Malik et al. [6] applied PCA in conjunction with nonlinear regression models to derive reduced chemistry for methane and propane combustion. Rahimi et al. [17] derived reduced chemistry model for ethylene oxidization from solutions of partially stirred reactor (PaSR). In the 0-D reactors, the effects of flow fields are approximately described by using two time scales, *i.e.*, residence time scale τ_r and mixing time scale τ_m, and the composition of inflow streams. Ren and Pope [18] studied the effects of different mixing models and time scales in PaSR reactors. Recently, Zhang et al. [16] derived reduced chemistry for syngas CO/H_2 combustion from datasets composed of 2.3 millions of steady PSR reactors covering a wide range of residence time scale and composition of inflow stream to explore the state space that might be encountered in 3-D flows. The datasets composed of a large ensemble of the 0-D/1-D reactors at various mixture conditions and time scales can be used to derive ROMs for kinetics which are then verified and improved by using results from 3-D DNS cases. Such a workflow provides an efficient data-based workflow for developing reduced order combustion models, which can then be deployed in 3-D device scale CFD simulations for validation against experiments.

In this work, we derive reduced chemistry models from a large ensemble of canonical reactors with detailed chemistry and verify the ability to represent the state space realized in 3-D DNS with detailed chemistry. The reduced chemistry models in Refs. [16, 19] were based on PCA and AE neural networks and were derived from the thermo-chemical dataset of CO/H_2 combustion in 0-D steady PSR at various inflow mixture conditions and residence time scales, designed based on the target 3-D high-fidelity DNS cases in Ref. [20]. In this work, we extend the approach to develop ROMs using datasets obtained from PaSR. We conduct an a priori examination of the ROMs in the DNS flames. The 3-D DNS cases are temporally evolving planar turbulent jet flames at Reynolds numbers varying from 2510 to 9079 and showing strong flame-turbulence interactions with local extinction and re-ignition. In the results section, we present the examination results of the reduced models with regard to the reconstruction accuracy of thermo-chemical state variables and the performance in capturing the local extinction and re-ignition behaviors.

II. Methodology

In the present work, the reduced order models for chemical kinetics are developed by using data-driven dimension reduction methods, *i.e.*, AE neural networks and PCA. The datasets used are composed of solutions of a large ensemble of 0-D stirred reactors with full chemical kinetics. The objective of the paper is to evaluate the accuracy of the reduced models derived from 0-D reactors in 3-D DNS turbulent jet flames. In the following, we summarize the AE and PCA methods, the configurations of the stirred reactors and the DNS flames.

A. Reduced Order Models

AE and PCA are two popular dimension reduction techniques that have been applied in various scientific problems, such as anomaly detection [21], molecular dynamics simulations [22] and biosciences [23]. Essentially, they are used to find the mappings between the original full system $\vec{\boldsymbol{x}} = [x_1, x_2, \dots, x_{n_x}]^\intercal$ and the reduced system $\vec{\boldsymbol{z}} = [z_1, z_2, \dots, z_{n_z}]^\intercal$ with much smaller number of dimensions, *i.e.*, $\mathscr{F} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ and $\mathscr{G} : \mathbb{R}^{n_z} \to \mathbb{R}^{n_x}$, where $n_z \ll n_x$. In PCA, linear mappings between $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{z}}$ are defined based on a subset of eigenvectors that preserve the largest data variance. In AE, nonlinear mappings are defined by bottleneck neural networks consisting of an encoder, a decoder and a bottleneck latent/code layer in-between. The encoder part takes the high-dimensional $\vec{\boldsymbol{x}}$ as input and compresses it to the reduced $\vec{\boldsymbol{z}}$, and the decoder takes $\vec{\boldsymbol{z}}$ and reconstructs it back to the original space as $\hat{\vec{\boldsymbol{x}}}$. The AE model learns by finding the optimal weights that minimize the difference between the reconstructed output data $\hat{\boldsymbol{X}} = [\hat{\vec{\boldsymbol{x}}}_1, \hat{\vec{\boldsymbol{x}}}_2, \dots, \hat{\vec{\boldsymbol{x}}}_{n_{\text{samp}}}]$ and the original input data $\boldsymbol{X} = [\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \dots, \vec{\boldsymbol{x}}_{n_{\text{samp}}}]$ where n_{samp} is the number of samples. PCA can be reproduced by AE with linear activation functions. More details about AE and PCA can be found in [16].

In the reduced chemistry work, $\vec{x}$ is taken as the scaled state vector in the full thermo-chemical space, $\vec{x} = 2(\vec{\phi} - \vec{\phi}_{\min})/(\vec{\phi}_{\max} - \vec{\phi}_{\min}) - 1$ with the state vector being $\vec{\phi} = [\rho, T, Y_1, Y_2, \dots, Y_{n_s}]^{\mathrm{T}}$, where ρ is density, T is temperature, and Y_α ($\alpha = 1, \dots, n_s$) are the mass fractions of the n_s chemical species. The dimensionality $n_x = n_s + 2$ is very high when complex detailed chemical kinetic model is used, which also represents the large number of transport equations needed to be solved in CFD. With AE and PCA reduced models, the thermo-chemical state of a chemically reacting system can be represented by $\vec{z}$ and the number of transport equations also reduces from n_x to n_z. For CO/H_2 combustion, it has been found in [16] that AE model can reduce the dimensionality from $n_x = 12$ to $n_z = 2$.

B. Configurations of canonical stirred reactors

The canonical stirred reactors used to derive reduced chemical kinetic models are PSR [24] and PaSR [18, 25–28]. The configuration of PaSR is solely presented in the following considering that PSR is a limiting case of PaSR and a summary of PSR can be found in our previous work [16].

The PaSR is a 0-D reactor configuration with inlets and outlets. The fluid enters and exits the reactor at the mass flow rate, $\dot{m}$. Both the total mass of the fluid inside reactor, M, and the mass flow rate, $\dot{m}$, are held constant in the work. The residence time scale τ_r is defined as the ratio, $\tau_r = M/\dot{m}$. Inside the reactor, the fluid exhibits a coexistence of nonhomogeneous thermo-chemical states and undergoes turbulent mixing at a rate determined by the mixing time scale, τ_m.

