

# Learning physics-based reduced-order models for a single-injector combustion process

Renee Swischuk<sup>\*1</sup>, Boris Kramer<sup>†2</sup>, Cheng Huang<sup>‡3</sup>, and Karen Willcox<sup>§4</sup>

<sup>1</sup>*Massachusetts Institute of Technology, Cambridge, MA, 02139*

<sup>2</sup>*University of California, San Diego, CA, 92122*

<sup>3</sup>*University of Michigan, Ann Arbor, MI, 48109*

<sup>4</sup>*University of Texas at Austin, Austin, TX, 78712*

This paper presents a physics-based data-driven method to learn predictive reduced-order models (ROMs) from high-fidelity simulations, and illustrates it in the challenging context of a single-injector combustion process. The method combines the perspectives of model reduction and machine learning. Model reduction brings in the physics of the problem, constraining the ROM predictions to lie on a subspace defined by the governing equations. This is achieved by defining the ROM in proper orthogonal decomposition (POD) coordinates, which embed the rich physics information contained in solution snapshots of a high-fidelity computational fluid dynamics (CFD) model. The machine learning perspective brings the flexibility to use transformed physical variables to define the POD basis. This is in contrast to traditional model reduction approaches that are constrained to use the physical variables of the high-fidelity code. Combining the two perspectives, the approach identifies a set of transformed physical variables that expose quadratic structure in the combustion governing equations and learns a quadratic ROM from transformed snapshot data. This learning does not require access to the high-fidelity model implementation. Numerical experiments show that the ROM accurately predicts temperature, pressure, velocity, species concentrations, and the limit-cycle amplitude, with speedups of more than five orders of magnitude over high-fidelity models. Our ROM simulation is shown to be predictive 200% past the training interval. ROM-predicted pressure traces accurately match the phase of the pressure signal and yield good approximations of the limit-cycle amplitude.

## Nomenclature

<b>A</b>	= System matrix for linear part
<b>B</b>	= Input matrix
<b>H</b>	= Matricized quadratic tensor
<b>q(t)</b>	= State vector in finite dimensions
<b>u(t)</b>	= External input vector
<b>Q</b>	= Snapshot matrix
<b>V</b>	= Matrix of POD basis vectors
$x, y$	= Spatial coordinates
$d$	= Number of physical variables
$n_x$	= Spatial discretization dimension
$r$	= Reduced model dimension
$t$	= Time
$\vec{q}_p(t, x, y), \vec{q}_c(t, x, y), \vec{q}_L(t, x, y)$	= State vector in primitive, conservative, and learning variables
$p(t, x, y), \mathbf{p}(t)$	= Pressure, continuous and discretized
$T(t, x, y), \mathbf{T}(t)$	= Temperature, continuous and discretized
$\rho(t, x, y), \boldsymbol{\rho}(t)$	= Density, continuous and discretized

<sup>\*</sup>Graduate Student, MIT Center for Computational Engineering, swischuk@mit.edu, Student Member AIAA.

<sup>†</sup>Assistant Professor, Department of Mechanical and Aerospace Engineering, bmramer@ucsd.edu, Member AIAA

<sup>‡</sup>Assistant Research Scientist, Department of Aerospace Engineering, huangche@umich.edu, Member AIAA

<sup>§</sup>Director, Oden Institute for Computational Engineering and Sciences, kwillcox@oden.utexas.edu, AIAA Fellow.

$\xi(t, x, y), \xi(t)$	= Specific volume, continuous and discretized
$v_x(t, x, y), \mathbf{v}_x(t)$	= Velocity in $x$ direction, continuous and discretized
$Y_l(t, x, y)$	= Species mass fraction, $l = 1, 2, \dots, n_{\text{sp}}$
$c_l(t, x, y)$	= Species molar concentrations, $l = 1, 2, \dots, n_{\text{sp}}$ , also denoted as $[S]$
$\otimes$	= Kronecker product
$\hat{\cdot}$	= Notation for ROM quantities

#### Abbreviations

CFD	= Computational Fluid Dynamics
GEMS	= General equation and mesh solver; a CFD code
PDE	= Partial differential equation
POD	= Proper orthogonal decomposition
ROM	= Reduced-order model

## I. Introduction

This paper presents an approach to learning low-dimensional surrogate models for a complex, nonlinear, multi-physics, multi-scale dynamical system in form of a multi-species combustion process. The need for repeated model evaluations in optimization, design, uncertainty quantification and control of aerospace systems has driven the development of reduced-order models (ROMs) for applications in aerodynamics [1–7], reacting flows [8–10] and combustion [11–13]. ROMs combine the rich information embedded in high-fidelity simulations with the efficiency of low-dimensional surrogate models; yet, effective and robust ROM methods for nonlinear, multi-scale applications such as combustion have remained an open challenge.

Most existing nonlinear model reduction methods are intrusive—that is, they derive the ROM by projecting the high-fidelity model operators onto a low-dimensional subspace. In doing so, the physics of the problem is embedded in the reduced-order representation. The proper orthogonal decomposition (POD) [14, 15] is the most common way to define the low-dimensional subspace, using the singular value decomposition to identify low-dimensional structure based on training data. For some problems, the projection approach is amenable to rigorous error analysis and structure-preservation guarantees [16–19], but these rigorous guarantees do not apply to nonlinear, multi-physics, multi-scale models, for which projection-based ROMs remain challenging to implement (due to the need for access to the high-dimensional operators). The compressible flow setting of the combustion process poses numerous problems with respect to stability of the projection-based ROMs, see [20–25] for several approaches to address this stability problem. Furthermore, ROMs for these problems typically require relatively high dimensionality (and thus high cost) to avoid problems with robustness and stability [12, 26]. For instance, Huang et al. [11] construct two separate ROMs for the same single-injector combustion simulation as presented herein, one that uses POD and the other uses least-squares Petrov-Galerkin projection. That work finds that well over 100 modes are necessary to obtain stable ROMs and sufficient accuracy.

There is increasing attention to non-intrusive model reduction methods (sometimes called black-box or data-driven methods) that learn a model based on training data, without requiring explicit access to the high-fidelity model operators. The non-intrusive philosophy aligns directly with the field of machine learning, where representations such as neural networks have been shown to induce nonlinear model forms that can approximate many physical processes [27]. However, neural networks require a large amount of training data, limiting their utility when the data comes from expensive large-scale partial differential equation (PDE) simulations [28]. Moreover, the parametrization of the learned low-dimensional model is critical to the predictive accuracy and success of the learned model—in particular, it is critical to determining whether the ROM can issue reliable predictions in regimes outside of the training data.

For large-scale PDE models, an important class of non-intrusive learning approaches tackle this challenge of model parametrization by embedding the structure of the problem into the learning formulation. Some approaches use sparse learning techniques to identify PDE model terms that explain the data [29–31]. Dynamic mode decomposition [32, 33] extracts spectral information of the infinite dimensional linear Koopman operator from observed data of the nonlinear system. This spectral information can then be used to build data-driven predictive models. When the model can be expressed in the form of a dynamical system with polynomial terms, then the learning problem can be formulated as a parameter estimation problem, as in the operator inference approach of [34]. An important advantage of non-intrusive learning approaches is that the user has the flexibility to choose the variables that drive the learning. This opens the way for variable transformations that expose system structure and, in doing so, transform the ROM learning task into a structured form. In some cases, the governing PDEs naturally admit variable transformations that reveal polynomial

form, such as the specific volume representation of the Euler equations [35]. More generally, one can introduce new auxiliary variables to the problem—known as *lifting*—to produce a system that is polynomial in its expanded set of state variables [36–38]. This allows for a much broader class of nonlinear systems to be learned using the operator inference framework.

In this work, we build on the operator inference framework to learn structured, polynomial ROMs from simulated snapshot data of a single-injector combustor model. While in our case the polynomial model parametrization is a model approximation, we show that the predictive capabilities of the learned ROM are excellent beyond the training data. Our proposed approach follows the steps below, which we develop in detail in the ensuing sections:

- 1) We obtain high-dimensional simulation snapshot data for a spatially two-dimensional combustion process from the General Equation and Mesh Solver (GEMS) CFD code [39] developed at Purdue University. The governing equations and combustion problem setup are described in Section II.
- 2) We identify a set of state variables in which many of the terms in the governing equations have quadratic form. We transform the snapshot data to these new state variables, as described in Section III.C.
- 3) We use operator inference to learn a ROM that evolves the combustion dynamics in a low-dimensional subspace. Details of the model learning are given in Section III and Section IV.B.

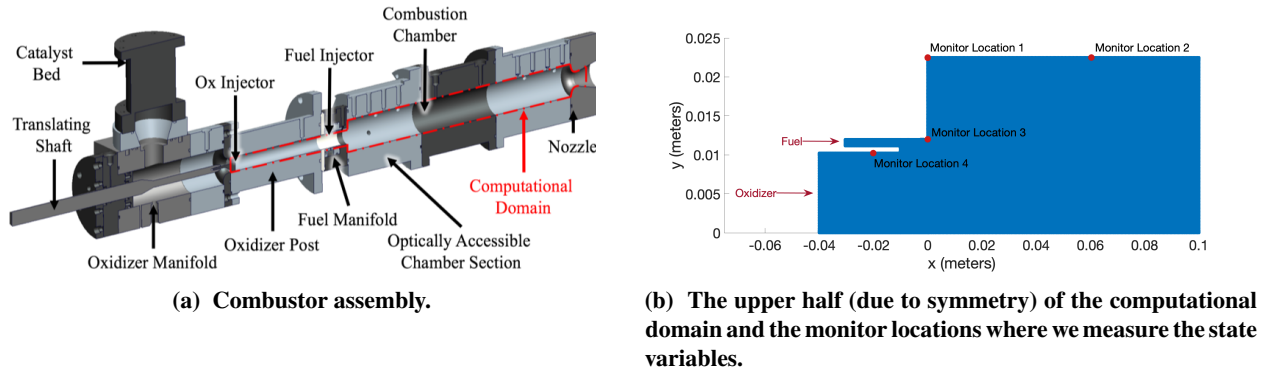
We present numerical results comparing our learned ROMs with GEMS test data in Section IV and conclude the paper in Section V.