The state distribution of a PaSR reactor can be described by a joint probability density function (PDF) $f(\vec{\psi}; t)$ or mass density function (MDF) $\mathcal{F}(\vec{\psi}; t) = \rho f$ [29] of the thermo-chemical state vector $\vec{\phi}$ with $\vec{\psi}$ being the corresponding sample space vector. The state vector $\vec{\phi} = [\phi_1, \phi_2, \dots, \phi_{n_s+1}]$ consists of mass fractions of species and enthalpy/temperature for combustion under atmospheric pressure in this work. The governing equation of the MDF, $\mathcal{F}(\vec{\psi}; t)$, in the PaSR reactor is

$$\frac{\partial \mathcal{F}}{\partial t} = -\sum_\alpha \frac{\partial}{\partial \psi_\alpha}\left[\left(\omega_\alpha(\vec{\psi}) - \frac{\psi_\alpha - \widetilde{\phi}_\alpha}{\tau_m}\right)\mathcal{F}\right] + \frac{\mathcal{F}_{\mathrm{in}} - \mathcal{F}}{\tau_r}, \tag{1}$$

where $\mathcal{F}_{\mathrm{in}}$ is the MDF of inflow streams and $\widetilde{\square}$ is the density-weighted variable mean, $\widetilde{\square} = \langle \rho \square \rangle / \langle \rho \rangle$. The first term on the right hand side (RHS) represents the MDF changes due to chemical reaction ω_α and mixing where the Interaction by Exchange with the Mean (IEM) model [30, 31] with a mixing time scale τ_m is applied. At the limiting case of fast mixing with infinitesimal τ_m, the PaSR tends to reproduce the PSR with a delta-distribution for $\mathcal{F}$, while at the opposite extreme with infinite τ_m, it tends to behave like the plug-flow reactor with almost no mixing.

For a given set of τ_m and τ_r values, the solution of Eq. (1) is specified by the initial condition, $\mathcal{F}(\vec{\psi}; 0) = \mathcal{F}_0$, and the inflow MDF, $\mathcal{F}_{\mathrm{in}}$. In the work by Correa [25, 26], a premixed mixture as the inflow condition is considered to study lean premixed combustion. Mobini and Bilger [32] proposed the imperfectly stirred reactor (ISR) that has a similar configuration to PaSR and allows the inflow to have a range of conditions from completely unmixed to partly premixed. Rahimi *et al.* [17] and Ren and Pope [18] used two inflow streams with $\mathcal{F}_{\mathrm{in}}$ being represented as a two-delta function. The two inflow streams are a fresh stream of unburned fuel/air mixture and a pilot stream at the equilibrium state of the fresh mixture in [17] and a fresh fuel stream and a fresh oxidizer stream in [18]. For the results reported in the paper, two inflow streams, *i.e.*, a fuel stream $\vec{\phi}_{\mathrm{F,in}}$ and an oxidizer stream $\vec{\phi}_{\mathrm{O,in}}$ at a global equivalence ratio Φ, are considered, or equivalently, $\mathcal{F}_{\mathrm{in}}(\vec{\psi}) = \rho\left[\gamma\delta(\vec{\psi} - \vec{\phi}_{\mathrm{F,in}}) + (1-\gamma)\delta(\vec{\psi} - \vec{\phi}_{\mathrm{O,in}})\right]$ with γ determined by equivalence ratio Φ, $\gamma = \Phi/\left(\Phi + (1/\xi_{\mathrm{st}} - 1)\right)$, where ξ_{st} is the stoichiometric mixture fraction. The initial condition is taken as the equilibrium state at the Φ, $\mathcal{F}_0 = \rho\delta(\vec{\psi} - \vec{\phi}_{\mathrm{eq}})$.

Eq. (1) is a high-dimensional PDE that is typically solved by using Monte Carlo particle method, where the MDF of PaSR is discretely represented by a stochastically equivalent particle system with N_p particles. Each particle has a set of properties that are mass $m^{(n)}$ and thermo-chemical state vector $\vec{\phi}^{(n)}$ ($n = 1, \dots, N_p$). In this work, every particle has equal mass, $m^{(n)} = m_p = M/N_p$. A system of differential equations can be developed to describe the evolving of particle states and hence the discrete MDF. Specifically, the governing equation for a particle state vector, $\vec{\phi}^{(n)}$, is

$$\frac{d\vec{\phi}^{(n)}}{dt} = -\frac{\vec{\phi}^{(n)} - \vec{\widetilde{\phi}}}{\tau_m} + \vec{\omega}^{(n)}, \tag{2}$$

which represents the contributions of mixing and chemical reaction to the state change, corresponding to the first term on the RHS of Eq. (1). The second term on the RHS of Eq. (1) represents the changes of MDF inside PaSR caused by

the inflow and outflow streams. When solving particle equations, at every time step, N_{in} particles at the inflow state ($\vec{\phi}_{\text{F,in}}$ or $\vec{\phi}_{\text{O,in}}$) enter the reactor and the same number of particles inside are randomly selected to flow out of the reactor. The value of N_{in} is calculated as

$$N_{\text{in}} = \frac{\dot{m}\Delta t}{m_p} = \frac{M\Delta t}{\tau_r m_p} = \frac{N_p \Delta t}{\tau_r}, \tag{3}$$

where Δt is the time step size.

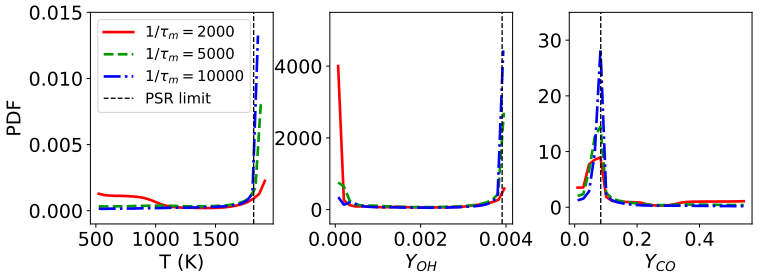

Fig. 1 PDFs of temperature and mass fractions of OH and CO from three PaSR reactors at $\Phi = 1.0$, $1/\tau_r = 2000$ (1/s), and $1/\tau_m = [2000, 5000, 10000]$ (1/s).

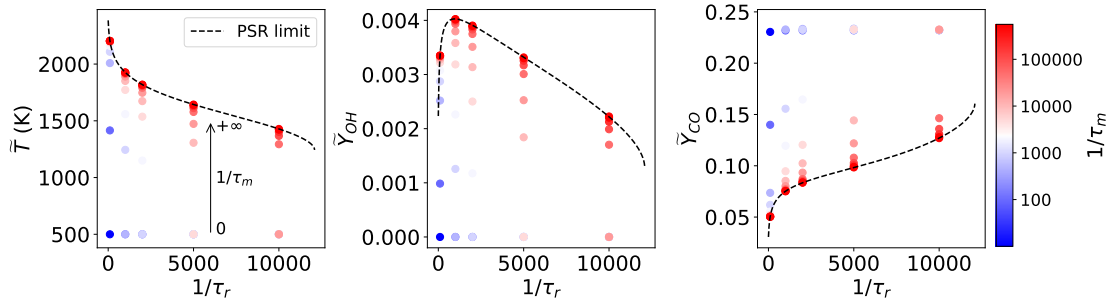

Fig. 2 Effects of mixing time scale τ_m on mean PaSR solutions (symbols) at $\Phi = 1.0$ and a few selected $1/\tau_r = [100, 1000, 2000, 5000, 10000]$ (1/s).

In this study, the same syngas fuel and oxidizer as in the DNS [20] are used, wherein the fuel is composed of 50% CO, 10% H_2 and 40% N_2 and the oxidizer is composed of 25% O_2 and 75% N_2 by volume. The same chemical mechanism for syngas oxidization as in [20] with 11 species and 21 reactions is used in the dataset generation. The number of particles is taken as $N_p = 5000$ and the time step size is $\Delta t = 1 \times 10^{-7}$ s. Fig. 1 shows some sample solutions of PaSR reactors at three mixing time scales $1/\tau_m = [2000, 5000, 10000]$ (1/s), and a fixed $\Phi = 1.0$ and $1/\tau_r = 2000$ (1/s). The PSR solution is also shown as the thin black dashed line. It can be seen that the PDF peaks of the scalars, *i.e.*, temperature T and mass fractions of OH and CO, Y_{OH} and Y_{CO}, move toward the PSR solution location with the mixing frequency $1/\tau_m$ increasing. The same observation can be made in Fig. 2, which shows the mean solutions, *i.e.*, mean temperature $\widetilde{T}$ and mean mass fractions of OH and CO, $\widetilde{Y}_{\text{OH}}$ and $\widetilde{Y}_{\text{CO}}$, of multiple PaSR reactors. For a given composition Φ and residence frequency $1/\tau_r$, the PaSR state moves from plug flow state to the PSR state as the mixing frequency $1/\tau_m$ increases from 0 to infinity. For a fixed $1/\tau_m$, temperature decreases with increasing $1/\tau_r$ due to less reaction time inside reactor or more heat loss to inflow/outflow streams, consistent with PSR. The datasets used to derive the reduced chemistry models are composed of PaSR solutions at various values of Φ, τ_r and τ_m within the range of interest.

C. Turbulent DNS jet flames

The 3D DNS flames considered here are the temporally evolving non-premixed planar jet flames [20]. The flames reside in turbulent fields formed of three streams that are the central fuel stream composed of 50% CO, 10% H_2 and 40% N_2 and the two surrounding oxidizer streams from the opposite direction composed of 25% O_2 and 75% N_2 by

volume. These flames are designed with a low Damköhler number, and exhibit strong flame-turbulence interactions with local extinction and re-ignition. Three cases, *i.e.*, L, M and H, are available with increasing jet Reynolds numbers from 2510 to 9079 and and also increasing degree of local extinction.

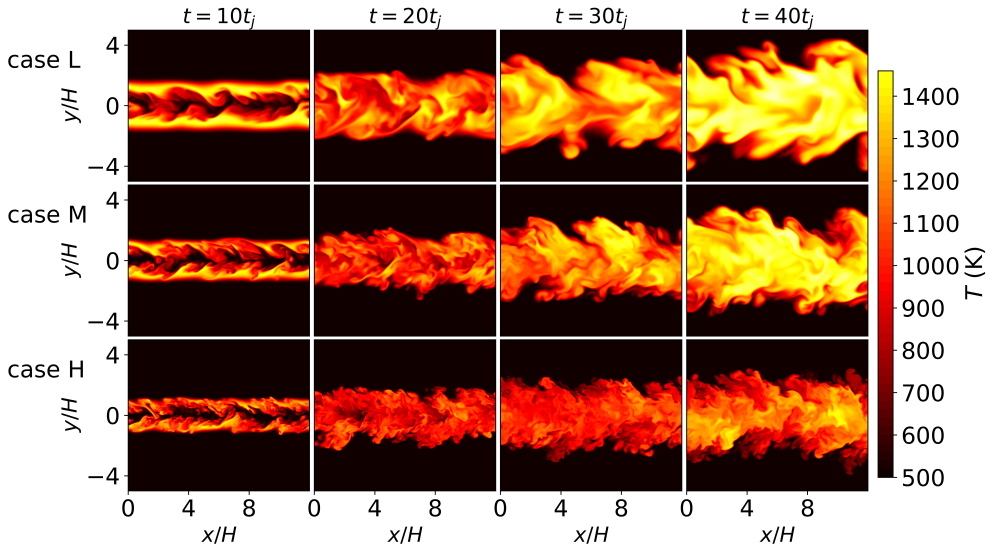

Fig. 3 Streamwise cuts of temperature T (K) at four times $t = [10t_j, 20t_j, 30t_j, 40t_j]$ for cases L, M and H. The jet time scale is $t_j = 5.0 \times 10^{-6}$ (s) for all cases and H is the planar jet height [20].