## II. Combustion model

Section II.A defines the computational domain under consideration, Section II.B presents the governing equations for the combustion model, and Section II.C briefly summarizes the numerical implementation. The combustion model follows the implementation of the General Equation and Mesh Solver (GEMS) CFD code [39] and more details can be found in [40]. The GEMS code has been successfully used for rocket engine simulations [11] and in high-pressure gas turbines [41].

### A. Computational domain

A single-injector combustor as in [42] is shown in Figure 1a, with the computational domain outlined in red dashed lines. Our domain is a simplified two-dimensional version of the computational domain, shown in Figure 1b, which also shows the four locations where we monitor the state variables.



**Fig. 1** Setup and geometry of single-injector combustor.

### B. Governing equations

The dynamics of the combustor are governed by the conservation equations for mass, momentum, energy and species mass fractions. For this two-dimensional problem, the conservation equations are

$$\frac{\partial \vec{q}_c}{\partial t} + \nabla \cdot (\vec{K} - \vec{K}_v) = \vec{S} \quad (1)$$

and they describe the evolution of the conservative variables

$$\vec{q}_c = [\rho \quad \rho v_x \quad \rho v_y \quad \rho e \quad \rho Y_1 \quad \dots \quad \rho Y_{n_{sp}}]^\top,$$

where  $\rho$  is the density ( $\frac{\text{kg}}{\text{m}^3}$ ),  $v_x$  and  $v_y$  are the  $x$  and  $y$  velocity ( $\frac{\text{m}}{\text{s}}$ ),  $e$  is the total energy ( $\frac{\text{J}}{\text{m}^3}$ ), and  $Y_l$  is the  $l$ th species mass fraction with  $l = 1, 2, \dots, n_{sp}$  and  $n_{sp}$  is defined as the number of chemical species that are included in the model.

The total energy is defined as

$$e = \sum_{l=1}^{n_{sp}} h_l Y_l + \frac{1}{2} (v_x^2 + v_y^2) - \frac{p}{\rho} = h^0 - \frac{p}{\rho}, \quad (2)$$

where pressure  $p$  is given in (Pa),  $h_l = h_l(T)$  is the enthalpy corresponding to the  $l$ th species and is a highly nonlinear function of temperature,  $T$ , and  $h^0$  is the stagnation enthalpy. The inviscid flux  $\vec{K}$  and viscous flux  $\vec{K}_v$  in Eq. (1) are

$$\vec{K} = \begin{bmatrix} \rho v_x \\ \rho v_x^2 + p \\ \rho v_x v_y \\ \rho v_x e + p v_x \\ \rho v_x Y_1 \\ \vdots \\ \rho v_x Y_{n_{sp}} \end{bmatrix} \vec{i} + \begin{bmatrix} \rho v_y \\ \rho v_x v_y \\ \rho v_y^2 + p \\ \rho v_y e + p v_y \\ \rho v_y Y_1 \\ \vdots \\ \rho v_y Y_{n_{sp}} \end{bmatrix} \vec{j}, \quad \vec{K}_v = \begin{bmatrix} 0 \\ \tau_{xx} \\ \tau_{yx} \\ \tau_{xx} v_x + \tau_{yx} v_y - j_x^q \\ -j_{1,x}^m \\ \vdots \\ -j_{n_{sp},x}^m \end{bmatrix} \vec{i} + \begin{bmatrix} 0 \\ \tau_{xy} \\ \tau_{yy} \\ \tau_{xy} v_x + \tau_{yy} v_y - j_y^q \\ -j_{1,y}^m \\ \vdots \\ -j_{n_{sp},y}^m \end{bmatrix} \vec{j}.$$

The two-dimensional viscous shear tensor is defined as

$$\tau = \begin{bmatrix} \tau_{xx} & \tau_{xy} \\ \tau_{xy} & \tau_{yy} \end{bmatrix} = \hat{\mu} \begin{bmatrix} \frac{1}{3} \frac{\partial v_x}{\partial x} & \frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x} \\ \frac{\partial v_y}{\partial x} + \frac{\partial v_x}{\partial y} & \frac{1}{3} \frac{\partial v_y}{\partial y} \end{bmatrix},$$

where  $\hat{\mu}$  is the mixture viscosity coefficient. The diffusive heat flux vector is defined as

$$\vec{j}^q = \begin{bmatrix} j_x^q & j_y^q \end{bmatrix}^\top = -\kappa \nabla T + \rho \sum_{l=1}^{n_{sp}} D_l h_l \nabla Y_l, \quad (3)$$

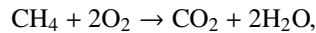
where  $\kappa$  quantifies thermal conductivity and  $D_l$  is the diffusion coefficient for the  $l$ th species into the mixture, which is an approximation used to model the multi-component diffusion as the binary diffusion of each species into a mixture. The two terms in the definition of the heat flux (Eq. (3)) represent heat transfer due to conductivity and species diffusion. The diffusive mass flux vector of species  $l$  is modeled as

$$\vec{j}_l^m = \begin{bmatrix} j_{l,x}^m & j_{l,y}^m \end{bmatrix}^\top = \left[ \rho D_l \frac{\partial Y_l}{\partial x} \quad \rho D_l \frac{\partial Y_l}{\partial y} \right]^\top.$$

The source term,  $\vec{S}$  in Eq. (1) is

$$\vec{S} = \begin{bmatrix} 0 & 0 & 0 & 0 & \dot{\omega}_1 & \dots & \dot{\omega}_{n_{sp}} \end{bmatrix}^\top \quad (4)$$

and is defined by considering a 1-step combustion reaction governed by



as presented in [43], with  $n_{sp} = 4$ . The corresponding general stoichiometric equation is defined as  $0 = \sum_{l=1}^{n_{sp}} \nu_l \chi_l$ , where  $\chi_1 = \text{CH}_4$ ,  $\chi_2 = \text{O}_2$ ,  $\chi_3 = \text{CO}_2$ ,  $\chi_4 = \text{H}_2\text{O}$  and  $\nu_l$  is the net stoichiometric coefficients of each species with  $\nu_1 = -1$ ,  $\nu_2 = -2$ ,  $\nu_3 = 1$  and  $\nu_4 = 2$ . The molar concentration of the  $l$ th species is denoted by  $c_l$ . In our case,  $l \in \{1, 2, 3, 4\}$ , so  $c_1 = [\text{CH}_4]$ ,  $c_2 = [\text{O}_2]$ ,  $c_3 = [\text{CO}_2]$ , and  $c_4 = [\text{H}_2\text{O}]$  are the molar concentrations. Here, we use the standard bracket notation  $[\cdot]$  to indicate molar concentration of a species. The general relationship between a species molar concentration,  $c_l$ , and a species mass fraction,  $Y_l$ , is

$$Y_l = \frac{c_l M_l}{\rho}, \quad (5)$$

where  $Y_1$  is the mass fraction of  $\text{CH}_4$ ,  $Y_2$  is the mass fraction of  $\text{O}_2$ ,  $Y_3$  is the mass fraction of  $\text{CO}_2$  and  $Y_4$  is the mass fraction of  $\text{H}_2\text{O}$ . The production rate of the  $l$ th species in the source term  $\vec{S}$  in Eq. 4 is modeled as

$$\dot{\omega}_l = M_l \frac{\partial c_l^{\text{reaction}}}{\partial t} = \nu_l \Gamma_r, \quad (6)$$

where  $c_l^{\text{reaction}}$  are chemical reaction source terms whose dynamics are described below and  $\Gamma_r$  is the reaction rate. The molar mass of  $\text{CH}_4$  is  $M_1 = 16.04 \frac{\text{g}}{\text{mol}}$ , the molar mass of  $\text{O}_2$  is  $M_2 = 32.0 \frac{\text{g}}{\text{mol}}$ , the molar mass of  $\text{CO}_2$  is  $M_3 = 44.01 \frac{\text{g}}{\text{mol}}$ , and the molar mass of  $\text{H}_2\text{O}$  is  $M_4 = 18.0 \frac{\text{g}}{\text{mol}}$ . The reaction rate is approximated by

$$\Gamma_r = k \prod_{l=1}^{n_{\text{reactant}}} c_l^{o_l},$$

where  $n_{\text{reactant}} = 2$  is the number of reactants,  $k$  is the rate coefficient and  $o_l$  is the reaction order of the  $l$ th reactant. In our case  $o_1 = 0.2$  and  $o_2 = 1.3$ . The rate coefficient,  $k$ , is described by the Arrhenius equation as

$$k = A \exp\left(\frac{-E_a}{R_u T}\right), \quad (7)$$

where  $R_u = 8.314 \frac{\text{J}}{\text{mol K}}$  is the universal gas constant,  $A = 2 \times 10^{10}$  is the pre-exponential constant and  $E_a = 2.025 \times 10^5$  is the energy required to reach a chemical reaction, measured in Joules and referred to as the activation energy. In this work, we use the ideal gas state equation that relates density and pressure to temperature

$$\rho = \frac{p}{RT}, \quad (8)$$

where  $R = \frac{R_u}{M}$  and  $M = \left(\sum_{l=1}^{n_{sp}} \left(\frac{Y_l}{M_l}\right)\right)^{-1}$  is the average molar mass of the mixtures. Thus, we can obtain temperature via  $T = \frac{p}{\rho R(Y_l)}$  from the states  $\rho, p, Y_l$ .

At the downstream end of the combustor, we impose a non-reflecting boundary condition while maintaining the chamber pressure via

$$p_{\text{back}}(t) = p_{\text{back,ref}}[1 + A \sin(2\pi f t)], \quad (9)$$

where  $p_{\text{back,ref}} = 1.0 \times 10^6$  Pa,  $A = 0.1$  and  $f = 5000\text{Hz}$ . The top and bottom wall boundary conditions are no-slip conditions, and for the upstream boundary we impose constant mass flow at the inlets.