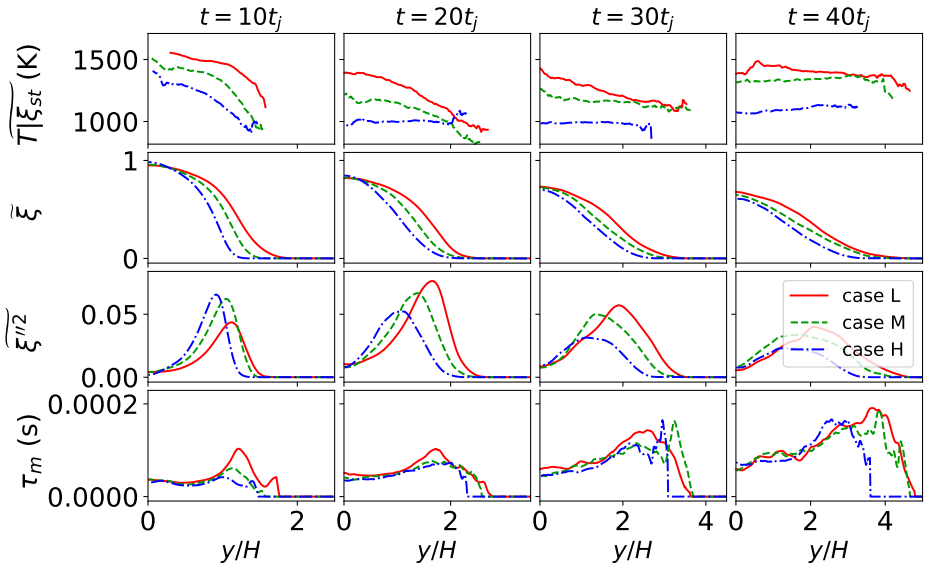

Fig. 4 Spatial profiles of stoichiometric conditional mean temperature $\widetilde{T|\xi_{st}}$, mean and variance mixture fraction $\widetilde{\xi}$, $\widetilde{\xi''^2}$, and mixing time scale τ_m at four times $t = [10t_j, 20t_j, 30t_j, 40t_j]$ for cases L, M and H.

DNS results at four times $t = [10t_j, 20t_j, 30t_j, 40t_j]$, where t_j is the transient jet time scale, are considered in the current study. Fig. 3 shows the instantaneous temperature contours of the three cases, L, M and H, at the four times. It is seen that from case L to H, the flame temperature decreases and increasingly fine-scale flame structures are observed. For each case, the flame evolves from the initial laminar flamelet solution to fully turbulent state under the intense turbulent mixing at later times. From the spatial profiles of stoichiometric conditional mean temperature $\widetilde{T|\xi_{st}}$ in the first row of Fig. 4, it is seen that the conditional mean flame temperature in all cases decreases from the initial laminar

value, reaches to the minimum value at $20t_j \sim 30t_j$, and increases at $40t_j$ due to re-ignition. It is also reported in [20] that the flames show approximately maximum amount of local extinction at around $t = 20t_j$ and various levels of re-ignition at $t = 40t_j$ where case L is largely re-ignited while cases M and H are not fully re-ignited yet. It is of interest to know if the reduced chemistry models can reproduce the local extinction and re-ignition behaviors in the DNS flames. Also shown in Fig. 4 are the spatial profiles of mean and variance mixture fraction $\widetilde{\xi}$, $\widetilde{\xi''^2}$, and mixing time scale τ_m. Mixing time scale is defined based on mixture fraction as $\tau_m = \frac{2\widetilde{\xi''^2}}{\widetilde{\chi}_\xi}$ with the residual scalar dissipation rate $\widetilde{\chi}_\xi = \langle 2\rho D_\xi \nabla\xi \cdot \nabla\xi \rangle / \langle\rho\rangle - 2\widetilde{D}_\xi \nabla\widetilde{\xi} \cdot \nabla\widetilde{\xi}$. These statistical information can be used to guide the PaSR/PSR parametric configurations in the generation of datasets.

D. Model examination

In the paper, we derive the reduced chemical kinetic models from datasets composed of solutions of 0-D stirred reactors in section II.B using the methods in section II.A, and examine the performance of the reduced models in the 3-D DNS flames in section II.C.

Two sets of AE reduced models are examined in the following. One set is from Refs. [16, 19] derived based on datasets from steady PSR reactors and is referred to as AE-PSR hereinafter. The AE-PSR model has a simple architecture with one hidden layer for the encoder and decoder parts, respectively. In total, it has five dense layers, *i.e.*, input layer, hidden layer with n_x neurons in the encoder, latent/code layer with n_z neurons, hidden layer with n_x neurons in the decoder and output layer. The hyperbolic tangent (Tanh) activation function is used. The loss function is defined as, $\text{Loss}_{\text{AE}} = \text{MSE}(\boldsymbol{X}, \hat{\boldsymbol{X}}) + \varepsilon_{\text{ele}}$, where the first term represents the reconstruction accuracy as the mean square error (MSE) between the original dataset $\boldsymbol{X}$ and the reconstructed dataset $\hat{\boldsymbol{X}}$. The second term, $\varepsilon_{\text{ele}} = \max(|\boldsymbol{Y}_C - \hat{\boldsymbol{Y}}_C|/\boldsymbol{Y}_C, |\boldsymbol{Y}_H - \hat{\boldsymbol{Y}}_H|/\boldsymbol{Y}_H, |\boldsymbol{Y}_O - \hat{\boldsymbol{Y}}_O|/\boldsymbol{Y}_O, |\boldsymbol{Y}_N - \hat{\boldsymbol{Y}}_N|/\boldsymbol{Y}_N)$, represents the penalty due to elemental conservation error, where $\boldsymbol{Y}_i$ is the mass fraction of element $i = \text{C/H/O/N}$. The model is implemented in Keras 2.3.0 with TensorFlow backend [33]. It is trained on training dataset composed of about 1.14 millions of steady PSR solutions and is evaluated on the test dataset composed of about 0.49 millions of steady PSR solutions. From the evaluation results in [16, 19], it has been concluded that the AE model with $n_z = 2$ can represent the syngas combustion in steady PSR reactors with a wide range of parametric conditions accurately. The other set of AE model, which is designed to have the same model architecture as AE-PSR, is trained with new datasets composed of statistically stationary PaSR solutions and is referred to as AE-PaSR. Solution data of 2084 PaSR reactors are collected, or equivalently 10.42 millions of discrete combustion states, 70% of which are used to train the new AE-PaSR model.

III. Results

This section presents the examination results of the trained AE-PSR and AE-PaSR reduced models with $n_z = 2$ in the 3-D DNS turbulent flames. Results of the corresponding linear PCA models are also provided as a reference. The models are examined from two aspects, *i.e.*, the reconstruction accuracy of thermo-chemical state variables and the reproduction accuracy of local extinction and re-ignition events.

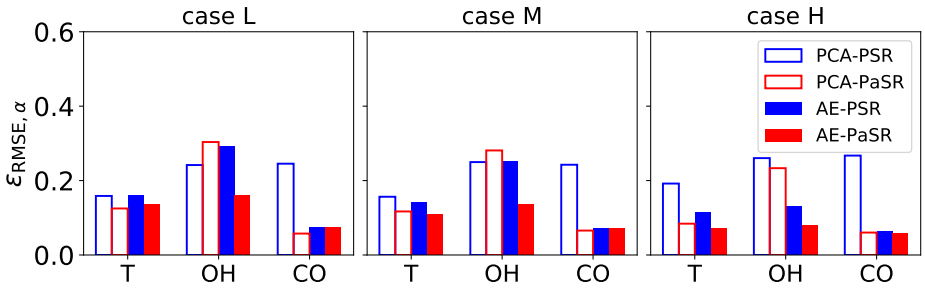

Fig. 5 Global relative RMSE error of T and mass fractions of OH and CO for cases L, M and H.

Fig. 5 shows the global relative Root Mean Square Error (RMSE),

$$\varepsilon_{\mathrm{RMSE},\alpha} = \frac{1}{\phi_{\alpha,\max} - \phi_{\alpha,\min}} \sqrt{\frac{1}{N_{\mathrm{DNS}}} \sum_{i=1}^{N_{\mathrm{DNS}}} \left(\phi_\alpha^{(i)} - \hat{\phi}_\alpha^{(i)}\right)^2}, \tag{4}$$

of scalar $\boldsymbol{\phi}_\alpha = T/Y_{\mathrm{OH}}/Y_{\mathrm{CO}}$ for all three DNS cases L, M and H, where $\phi_\alpha^{(i)}$ is the DNS value in one grid cell and $\hat{\phi}_\alpha^{(i)}$ is the corresponding reconstructed result using the reduced models, *i.e.*, PCA-PSR, PCA-PaSR, AE-PSR and AE-PaSR. The global relative RMSE error $\varepsilon_{\mathrm{RMSE},\alpha}$ is calculated based on values from the DNS cells ($i = 1, \ldots, N_{\mathrm{DNS}}$) at all four times $t = [10t_j, 20t_j, 30t_j, 40t_j]$ that are in the main jet flow area roughly defined as $\tilde{\xi}^{(i)} > 0.001$. From Fig. 5, it is seen that the reduced models, especially the AE reduced models, trained based on PaSR reactors (red bars) generally show improved reconstruction accuracy compared to the ones trained based on PSR reactors (blue bars). The improvement is as expected considering that the interactions of chemistry and turbulent mixing have been incorporated in the PaSR reactors and therefore the combustion states from the PaSR reactors are expected to be a better representation of combustion manifolds of the DNS flames. Some exceptions are observed for PCA reduced models in case L and M for OH reconstruction where PCA-PaSR shows a slightly larger error than PCA-PSR. This can be explained by the inadequacy of the generic linear PCA with $n_z = 2$ in representing the complex combustion states in the datasets as reported in [16]. More advanced PCA models, such as in [12], are available but beyond the scope of this paper.

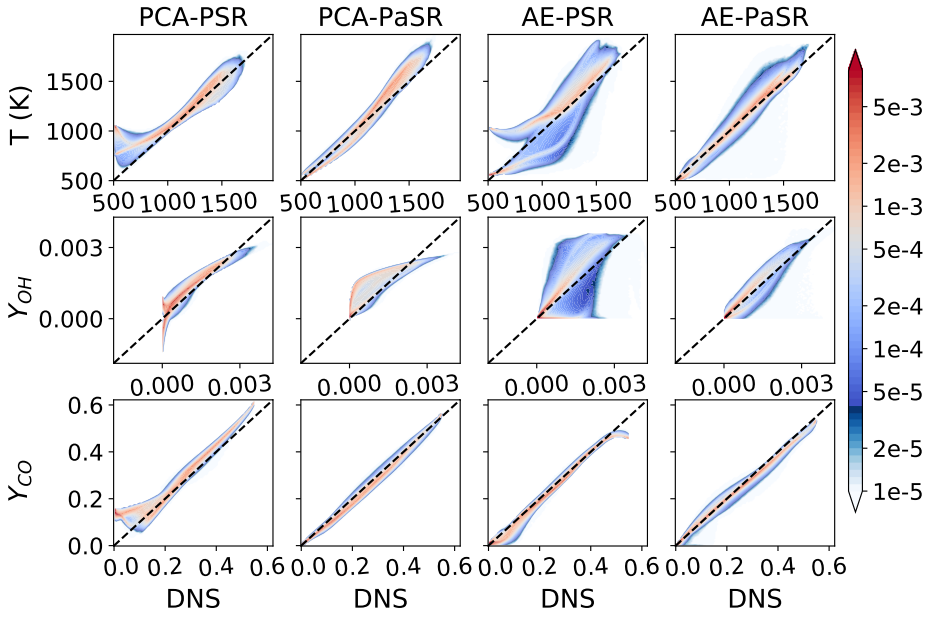

Fig. 6 Normalized joint histogram contours of T and mass fractions of OH and CO from DNS, and from PCA and AE models, for case L.

Figs. 6 and 7 show the normalized joint histogram contours of temperature and mass fractions of OH and CO from DNS and from the reduced models, *i.e.*, PCA-PSR, PCA-PaSR, AE-PSR and AE-PaSR, for two cases L and H. The histogram are calculated based on DNS results from all four times. From the figures, it is seen that all the reduced models perform well in reconstructing CO for both DNS flames. For temperature reconstruction in the first row, the reduced models trained from PaSR reactors show significant improvement compared to the ones from PSR reactors, especially at the relatively low temperature regions where both the PCA-PSR model in the first column and the AE-PSR model in the third column show an over-prediction with a two-branch structure. The multi-mode behaviors are caused by the solution distribution of steady PSR reactors, which is shown as the S-curve with three branches [34] and consequently the combustion states between the branches are missing in the PSR datasets used to train PCA-PSR and AE-PSR models. The top peak histogram locations for $T_{\mathrm{DNS}} = 500$ K of PCA-PSR and AE-PSR models are around $T = 1000$ K, which is roughly the lower end of the middle branch on the S-curve as shown in [16, 19]. For the reconstruction of OH, the reduced models trained from PaSR reactors also show significant improvement compared to