### C. Numerical model

GEMS uses the finite volume method to discretize the conservation equations (1). The primitive variables  $\vec{q}_p = [p \ v_x \ v_y \ T \ Y_1 \ \dots \ Y_{n_{sp}}]^\top$  are chosen as solution variables in GEMS, since they allow for easier computation of thermal properties and provide more flexibility when extending to complex fluid problems like liquid and supercritical fluids. For a spatial discretization with  $n_x$  cells, this results in a  $dn_x$ -dimensional system of nonlinear ordinary differential equations (ODEs)

$$\frac{d\mathbf{q}}{dt} = \mathbf{G}(\mathbf{q}, \mathbf{u}(t)), \quad \mathbf{q}(0) = \mathbf{q}_0, \quad (10)$$

for  $0 < t \leq T$ , where  $d$  is the number of unknowns in the PDE governing equations and here  $d = 8$  (four flow variables and four species concentrations). In Eq. (10),  $\mathbf{q}(t) \in \mathbb{R}^{dn_x}$  is the discretized state vector at time  $t$  (for GEMS, it is the discretization of the primitive variables  $\vec{q}_p = [p \ v_x \ v_y \ T \ Y_1 \ \dots \ Y_{n_{sp}}]^\top$ ),  $\mathbf{q}_0$  are the specified initial conditions, and  $\frac{d\mathbf{q}}{dt}$  is the time derivative of the state vector at time  $t$ . The  $m$  inputs  $\mathbf{u}(t) \in \mathbb{R}^m$  arise from the time-dependent boundary condition, defined in Eq. (9), applied at the combustor downstream end. The nonlinear function  $\mathbf{G} : \mathbb{R}^{dn_x} \times \mathbb{R}^m \rightarrow \mathbb{R}^{dn_x}$  maps the discretized states  $\mathbf{q}$  and the input  $\mathbf{u}$  to the state time derivatives, representing the spatial discretization of the governing equations described in Section II.B.

Solving these high-dimensional nonlinear ODEs is expensive, motivating the derivation of a ROM that can yield approximate solutions at reduced cost. The nonlinear multi-scale dynamics represented by these equations makes this a challenging task. To maintain computational efficiency in the ROM, state-of-the-art nonlinear model reduction methods combine POD with a sparse interpolation method (often called hyperreduction) by evaluating the nonlinear functions only at a select number of points. For instance, POD together with the discrete empirical interpolation method (DEIM)

has had some success, but also encountered problems in combustion applications [12]. Of particular challenge is the need to include a large number of interpolation points in the POD-DEIM approximation, which means that the ROM loses its computational efficiency. Robustness and stability of the POD-DEIM models is also a challenge [26]. In the next section, we present a different approach that uses non-intrusive ROM learning to enable variable transformations that expose system structure. This structure is then exploited in the derivation of the ROM and removes the need for the DEIM approximation.

### III. Non-intrusive learning of a combustion reduced model

This section presents our approach to learn ROMs for the unsteady combustion dynamics simulation from GEMS. Section III.A writes a general nonlinear system in a form that exposes the underlying structure of the governing equations and shows how projection preserves that structure. Section III.B presents the operator inference approach from [34], which learns structured ROM operators from simulation data. Section III.C describes variable transformations that lead to the desired polynomial structure for the combustion governing equations presented in Section II. These transformations yield the structure needed to apply the operator inference approach.

#### A. Projection preserves polynomial structure in the governing equations

Consider a large-scale system of nonlinear ODEs written in polynomial form

$$\frac{d\mathbf{q}}{dt} = \mathbf{A}\mathbf{q} + \mathbf{H}(\mathbf{q} \otimes \mathbf{q}) + \mathbf{C}(\mathbf{q} \otimes \mathbf{q} \otimes \mathbf{q}) + \mathbf{B}\mathbf{u} + \mathbf{c} + \text{HOT}. \quad (11)$$

Relating this equation to the general nonlinear system in Eq. (10), we see that  $\mathbf{A}\mathbf{q}$  are the terms in  $\mathbf{G}(\cdot)$  that are linear in the state  $\mathbf{q}$ , with  $\mathbf{A} \in \mathbb{R}^{dn_x \times dn_x}$ ;  $\mathbf{H}(\mathbf{q} \otimes \mathbf{q})$  are the terms in  $\mathbf{G}(\cdot)$  that are quadratic in  $\mathbf{q}$ , with  $\mathbf{H} \in \mathbb{R}^{dn_x \times (dn_x)^2}$ ;  $\mathbf{C}(\mathbf{q} \otimes \mathbf{q} \otimes \mathbf{q})$  are the terms in  $\mathbf{G}(\cdot)$  that are cubic in  $\mathbf{q}$ , with  $\mathbf{C} \in \mathbb{R}^{dn_x \times (dn_x)^3}$ ;  $\mathbf{B}\mathbf{u}$  are the terms in  $\mathbf{G}(\cdot)$  that are linear in the input  $\mathbf{u}$ , with  $\mathbf{B} \in \mathbb{R}^{dn_x \times m}$ ; and  $\mathbf{c} \in \mathbb{R}^{dn_x}$  are constant terms in  $\mathbf{G}(\cdot)$  that do not depend on state or input. The abbreviation “HOT” in Eq. (11) denotes higher-order terms, and represents terms that are quartic and higher order, as well as any other nonlinear terms that cannot be represented in polynomial form.

We emphasize that we are not (yet) introducing approximations—rather, we are explicitly writing out the discretized equations in the form (11) to expose the system structure that arises from the form of the terms in the governing PDEs. For example, a term such as  $\frac{\partial}{\partial x} \rho v_x$  in Eq. (1) is linear in the state  $\rho v_x$ , while a term such as  $\frac{\partial}{\partial x} \rho v_x Y_1$  is quadratic in the states  $\rho v_x$  and  $Y_1$ . Also note that the term  $\frac{\partial}{\partial x} \rho v_x$  is quadratic in the states  $\rho$  and  $v_x$ , highlighting the important point that the structure of the nonlinear model depends on the particular choice of state variables.

A projection-based ROM of Eq. (11) preserves the polynomial structure. Approximating the high-dimensional state  $\mathbf{q}$  in a low-dimensional basis  $\mathbf{V} \in \mathbb{R}^{dn_x \times r}$ , with  $r \ll dn_x$ , we write  $\mathbf{q} \approx \mathbf{V}\hat{\mathbf{q}}$ . Using a Galerkin projection, this yields the ROM of Eq. (11) as

$$\frac{d\hat{\mathbf{q}}}{dt} = \hat{\mathbf{A}}\hat{\mathbf{q}} + \hat{\mathbf{H}}(\hat{\mathbf{q}} \otimes \hat{\mathbf{q}}) + \hat{\mathbf{C}}(\hat{\mathbf{q}} \otimes \hat{\mathbf{q}} \otimes \hat{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{u} + \hat{\mathbf{c}} + \text{HOT}, \quad (12)$$

where  $\hat{\mathbf{A}} = \mathbf{V}^T \mathbf{A} \mathbf{V} \in \mathbb{R}^{r \times r}$ ,  $\hat{\mathbf{H}} = \mathbf{V}^T \mathbf{H} (\mathbf{V} \otimes \mathbf{V}) \in \mathbb{R}^{r \times r^2}$ ,  $\hat{\mathbf{C}} = \mathbf{V}^T \mathbf{C} (\mathbf{V} \otimes \mathbf{V} \otimes \mathbf{V}) \in \mathbb{R}^{r \times r^3}$ , and  $\hat{\mathbf{B}} = \mathbf{V}^T \mathbf{B} \in \mathbb{R}^{r \times m}$  are the ROM operators corresponding respectively to  $\mathbf{A}$ ,  $\mathbf{H}$ ,  $\mathbf{C}$ , and  $\mathbf{B}$ , and  $\hat{\mathbf{c}} = \mathbf{V}^T \mathbf{c} \in \mathbb{R}^r$  is a constant vector. We note again that projection preserves polynomial structure, that is, (12) has the same polynomial form as (11), but in the reduced subspace defined by  $\mathbf{V}$ .

In what follows, we will work with a quadratic system in order to simplify notation. We note that the least squares learning approach described below applies directly to cubic, quartic and all higher-order polynomial terms (although it should be noted that the number of elements in the ROM operators scales with  $r^4$  for the cubic operator,  $r^5$  for the quartic operator, etc.). Higher-order terms often exhibit significant block-sparsity that can be exploited in numerical implementations, which limits the growth of computational cost to solve the ROM. For terms in the governing equations that are not in polynomial form (such as terms involving  $\frac{1}{p}$ , and the Arrhenius reaction terms) we discuss in Section III.C the introduction of variable transformations and auxiliary variables via the process of lifting [36, 37] to convert these terms to polynomial form.

## B. Operator inference for learning reduced models

Here we summarize the steps of the operator inference approach from [34]. First, we collect  $K$  snapshots of the state by solving the high-fidelity model. We store the snapshots and the inputs used to generate them in the matrices:

$$\mathbf{Q} = [\mathbf{q}_0 \dots \mathbf{q}_K] \in \mathbb{R}^{dn_x \times K}, \quad \mathbf{U} = [\mathbf{u}_0, \dots, \mathbf{u}_K] \in \mathbb{R}^{m \times K},$$

where  $\mathbf{u}_i \equiv \mathbf{u}(t_i)$  and  $\mathbf{q}_i \equiv \mathbf{q}(t_i)$  with  $0 = t_0 < t_1 < \dots < t_K = T$ . In general,  $dn_x \gg K$ , so the matrix  $\mathbf{Q}$  is tall and skinny. Second, we identify the low-dimensional subspace in which we will learn the ROM. In this work, we use the POD to define the low-dimensional subspace, by computing the singular value decomposition of the snapshot matrix

$$\mathbf{Q} = \mathbf{V}\mathbf{\Sigma}\mathbf{W}^\top,$$

where  $\mathbf{V} \in \mathbb{R}^{dn_x \times K}$ ,  $\mathbf{\Sigma} \in \mathbb{R}^{K \times K}$  and  $\mathbf{W} \in \mathbb{R}^{K \times K}$ . The  $r \ll dn_x$  dimensional POD basis,  $\mathbf{V}_r = [\mathbf{v}_1, \dots, \mathbf{v}_r]$ , is given by the first  $r$  columns of  $\mathbf{V}$ . Third, we project the state snapshot data onto the POD subspace spanned by the columns of  $\mathbf{V}_r$  and obtain the reduced snapshot matrices

$$\hat{\mathbf{Q}} = \mathbf{V}_r^\top \mathbf{Q} = [\hat{\mathbf{q}}_0 \dots \hat{\mathbf{q}}_K] \in \mathbb{R}^{r \times K}, \quad \dot{\hat{\mathbf{Q}}} = [\dot{\hat{\mathbf{q}}}_0 \quad \dot{\hat{\mathbf{q}}}_1 \quad \dots \quad \dot{\hat{\mathbf{q}}}_K] \in \mathbb{R}^{r \times K},$$

where the columns of  $\dot{\hat{\mathbf{Q}}}$  are computed from  $\hat{\mathbf{Q}}$  using any time derivative approximation (see, e.g., [44–46]), or can be obtained—if available—by collecting and projecting snapshots of  $\mathbf{G}(\mathbf{q}_i, \mathbf{u}_i)$ .

Operator inference solves a least squares problem to find the reduced operators that yield the ROM that best matches the projected snapshot data in a minimum residual sense. For the quadratic ROM

$$\dot{\hat{\mathbf{q}}} = \hat{\mathbf{A}}\hat{\mathbf{q}} + \hat{\mathbf{H}}(\hat{\mathbf{q}} \otimes \hat{\mathbf{q}}) + \hat{\mathbf{B}}\mathbf{u} + \hat{\mathbf{c}}, \quad (13)$$

operator inference solves the least squares problem

$$\min_{\hat{\mathbf{A}} \in \mathbb{R}^{r \times r}, \hat{\mathbf{H}} \in \mathbb{R}^{r \times r^2}, \hat{\mathbf{B}} \in \mathbb{R}^{r \times m}, \hat{\mathbf{c}} \in \mathbb{R}^r} \left\| \hat{\mathbf{Q}}^\top \hat{\mathbf{A}}^\top + (\hat{\mathbf{Q}} \otimes \hat{\mathbf{Q}})^\top \hat{\mathbf{H}}^\top + \mathbf{U}^\top \hat{\mathbf{B}}^\top + \mathbf{1}_K \hat{\mathbf{c}}^\top - \dot{\hat{\mathbf{Q}}}^\top \right\|_2^2,$$

where  $\mathbf{1}_K \in \mathbb{R}^K$  is the length  $K$  column vector with all entries set to unity. Note that this least squares problem is linear in the coefficients of the unknown ROM operators  $\hat{\mathbf{A}}$ ,  $\hat{\mathbf{H}}$ ,  $\hat{\mathbf{B}}$  and  $\hat{\mathbf{c}}$ . Also note that the operator inference approach permits us to compute the ROM operators  $\hat{\mathbf{A}}$ ,  $\hat{\mathbf{H}}$ ,  $\hat{\mathbf{B}}$  and  $\hat{\mathbf{c}}$  without needing explicit access to the original high-dimensional operators  $\mathbf{A}$ ,  $\mathbf{H}$ ,  $\mathbf{B}$  and  $\mathbf{c}$ .

We combine the unknown operators of Eq. (13) in the matrix

$$\mathbf{O} = [\hat{\mathbf{A}} \quad \hat{\mathbf{H}} \quad \hat{\mathbf{B}} \quad \hat{\mathbf{c}}] \in \mathbb{R}^{r \times (r+r^2+m+1)},$$

and the known low-dimensional data in the data matrix

$$\mathbf{D} = [\hat{\mathbf{Q}}^\top \quad (\hat{\mathbf{Q}} \otimes \hat{\mathbf{Q}})^\top \quad \mathbf{U}^\top \quad \mathbf{1}_K] \in \mathbb{R}^{K \times (r+r^2+m+1)}, \quad (14)$$

and then solve the minimization problem

$$\min_{\mathbf{O} \in \mathbb{R}^{r \times (r+r^2+m+1)}} \left\| \mathbf{D}\mathbf{O}^\top - \dot{\hat{\mathbf{Q}}}^\top \right\|_2^2. \quad (15)$$

For  $K > r+r^2+m+1$  this overdetermined linear least-squares problem has a unique solution [47, Sec. 5.3]. It was proven in [34] that Eq. (15) can be written as  $r$  independent least squares problems of the form  $\min_{\mathbf{o}_i \in \mathbb{R}^{r+r^2+m+1}} \|\mathbf{D}\mathbf{o}_i - \mathbf{r}_i\|_2^2$ ,

for  $i = 1, \dots, r$ , where  $\mathbf{o}_i$  is a column of  $\mathbf{O}^\top$  (row of  $\mathbf{O}$ ) and  $\mathbf{r}_i$  is a column of  $\dot{\hat{\mathbf{Q}}}^\top$ . This makes the operator inference approach efficient and scalable.

Regularization becomes necessary to avoid overfitting and to infer operators that produce a stable ROM. In this work, we use an  $L_2$  regularization penalty on the off-diagonal elements of the operator  $\hat{\mathbf{A}}$  and on all elements of the remaining operators. With this regularization, our least squares problem becomes

$$\min_{\mathbf{o}_i \in \mathbb{R}^{r+r^2+m+1}} \|\mathbf{D}\mathbf{o}_i - \mathbf{r}_i\|_2^2 + \lambda \|\mathbf{P}_i \mathbf{o}_i\|_2^2 \quad \text{for } i = 1, \dots, r, \quad (16)$$

where  $\lambda$  is the regularization parameter and  $\mathbf{P}_i$  is the  $r+r^2+m+1$  identity matrix with the  $i$ th diagonal set to zero so that we avoid regularizing the diagonal elements of  $\mathbf{A}$ . It should be noted that the regularization parameter,  $\lambda$ , is problem specific and should be chosen accordingly. In Section IV.B, we discuss details of the operator inference implementation, a method for selecting  $\lambda$ , and the removal of redundant terms in the least squares problem in Eq. (16).

### C. A structure-exploiting ROM learning formulation for GEMS

A key contribution of this work is to recognize that the non-intrusive operator inference approach gives us complete flexibility in the set of physical variables we work with to define the ROM. We can identify choices of physical variables that expose the desired polynomial structure in the governing equations, and then extract snapshots for those variables by applying transformations to the snapshot data—we do not need to make any modifications to the high-fidelity CFD simulation model itself. In theory, a classical intrusive ROM approach could work with transformed variables (e.g., in the work of [24]); however, this would involve rewriting the high-fidelity simulator, a task that would be not only time-consuming but also fraught with mathematical pitfalls, especially for unusual choices of variables. This is where the data-driven perspective of machine learning becomes extremely valuable.

The Euler equations admit a quadratic representation in the specific volume variables; in that case, a transformation of the snapshots from conservative (or primitive) variables to specific volume variables can be exploited to create quadratic ROMs [35]. Other PDEs may not admit polynomial structure via such straightforward transformations, but the process of *lifting* the equations via the introduction of new auxiliary variables can produce a set of coordinates in which the governing equations become polynomial in the lifted state [36–38]. For example, the tubular reactor example of [37] includes Arrhenius-type reaction terms similar to those in Eq. (7). The introduction of auxiliary variables permits the governing equations to be written equivalently with quartic nonlinearity in the lifted variables.\*

Lifting to polynomial form for the GEMS equations described in Section II is made difficult by several of the terms, in particular through some of the gas thermal properties such as the nonlinear dependence of enthalpy on temperature. A complete lifting that converts all equations to a polynomial form is possible, but would require the introduction of a large number of auxiliary variables and would also result in the introduction of some algebraic equations. However, as the analysis below shows, the GEMS governing equations admit a transformation for which many terms in the governing equations take polynomial form when we use the variables

$$\vec{q}_L = \begin{bmatrix} p & v_x & v_y & \xi & c_1 & c_2 & c_3 & c_4 \end{bmatrix}^\top. \quad (17)$$

Here  $\xi = \frac{1}{\rho}$  is the specific volume, and recall that  $c_1 = [\text{CH}_4]$ ,  $c_2 = [\text{O}_2]$ ,  $c_3 = [\text{CO}_2]$ , and  $c_4 = [\text{H}_2\text{O}]$  are the molar concentrations with  $c_l = \frac{\rho Y_l}{M_l}$ .

Below, we derive the governing PDEs for specific volume  $\xi$  and velocities  $v_x, v_y$ . These three governing PDEs all turn out to be quadratic in the learning variables  $\vec{q}_L$ . In Appendix A we present the lifting transformations for the source term dynamics  $c_l^{\text{reaction}}$  in the vector  $\vec{S}$  in Eq. (4). In Appendix B we derive the equations governing the pressure  $p$  and the species molar concentrations  $c_i$ . These equations have some terms that are not polynomial in the chosen learning variables  $\vec{q}_L$ .

To keep notation clean in application of the chain rule, let the conservative variables be denoted as  $g_1 = \rho$ ,  $g_2 = \rho v_x$ ,  $g_3 = \rho v_y$ ,  $g_4 = \rho e$ ,  $g_5 = \rho Y_1$ ,  $g_6 = \rho Y_2$ ,  $g_7 = \rho Y_3$ ,  $g_8 = \rho Y_4$ . Throughout, we frequently use the relationship

$$\frac{\partial \xi}{\partial x} = \frac{\partial}{\partial x} \frac{1}{\rho} = -\frac{1}{\rho^2} \frac{\partial \rho}{\partial x} = -\xi^2 \frac{\partial \rho}{\partial x}, \quad (18)$$

and similarly for  $\frac{\partial \xi}{\partial y}$ . Note also that we are assuming the existence of these partial derivatives, that is, we do not consider the case of problems with discontinuities.