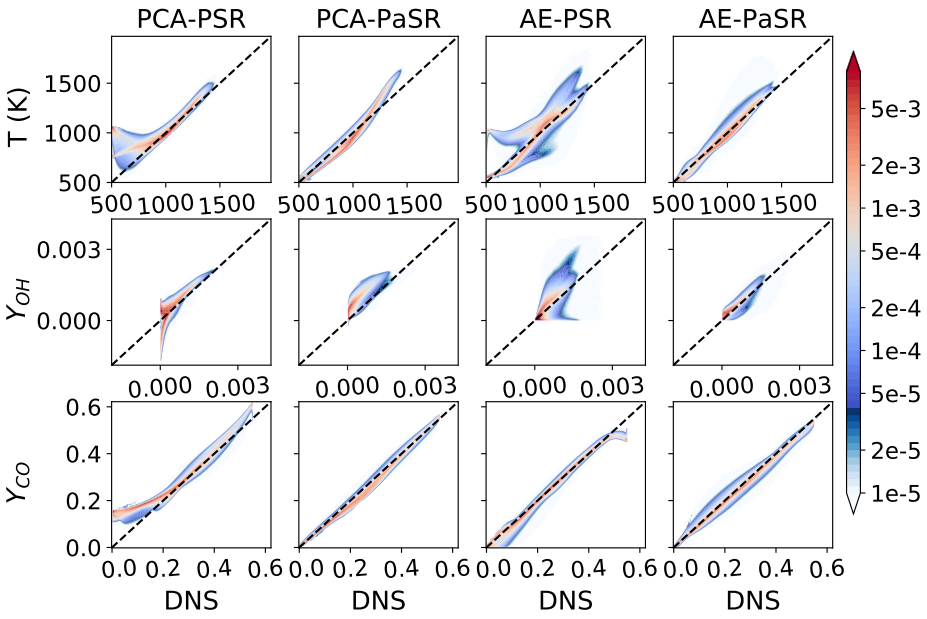

Fig. 7 Normalized joint histogram contours of T and mass fractions of OH and CO from DNS, and from PCA and AE models, for case H.

the ones from PSR reactors, even for case L in which we observed the opposite trend for the global RMSE error of PCA reduced models in Fig. 5. The PCA-PSR models in the first column of Figs. 6 and 7 are both observed to predict nonphysical negative OH mass fractions. The nonphysical behaviors are largely eliminated in the PCA-PaSR models.

Thus far, we have compared the reconstruction accuracy of the four reduced models, PCA-PSR, PCA-PaSR, AE-PSR and AE-PaSR. It is observed that the reduced models trained from PaSR reactors, *i.e.*, PCA-PaSR and AE-PaSR, generally show significant improvement over the PCA-PSR and AE-PSR reduced models. Among all the models, AE-PaSR performs the best. The above observations are made based on DNS results from all the four times $t = [10t_j, 20t_j, 30t_j, 40t_j]$. It is known that the DNS flames are temporally evolving flames experiencing various level of local extinction and re-ignition. In the next, we assess if the reduced models can capture the local extinction and re-ignition temporal behaviors.

Fig. 8 shows the PDF of OH mass fraction conditional on mixture fraction at two times $t = [20t_j, 40t_j]$ for the three cases L, M and H. The time $t = 20t_j$ approximately is when the maximum level of local extinction occurs and the later time $t = 40t_j$ corresponds to the time when the flame is re-ignited, as reported in [20]. The same observations can be made from the DNS conditional PDF results in the first row of Fig. 8. The PDF peak locations move toward higher OH mass fraction locations from $t = 20t_j$ to $t = 40t_j$. Also, from case L, to M and H, more local extinction is observed with PDF peaks moving toward low OH mass fraction direction. Plots from the second row to fifth row show the reconstruction results of model PCA-PSR, PCA-PaSR, AE-PSR and AE-PaSR, respectively. Overall, all the reduced models capture the increasing trend of PDF peak locations from $20t_j$ to $40t_j$ as in DNS, as well as the decreasing trend of PDF peak locations from case L to M and H. PCA reduced models are observed to predict narrower conditional PDF distributions than DNS and AE reduced models, which is attributed to the limited combustion states that a linear model can represent. A large proportion of OH mass fraction values are predicted to be negative in PCA-PSR but very few in PCA-PaSR, consistent with the observations in Figs. 6 and 7. The AE-PSR model in the second last row of Fig. 8 overpredicts the maximum OH mass fraction and shows a large amount of local extinction events with zero OH mass fractions at lean mixture fraction regions for all the three cases at $20t_j$, compared to DNS. The agreement with DNS is improved in the AE-PaSR model shown in the last row, in which the maximum values are decreased and more nonzero OH mass fractions are found at the fuel lean regions at $20t_j$. Overall, the reduced models derived from PaSR show better accuracy and the AE reduced models outperform the PCA reduced models. The observations are more noticeable in Fig. 9.

Fig. 9 shows the comparison of conditional mean OH mass fractions on mixture fraction, $\widetilde{Y_{\mathrm{OH}}|\xi}$, from DNS (empty circles) and from the four models (lines), for the same cases as in Fig. 8. All models are observed to capture the trend