**Specific Volume**  $\xi = 1/\rho$ . We use the constitutive relationship for the density  $\rho$  in Eq. (1) in the derivation:

$$\frac{\partial \xi}{\partial t} = \frac{\partial}{\partial t} \frac{1}{\rho} = -\frac{1}{\rho^2} \dot{\rho} = \xi^2 \nabla \cdot (\rho v_x \vec{i} + \rho v_y \vec{j}) = \xi^2 \left[ \frac{\partial \rho}{\partial x} v_x + \rho \frac{\partial v_x}{\partial x} \right] + \xi^2 \left[ v_y \frac{\partial \rho}{\partial y} + \rho \frac{\partial v_y}{\partial y} \right].$$

Inserting Eq. (18) into the above, we obtain

$$\frac{\partial \xi}{\partial t} = -\frac{\partial \xi}{\partial x} v_x + \xi \frac{\partial v_x}{\partial x} - v_y \frac{\partial \xi}{\partial y} + \xi \frac{\partial v_y}{\partial y},$$

which is quadratic in the learning variables  $\xi, v_x, v_y$ .

---

\*The Arrhenius reaction terms can be lifted further to quadratic form, but then require the inclusion of algebraic constraints, which makes the model reduction task more difficult, see [37].



**Velocities  $v_x, v_y$ :** We have

$$\frac{\partial v_x}{\partial t} = \frac{\partial}{\partial t} \frac{g_2}{g_1} = \frac{1}{g_1} \dot{g}_2 - g_2 \frac{1}{g_1^2} \dot{g}_1 = \frac{1}{\rho} \dot{g}_2 - \frac{v_x}{\rho} \dot{g}_1,$$

and from Eq. (1) we have that  $\dot{g}_1 = \dot{\rho} = -\nabla \cdot (\rho v_x \vec{i} + \rho v_y \vec{j})$  as well as  $\dot{g}_2 = (\rho \dot{v}_x) = \nabla \cdot (-(\rho v_x^2 + p) \vec{i} - (\rho v_x v_y) \vec{j} + \tau_{xx} \vec{i} + \tau_{xy} \vec{j})$ . Thus, we obtain

$$\begin{aligned} \frac{\partial v_x}{\partial t} &= \frac{1}{\rho} \nabla \cdot (-(\rho v_x^2 + p) \vec{i} - (\rho v_x v_y) \vec{j} + \tau_{xx} \vec{i} + \tau_{xy} \vec{j}) + \frac{v_x}{\rho} \nabla \cdot (\rho v_x \vec{i} + \rho v_y \vec{j}), \\ &= -\xi \frac{\partial \rho}{\partial x} v_x^2 - \frac{\partial v_x^2}{\partial x} - \xi \frac{\partial p}{\partial x} - \xi \frac{\partial \rho}{\partial y} v_x v_y - \frac{\partial v_x v_y}{\partial y} + \xi \left( \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} \right) + v_x^2 \xi \frac{\partial \rho}{\partial x} + v_x \frac{\partial v_x}{\partial x} + \xi v_x v_y \frac{\partial \rho}{\partial y} + v_x \frac{\partial v_y}{\partial y} \\ &= -\frac{\partial v_x^2}{\partial x} - \xi \frac{\partial p}{\partial x} - \frac{\partial v_x v_y}{\partial y} + \xi \left( \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} \right) + v_x \frac{\partial v_x}{\partial x} + v_x \frac{\partial v_y}{\partial y} \\ &= -\xi \frac{\partial p}{\partial x} - v_y \frac{\partial v_x}{\partial y} + \xi \left( \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} \right) - v_x \frac{\partial v_x}{\partial x} \end{aligned}$$

and we get a similar expression for  $\frac{\partial v_y}{\partial t}$ . Both dynamics are quadratic in the learning variables  $p, v_x, v_y, \xi$ .

As noted above, Appendix A and Appendix B present the derivations for the chemical source terms, pressure, and chemical species.

## IV. Numerical Results

We now apply the variable transformations and operator inference framework to learn a predictive ROM from GEMS high-fidelity combustion simulation data.<sup>†</sup> Section IV.A describes the problem setup and GEMS dataset. Section IV.B discusses implementation details and Section IV.C presents our numerical results. Additional numerical results can be found in [48].

### A. GEMS Dataset

The computational domain shown in Figure 1b is discretized with  $n_x = 38523$  spatial discretization points. Each CFD state solution thus has dimension  $dn_x = 308184$ . The problem considered here has fuel and oxidizer input streams with constant mass flow rates of  $5.0 \frac{\text{kg}}{\text{s}}$  and  $0.37 \frac{\text{kg}}{\text{s}}$ , respectively. The fuel is composed of gaseous methane and the oxidizer is 42% gaseous  $\text{O}_2$  and 58% gaseous  $\text{H}_2\text{O}$ , as described in [11]. The forcing input Eq. (9) is applied at the right side of the domain. For this simulation, the resulting Reynolds number is about 10,000, defined as  $Re = \frac{\rho v_x L}{\mu}$  where the density  $\rho$ , horizontal velocity  $v_x$  and viscosity  $\mu$  are evaluated at the inlet of the oxidizer post ( $x = -0.04\text{m}$  in Fig. 1b), and the characteristic length  $L$  is defined as the height of the oxidizer inlet. The highest Mach number is  $\approx 0.25$  and is evaluated inside the oxidizer post (from  $x = -0.04\text{m}$  to  $0\text{m}$  in Fig. 1b).

To generate training data, GEMS is simulated for a time duration of 1ms with a time step size of  $\Delta t = 1 \times 10^{-7}\text{s}$  resulting in  $K = 10000$  snapshots. The GEMS output is transformed to the variables given in Eq. (17). The recorded snapshot matrix is thus

$$\mathbf{Q} = [\mathbf{q}_0 \quad \mathbf{q}_1 \quad \dots \quad \mathbf{q}_K] \in \mathbb{R}^{dn_x \times K} = \mathbb{R}^{308184 \times 10000}$$

Our numerical experiments were parallelized on a cluster with two computing nodes. Each node has two 10-core Intel Xeon-E5 processors (20 cores per node) and 128 GB RAM. The training data generation took approximately 200h in CPU time for the 1ms, 10000 snapshots of high-fidelity CFD data.

The range of variable values for the training data is shown in Table 1. Note that the data covers a wide range of scales. Pressure is of the order  $10^6$  while species concentrations can be as low as  $10^{-12}$ . This large scaling difference presents a challenge when learning models from data. To deal with the numerical issues related to large differences in scaling and small species concentrations and velocities, we scale each variable to the interval  $[-1, 1]$ . Variables are scaled before computing the POD basis and projecting the data.

<sup>†</sup>Code for the operator inference framework is available at <https://test.pypi.org/project/operator-inference/> in Python and <https://github.com/elizqian/operator-inference> in Matlab.

**Table 1 Range of variable values for GEMS data.**

State variable	Minimum	Mean	Maximum
Pressure $p$ in Pa	$9.226 \times 10^5$	$1.142 \times 10^6$	$1.433 \times 10^6$
Velocity $v_x$ in $\frac{m}{s}$	-222.930	69.637	307.147
Velocity $v_y$ in $\frac{m}{s}$	-206.990	1.304	186.548
Specific volume $\xi = \rho^{-1}$ in $\frac{m^3}{kg}$	0.0533	0.220	0.674
Molar concentration $[CH_4]$	0.0	0.063	1.169
Molar concentration $[O_2]$	0.0	0.056	0.097
Molar concentration $[CO_2]$	0.0	0.002	0.012
Molar concentration $[H_2O]$	0.0	0.154	0.232

To obtain snapshots of the projected state time derivative, we approximate the derivative with a five-point approximation  $\hat{\mathbf{q}}_i = (-\hat{\mathbf{q}}_{i+2} + 8\hat{\mathbf{q}}_{i+1} - 8\hat{\mathbf{q}}_{i-1} + \hat{\mathbf{q}}_{i-2})/(12\Delta t)$ . This approximation is fourth order accurate. The first two and last two time derivatives are computed using first-order forward and backward Euler approximations, respectively.

## B. Learning a quadratic reduced-order model

To learn the operators of the quadratic ROM, we solve the regularized least squares problem shown in Eq. (16). We are using numpy's least squares solve `numpy.linalg.lstsq`. The algorithm is based on the LAPACK routine `xGELSD`. That routine is based on the SVD, which typically provides a stable implementation. In what follows next, we describe several important implementation details.