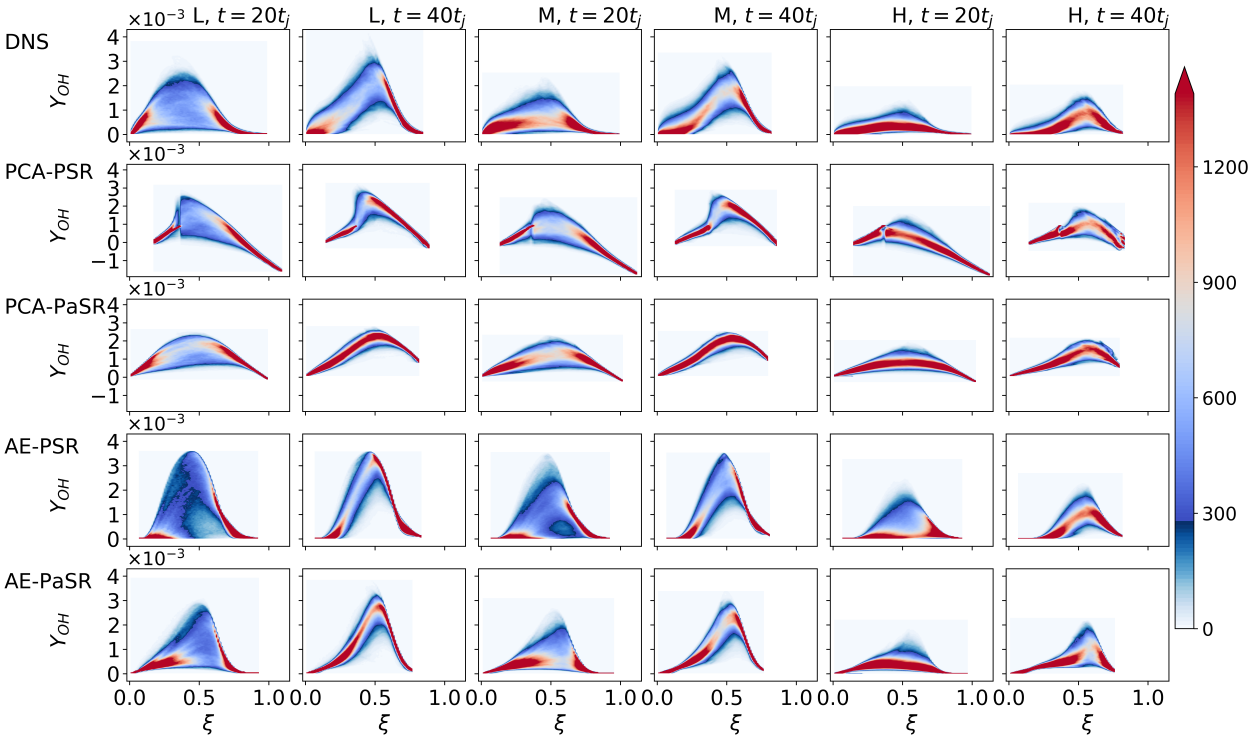

Fig. 8 Conditional PDF of OH mass fraction on mixture fraction ξ from DNS, PCA and AE models for cases L, M and H at times $t = [20t_j, 40t_j]$.

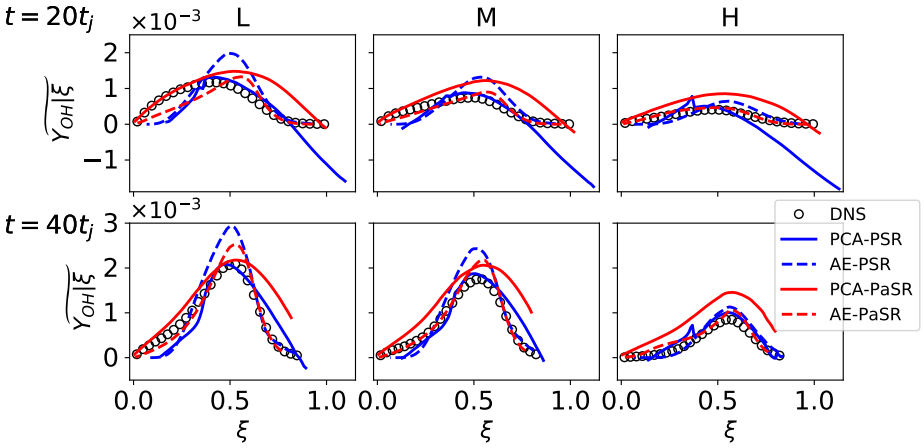

Fig. 9 Conditional mean of OH mass fraction on mixture fraction ξ from DNS, PCA and AE models for cases L, M and H at times $t = [20t_j, 40t_j]$.

from local extinction at $20t_j$ to re-ignition at $40t_j$ and the increasing amount of local extinction from case L, to M and H, with some deviations. The PCA-PSR model (blue solid lines) presents negative conditional mean values and the PCA-PaSR model (red solid lines) overpredicts the conditional mean values at almost all mixture fractions. The AE-PSR model (blue dashed lines) also overpredicts the conditional mean values around the stoichiometric conditions. The overprediction is alleviated in the AE-PaSR model (red dashed lines). Among the reduced models in Figs. 8 and 9, the AE-PaSR model shows the overall best agreement on conditional PDF distributions with DNS at all times for the three cases. The improvement in AE-PaSR could be attributed to the inclusion of mixing effects on the combustion state in PaSR, the significance of which has been widely acknowledged. Previous studies on PaSR [18, 28] have also

shown that the combustion states are sensitive to the mixing models. Currently the IEM mixing model [30, 31] is used. Explorations of other mixing models, such as the modified Curl model [35] and the Euclidean Minimum Spanning Tree (EMST) model [36], in constructing the PaSR datasets are expected to be able to improve the reduced models further.

Conclusions

Reduced models for chemical kinetics were derived from datasets composed of a large ensemble of 0-D PSR and PaSR reactors using AE neural network and PCA methods. The four reduced models, *i.e.*, PCA-PSR, PCA-PaSR, AE-PSR and AE-PaSR, were examined a priori in the 3-D temporally evolving jet flames featuring local extinction and re-ignition. It was found that for the syngas combustion in the DNS flames, the AE-PaSR model with two reduced variables can represent the thermo-chemical state satisfactorily. More specifically, it captures the local extinction and re-ignition behaviors and shows the best agreement with DNS data. It was also observed from the comparison among the reduced models and DNS results that the reduced models trained from PaSR datasets, *i.e.*, PCA-PaSR and AE-PaSR, show significant improvement over the models trained from PSR datasets, *i.e.*, PCA-PSR and AE-PSR. The improvement indicates the importance of considering the interactions of chemistry and turbulent mixing in dataset generation for reduced model construction. Interesting future work include considering other mixing models, 0-D stirred reactor network for training/test datasets and applying the reduced models in 3-D CFD simulations.

Acknowledgments

This work was supported by ORNL through the Laboratory Directed Research and Development (LDRD) Program. It was performed using resources of the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC0500OR22725. This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

[1] Peters, N., "Laminar diffusion flamelet models in non-premixed turbulent combustion," *Progress in Energy and Combustion Science*, Vol. 10, No. 3, 1984, pp. 319–339.

[2] Bilger, R., "Conditional moment closure for turbulent reacting flow," *Physics of Fluids A: Fluid Dynamics*, Vol. 5, No. 2, 1993, pp. 436–444.

[3] Van Oijen, J., and De Goey, L., "Modelling of premixed laminar flames using flamelet-generated manifolds," *Combustion Science and Technology*, Vol. 161, No. 1, 2000, pp. 113–137.

[4] Gicquel, O., Darabiha, N., and Thévenin, D., "Liminar premixed hydrogen/air counterflow flame simulations using flame prolongation of ILDM with differential diffusion," *Proceedings of the Combustion Institute*, Vol. 28, No. 2, 2000, pp. 1901–1908.

[5] Sutherland, J. C., and Parente, A., "Combustion modeling using principal component analysis," *Proceedings of the Combustion Institute*, Vol. 32, No. 1, 2009, pp. 1563–1570.

[6] Malik, M. R., Isaac, B. J., Coussement, A., Smith, P. J., and Parente, A., "Principal component analysis coupled with nonlinear regression for chemistry reduction," *Combustion and Flame*, Vol. 187, 2018, pp. 30–41.