### 1. Singular value decomposition implementation and POD basis selection

Due to the large size of this dataset, we implement the randomized singular value decomposition algorithm, introduced in [49], to compute the leading 500 singular values and vectors of the snapshot matrix. The randomized singular value decomposition algorithm can be implemented in a scalable way for large datasets as the data does not have to be read into single memory all at once. The POD basis is chosen as the  $r$  leading left singular vectors. The dimension  $r$  is typically chosen so that the cumulative energy contained in the subspace is greater than a user specified tolerance  $\epsilon$ , i.e.,

$$\frac{\sum_{k=1}^r \sigma_k^2}{\sum_{k=1}^{dn_x} \sigma_k^2} > \epsilon,$$

where  $\sigma_k^2$  are the squared singular values of the data matrix  $\mathbf{Q}$ . To guide the choice of  $r$  we also use the relative projection error

$$\mathcal{E}_{\text{proj}} = \frac{\|\mathbf{Q} - \mathbf{V}_r \mathbf{V}_r^\top \mathbf{Q}\|_F^2}{\|\mathbf{Q}\|_F^2} = 1 - \frac{\sum_{k=1}^r \sigma_k^2}{\sum_{k=1}^{dn_x} \sigma_k^2} \quad (19)$$

### 2. Removing redundant terms in least squares problem

There are redundant terms that arise in the Kronecker product  $\hat{\mathbf{Q}} \otimes \hat{\mathbf{Q}}$  in Eq. (14), which can cause the least squares problem to become ill-posed. To see this, consider  $q \otimes q$  with  $q = [q_1 \ q_2]^T$ , i.e., we have  $q \otimes q = [q_1^2 \ q_2 q_1 \ q_1 q_2 \ q_2^2]$ . Since we know where the repeated terms occur in the product we merely need to remove the redundant (repeated) terms before we solve the least-squares problem. Thus, the Kronecker product is replaced with the term

$$\hat{\mathbf{Q}}^2 = [\hat{\mathbf{q}}_0^2 \ \hat{\mathbf{q}}_1^2 \ \dots \ \hat{\mathbf{q}}_K^2] \in \mathbb{R}^{s \times K},$$

where  $s = \frac{r(r+1)}{2}$ . Each vector  $\hat{\mathbf{q}}_j^2$  is defined, according to [34], as

$$\hat{\mathbf{q}}_j^2 = \begin{bmatrix} \mathbf{q}_j^{(1)} \\ \vdots \\ \mathbf{q}_j^{(r)} \end{bmatrix} \in \mathbb{R}^s, \quad \text{where,} \quad \mathbf{q}_j^{(i)} = \hat{q}_{i,j} \begin{bmatrix} \hat{q}_{i,j} \\ \vdots \\ \hat{q}_{r,j} \end{bmatrix} \in \mathbb{R}^i,$$

and  $\hat{q}_{i,j}$  is the  $i$ th element of the vector  $\hat{\mathbf{q}}_j$ . Now, instead of learning the operator  $\hat{\mathbf{H}} \in \mathbb{R}^{r \times r^2}$ , the operator  $\hat{\mathbf{F}} \in \mathbb{R}^{r \times s}$  is learned, which satisfies the equivalent least squares problem

$$\min_{\hat{\mathbf{A}} \in \mathbb{R}^{r \times r}, \hat{\mathbf{F}} \in \mathbb{R}^{r \times s}, \hat{\mathbf{B}} \in \mathbb{R}^{r \times m}, \hat{\mathbf{c}} \in \mathbb{R}^r} \left\| \hat{\mathbf{Q}}^\top \hat{\mathbf{A}}^\top + (\hat{\mathbf{Q}}^2)^\top \hat{\mathbf{F}}^\top + \mathbf{U}^\top \hat{\mathbf{B}}^\top + \mathbf{1}_K \hat{\mathbf{c}}^\top - \dot{\hat{\mathbf{Q}}}^\top \right\|_2^2 \quad (20)$$

The least squares problem is again of the form (15) but we use the data matrix  $\mathbf{D} = \begin{bmatrix} \hat{\mathbf{Q}}^\top & (\hat{\mathbf{Q}}^2)^\top & \mathbf{U}^\top & \mathbf{1}_K \end{bmatrix} \in \mathbb{R}^{K \times (r+s+m+1)}$  and solve for the operators  $\mathbf{O} = [\hat{\mathbf{A}} \quad \hat{\mathbf{F}} \quad \hat{\mathbf{B}} \quad \hat{\mathbf{c}}] \in \mathbb{R}^{r \times (r+s+m+1)}$ . Once we have solved for the operator  $\hat{\mathbf{F}}$  we can easily transform it to obtain  $\hat{\mathbf{H}}$ .

### 3. Regularization

We use an  $L_2$ -regularization (also known as Tikhonov regularization or ridge regression) to solve the operator inference problem, as shown in Eq. (20). The regularization term introduces a trade-off between operators that fit the data well and operators with small values. This regularization is used to avoid overfitting to the data, which in this setting causes our learned ROMs to be unstable (solution blow up in finite time). The regularization parameter  $\lambda$  affects the performance of this algorithm—we require enough regularization to avoid overfitting, but if  $\lambda$  is too large, the data will be poorly fit.

To help determine appropriate values of  $\lambda$ , we consider the “L-curve” criterion discussed in [50]. The L-curve is a way of visualizing the effects of different values of  $\lambda$  on the norm of the residual (data fit) against the norm of the solution. The L-curve criterion recommends choosing a value for  $\lambda$  that lies in the corner of the curve, nearest the origin. In our numerical experiments, we compute the L-curve to help determine appropriate values for  $\lambda$ .

Regularization also helps to reduce the condition number of the regularized least squares data matrix,  $[\mathbf{D} \quad \lambda \mathbf{P}]^\top$ . In Figure 2 we show the condition number of the data matrix  $\mathbf{D}$  from Eq. (20) (with the data already scaled to  $[-1, 1]$ ) when we include 2500, 5000, 7500 and 10000 snapshots in the training set. The condition number is quite large for this application, yet it decreases as we add more training data. This effect is due to the fact that as we add more training data, the first  $r$  POD basis vectors become richer. This in turn means that the projected data in  $\mathbf{D}$  are richer for a given dimension of the POD basis. Our numerical experiments reinforced this finding by confirming that 10000 training snapshots were required to achieve a sufficiently rich POD basis to obtain accurate ROMs, an indication of the complexity of the combustor dynamics. While the condition number of the least-squares problem is high, for this example, it remains at manageable levels. However, if the condition number gets much larger, a careful treatment and consideration of proper data sampling and regularization will be required.

### 4. Error measures

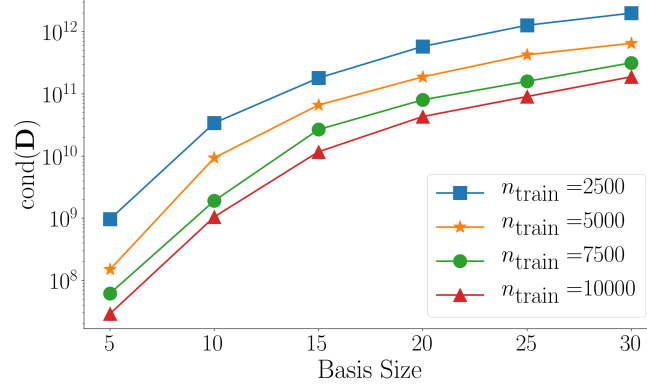
To evaluate the ROM performance, we define appropriate error metrics for each variable. The error is computed after the ROM solutions have been reconstructed back into the full CFD model dimension and scaled back to the original variable ranges. Recall Table 1, which showed the range of values for each variable. Below we provide details on how error is computed for each variable:

- For pressure and temperature, the values are always positive and well above zero, so we use a standard relative error, defined as

$$\mathcal{E}_{\text{relative}, i} = \frac{|\zeta_{\text{CFD}, i} - \zeta_{\text{ROM}, i}|}{|\zeta_{\text{CFD}, i}|} \quad (21)$$

where  $\zeta \in \{\mathbf{p}, \mathbf{T}\}$  and  $\zeta_{\cdot, i}$  denotes the  $i$ th entry of a vector, and  $\zeta_{\text{CFD}}$  is the CFD solution variable and  $\zeta_{\text{ROM}}$  is the solution variable obtained from the ROM simulation.

- Due to the small values of species concentrations (on the order of  $10^{-12}$ ), dividing by the true value can skew a small error. Similarly, velocities range from positive to negative, including zero. Thus, for species concentrations



**Fig. 2** The condition number of the data matrix,  $D$ , vs. basis size for different sized training sets.

and velocities, we use a normalized absolute error, defined as

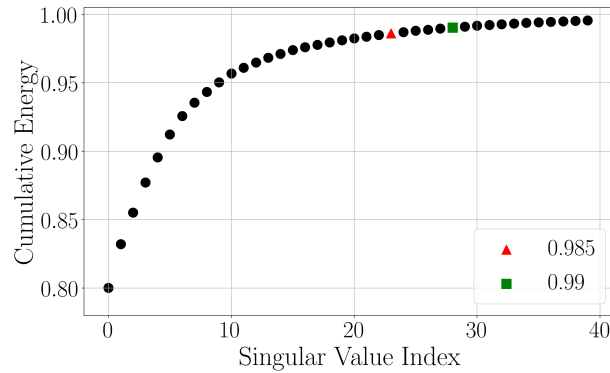
$$\mathcal{E}_{\text{nabs},i} = \frac{|\zeta_{\text{CFD},i} - \zeta_{\text{ROM},i}|}{\max_l (|\zeta_{\text{CFD},l}|)}, \quad (22)$$

where  $\zeta \in \{\mathbf{v}_x, \mathbf{v}_y, [\text{CH}_4], [\text{O}_2], [\text{CO}_2], [\text{H}_2\text{O}]\}$  and  $\max_l (|\zeta_{\text{CFD},l}|)$  denotes the maximum entry of  $|\zeta_{\text{CFD}}|$ , i.e., the maximum absolute value over the discretized spatial domain.

### C. Learned reduced model performance

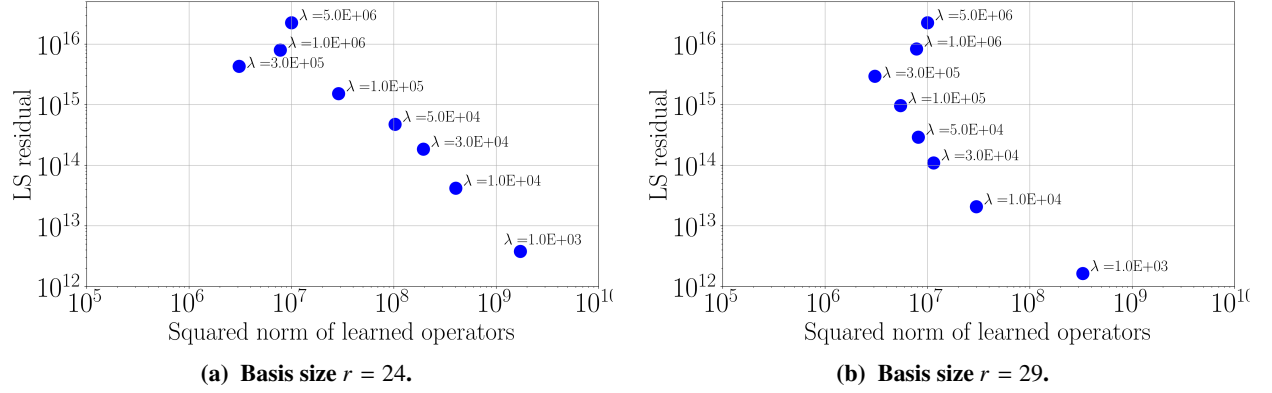
The given  $K = 10000$  snapshots representing 1ms of GEMS simulation data (scaled to  $[-1,1]$ ) are used to learn the ROM. We are also given another 2ms of testing data at the monitor locations shown in Figure 1b, which we use to assess the predictive capabilities of our learned ROMs beyond the range of training data.

The cumulative energy of the singular values of the snapshot data matrix  $\mathbf{Q}$  is shown in Figure 3. The singular values that correspond to a cumulative energy of 98.5% and 99% are indicated with the red triangle and green square, respectively. We use basis sizes of  $r = 24$ , capturing 98.5% of the total energy, and  $r = 29$ , capturing 99% of the total energy.



**Fig. 3** The cumulative energy computed with Eq. (19). The leading  $r = 24$  singular values capture 98.5% of the energy and the leading  $r = 29$  capture 99%.

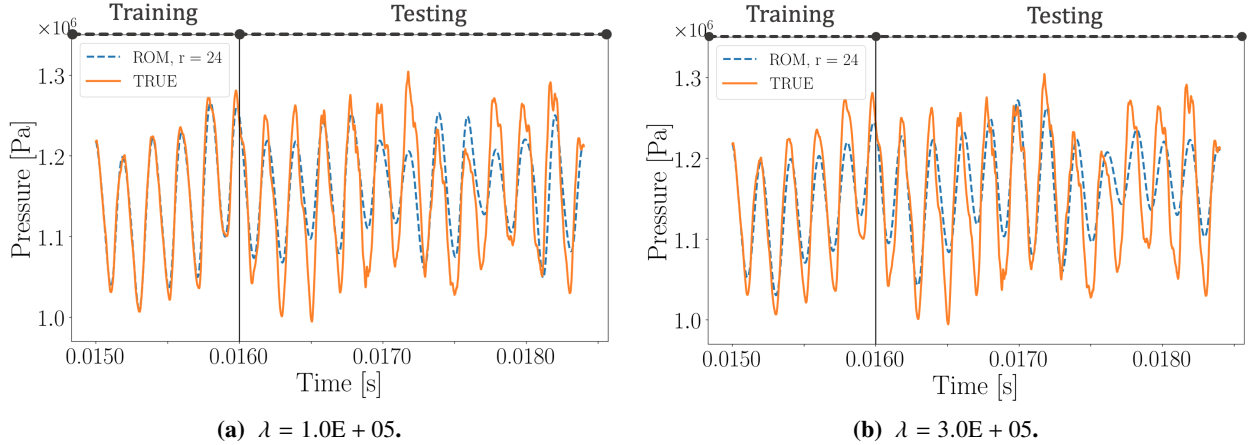
In Figures 4a, 4b we show the L-curve for each basis size. The L-curve for a basis of  $r = 24$  is somewhat uninformative in this case. The regularization parameters chosen were  $1.0\text{E}+05$  and  $\lambda = 3.0\text{E}+05$ , which lie in the upper part of the L-curve and are therefore a recommended choice by the L-curve criterion discussed above. For a basis



**Fig. 4** The L-curve for different basis sizes and regularization parameters  $\lambda$ . The horizontal axis shows the squared norm of learned operators,  $\|\mathbf{O}\|_2^2$ , and the vertical axis shows the least squares residual,  $\|\mathbf{D}\mathbf{O}^\top - \hat{\mathbf{Q}}^\top\|_2^2$ .

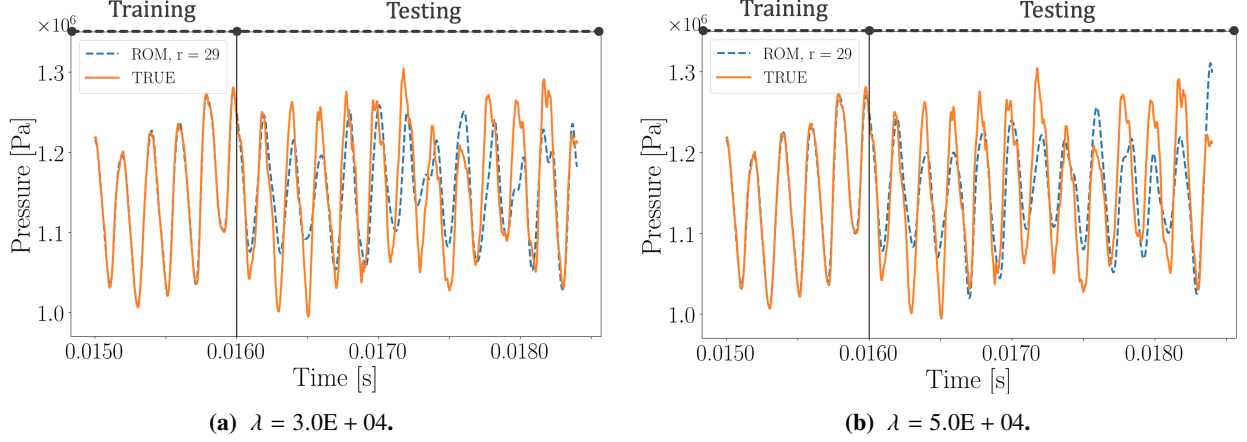
of size of  $r = 29$ , the L-curve indicates a regularization parameter around  $\lambda = 3.0\text{E}+04$ . Stable systems are produced for  $\lambda = 3.0\text{E}+04$  and  $5.0\text{E}+04$ .

We simulate the learned ROM for the two model sizes  $r = 24$  and  $r = 29$  with the same initial value and time step size ( $\Delta t = 1 \times 10^{-7}\text{s}$ ) as those used to generate the training set. Since the ROM was constructed from data scaled to  $[-1, 1]$ , we solve the ROM in the (scaled) subspace, and then reconstruct the dimensional quantities by reversing the scaling. Figures 5 and 6 compare the time trace of pressure computed by GEMS (our “truth” data) with the ROM predictions for 30000 time steps (10000 time steps used for training, 20000 time steps are pure prediction for the ROM) at the cell located at  $(0.0, 0.0225)$  in the domain (denoted as monitor location 1 in Figure 1b). The performance of the ROM on the training data (first 1ms of data) is good in both cases, although the  $r = 29$  ROM (Figure 6) is more accurate. For test data predictions beyond the training data, both ROMs yield accurate phase predictions and pressure oscillation amplitudes that are good approximations of the truth.



**Fig. 5** Pressure time traces for basis size  $r = 24$ . Training with 10000 snapshots. Black vertical line denotes the end of the training data and the beginning of the test data.

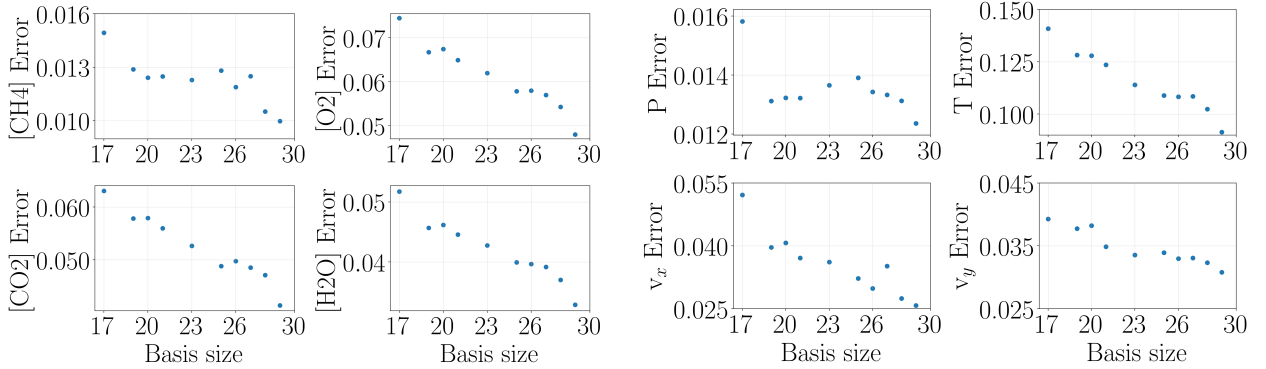
A Galerkin-projection-based POD method was applied to this GEMS model in Ref. [11]. The authors there found that a large number of modes ( $r > 100$ ) was necessary to obtain stable ROMs. Moreover, comparable accuracy, e.g., for the pressure probe predictions shown here, was only achieved with  $r = 200$  POD modes for the ROM (c.f. Fig. 15 in Ref. [11] with Figures 5–6 herein; both are recording pressure at the same probe location). Consequently, while our learned ROM does not resolve the full flow physics, we do obtain good predictability in time at much lower ROM dimension than in the classical Galerkin-POD approach used in Ref. [11]. The key innovation leading to this



**Fig. 6** Pressure time traces for basis size of  $r = 29$ . Training with 10000 snapshots. Black vertical line denotes the end of the training data and the beginning of the test data.

improvement is our use of variable transformations to build the ROM over a space for which the transformed governing equations have more polynomial structure. The non-intrusive operator inference approach is an enabler that makes these variable transformations practical from an implementation perspective, since the transformations are applied only to the snapshot data and not to the CFD model itself.