[7] Isaac, B. J., Thornock, J. N., Sutherland, J., Smith, P. J., and Parente, A., "Advanced regression methods for combustion modelling using principal components," *Combustion and Flame*, Vol. 162, No. 6, 2015, pp. 2592 – 2601.

[8] Mirgolbabaei, H., and Echekki, T., "A novel principal component analysis-based acceleration scheme for LES–ODT: An a priori study," *Combustion and Flame*, Vol. 160, No. 5, 2013, pp. 898–908.

[9] Coussement, A., Gicquel, O., and Parente, A., "MG-local-PCA method for reduced order combustion modeling," *Proceedings of the Combustion Institute*, Vol. 34, No. 1, 2013, pp. 1117–1123.

[10] Isaac, B. J., Coussement, A., Gicquel, O., Smith, P. J., and Parente, A., "Reduced-order PCA models for chemical reacting flows," *Combustion and Flame*, Vol. 161, No. 11, 2014, pp. 2785–2800.

[11] Parente, A., Sutherland, J. C., Tognotti, L., and Smith, P. J., "Identification of low-dimensional manifolds in turbulent flames," *Proceedings of the Combustion Institute*, Vol. 32, No. 1, 2009, pp. 1579–1586.

[12] Mirgolbabaei, H., and Echekki, T., "Nonlinear reduction of combustion composition space with kernel principal component analysis," *Combustion and Flame*, Vol. 161, No. 1, 2014, pp. 118–126.

[13] Mirgolbabaei, H., Echekki, T., and Smaoui, N., "A nonlinear principal component analysis approach for turbulent combustion composition space," *International Journal of Hydrogen Energy*, Vol. 39, No. 9, 2014, pp. 4622–4633.

[14] Ranade, R., Alqahtani, S., Farooq, A., and Echekki, T., "An ANN based hybrid chemistry framework for complex fuels," *Fuel*, Vol. 241, 2019, pp. 625–636.

[15] Mirgolbabaei, H., and Echekki, T., "The reconstruction of thermo-chemical scalars in combustion from a reduced set of their principal components," *Combustion and Flame*, Vol. 162, No. 5, 2015, pp. 1650–1652.

[16] Zhang, P., Sankaran, R., Stoyanov, M., Lebrun-Grandie, D., and Finney, C. E., "Reduced Models for Chemical Kinetics derived from Parallel Ensemble Simulations of Stirred Reactors," *AIAA Scitech 2020 Forum*, 2020, p. 0177.

[17] Rahimi, M., Esposito, G., Chelliah, H., and Pope, S., "Chemical Kinetic Model Reduction Based on Partially-Stirred Reactor Simulations with Comparable Chemical and Mixing Time Scales," *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, American Institute of Aeronautics and Astronautics, Grapevine (Dallas/Ft. Worth Region), Texas, 2013.

[18] Ren, Z., and Pope, S. B., "An investigation of the performance of turbulent mixing models," *Combustion and Flame*, Vol. 136, No. 1-2, 2004, pp. 208–216.

[19] Zhang, P., and Sankaran, R., "Reduced models for chemical kinetics using autoencoder neural networks," under review.

[20] Hawkes, E. R., Sankaran, R., Sutherland, J. C., and Chen, J. H., "Scalar mixing in direct numerical simulations of temporally evolving plane jet flames with skeletal CO/H_2 kinetics," *Proceedings of the Combustion Institute*, Vol. 31, No. 1, 2007, pp. 1633–1640.

[21] Sakurada, M., and Yairi, T., "Anomaly detection using autoencoders with nonlinear dimensionality reduction," *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, 2014, pp. 4–11.

[22] Wehmeyer, C., and Noé, F., "Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics," *The Journal of Chemical Physics*, Vol. 148, No. 24, 2018, p. 241703.

[23] Hubert, M., and Engelen, S., "Robust PCA and classification in biosciences," *Bioinformatics*, Vol. 20, No. 11, 2004, pp. 1728–1736.

[24] Shan, R., and Lu, T., "Ignition and extinction in perfectly stirred reactors with detailed chemistry," *Combustion and Flame*, Vol. 159, No. 6, 2012, pp. 2069–2076.

[25] Correa, S. M., and Braaten, M. E., "Parallel simulations of partially stirred methane combustion," *Combustion and Flame*, Vol. 94, No. 4, 1993, pp. 469–486.

[26] Correa, S., "Turbulence-chemistry interactions in the intermediate regime of premixed combustion," *Combustion and Flame*, Vol. 93, No. 1-2, 1993, pp. 41–60.

[27] Bhave, A., and Kraft, M., "Partially Stirred Reactor Model: Analytical Solutions and Numerical Convergence Study of a PDF/Monte Carlo Method," *SIAM Journal on Scientific Computing*, Vol. 25, No. 5, 2004, pp. 1798–1823.

[28] Chen, J.-Y., "Stochastic Modeling of Partially Stirred Reactors," *Combustion Science and Technology*, Vol. 122, No. 1-6, 1997, pp. 63–94.

[29] Pope, S. B., "PDF methods for turbulent reactive flows," *Progress in energy and combustion science*, Vol. 11, No. 2, 1985, pp. 119–192.

[30] Villermaux, J., and Devillon, J., "Représentation de la coalescence et de la redispersion des domaines de ségrégation dans un fluide par un modéle d'interaction phénoménologique," *Proc. Second Int. Symp. on Chemical Reaction Engineering, Elsevier, New York*, 1972, pp. 1–13.

[31] Dopazo, C., and O'Brien, E. E., "An approach to the autoignition of a turbulent mixture," *Acta Astronaut.*, Vol. 1, No. 9-10, 1974, pp. 1239–1266.

[32] Mobini, K., and Bilger, R. W., "Imperfectly Stirred Reactor Model Predictions of Reaction in a Burner with Strong Recirculation," *Combustion Science and Technology*, Vol. 176, No. 1, 2004, pp. 45–70.

[33] Chollet, F., "Keras," `https://github.com/fchollet/keras`, 2015.

[34] Law, C. K., *Combustion physics*, Cambridge University Press, 2010.

[35] Janicka, J., Kolbe, W., and Kollmann, W., "Closure of the transport equation for the probability density function of turbulent scalar fields," *Journal of Non-Equilibrium Thermodynamics*, Vol. 4, No. 1, 1979, pp. 47–66.

[36] Subramaniam, S., and Pope, S. B., "A mixing model for turbulent reactive flows based on Euclidean minimum spanning trees," *Combustion and Flame*, Vol. 115, No. 4, 1998, pp. 487–514.