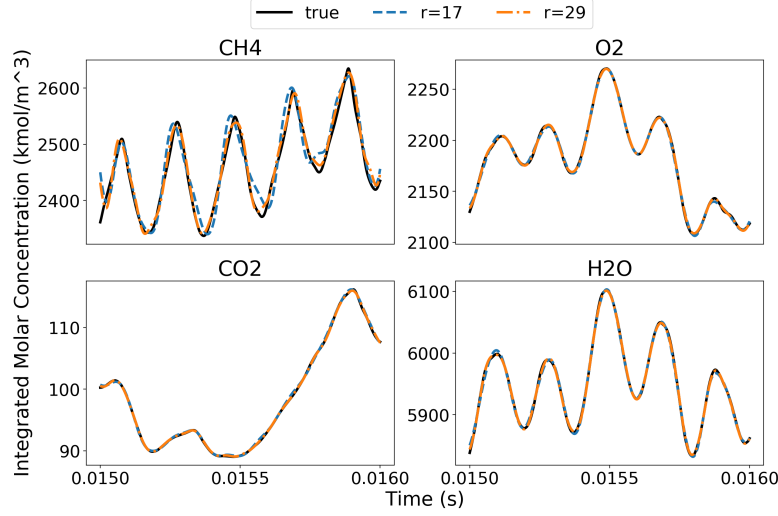
We also compute the average error of each variable over the entire domain at the last time step of the training set (the 10000th time step). The normalized absolute error, defined in Eq. (22), is shown for each species and for  $x$  and  $y$  velocity in Figure 7. This figure also shows the relative error, defined in Eq. (21), for pressure and temperature. These plots show that overall, the error is decreasing with an increasing basis size. At a basis size of 18, 22, and 24 the systems were unstable (solution blow up in finite time) and so these basis sizes are excluded from the figure. The cause of this, and the non-monotone error decay, may be due to the fact that the same regularization parameter was used for each of these,  $\lambda = 3.0E+04$ . Ideally, one would pick a parameter specific for the basis size, but here we used the same regularization parameter in order to give a fairer assessment of the ROM performance without using manual tuning to optimize our results. We note, however, that even with manual tuning of the regularization parameters we are not guaranteed monotone state-error decay for strongly nonlinear dynamical system ROMs.



**Fig. 7** Error measures vs. basis size, averaged over the spatial domain at the last time step of training data. Normalized absolute error (22) of species  $\text{CH}_4$ ,  $\text{O}_2$ ,  $\text{CO}_2$ ,  $\text{H}_2\text{O}$  and  $v_x$  and  $v_y$  velocities; Relative error (21) given for pressure and temperature.

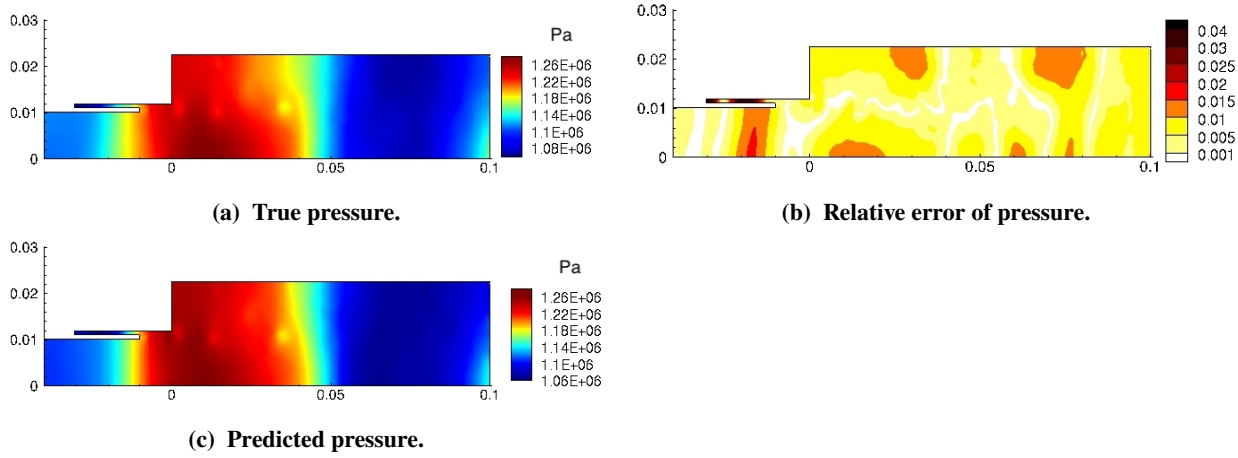
In Figure 8, we show the integrated species concentrations over time. To compute these, at each time step in our simulation, we sum all elements of a species vector. This measure monitors whether our ROM conserves species mass, an important feature of a physically meaningful simulation. As the discretization of the high-fidelity model becomes finer, point-wise error may become large and misleading if the mass is shifted slightly into the neighboring cells. The

integrated species concentration complements the evaluation of point-wise errors and provides a global view of the error in the domain. While  $\text{CH}_4$  conservation is tracked well qualitatively by the ROM, it does show the largest deviation out of the four species. This is a result of  $\text{CH}_4$  having the sharpest gradients ( $\text{CH}_4$  concentration ranging from 0 to 1) compared to the other three species, see also Figures 13–16 below, and their respective color bars.



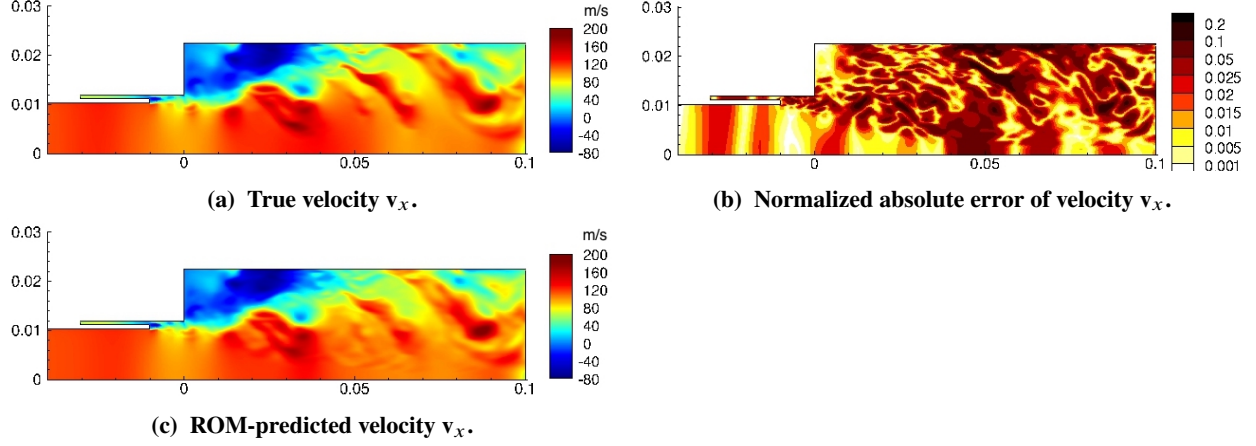
**Fig. 8** Integrated species at each time step for different basis sizes. Training with 10000 snapshots.

We also compare the state variables over the entire domain predicted by the learned ROM with the given GEMS data at the last time step  $K = 10000$  (which corresponds to  $t = 0.0159999\text{s}$ ). We provide the true field, the ROM-predicted field, and an error field for each variable in Figures 9–16. Again, for pressure and temperature, we use a relative error from Eq. (21). For  $x$  and  $y$  velocity and for species molar concentrations we use a normalized absolute error from Eq. (22). The plots show that the ROM predictions are, as expected, not perfect, but indeed they capture well the overall structure and many details of the solution fields.

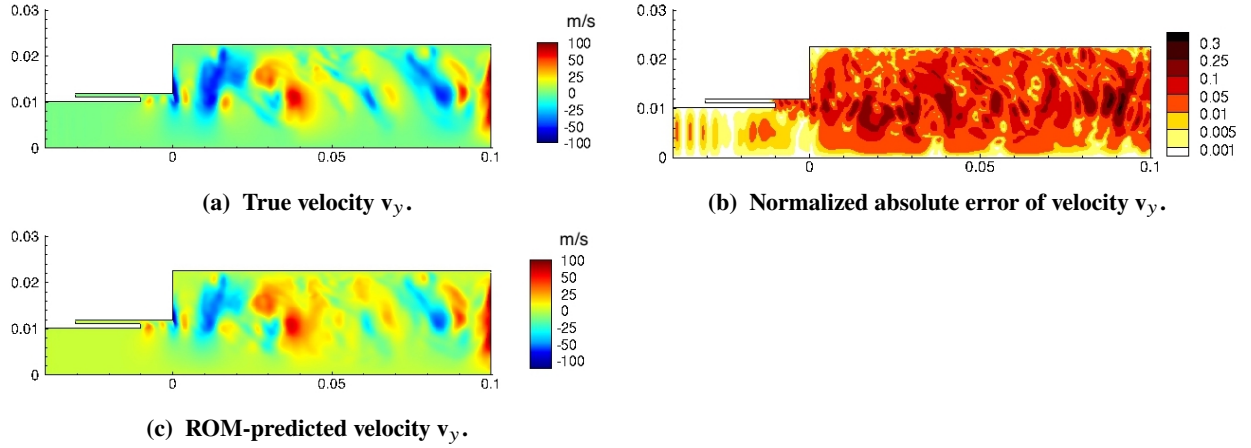


**Fig. 9** Predictive results for pressure at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0\text{E} + 04$ .

Table 2 shows timing results for the ROM generation and simulation, as performed using python 3.6.4 on a dual-core Intel i5 processor with 2.3 GHz and 16GB RAM. We report the following CPU times: solving the operator inference least squares problem (20); ROM runtime for the two different basis sizes for 3ms of real time prediction; and reconstruction of the high-dimensional, unscaled combustion variables. The latter is required since the ROM is evolving the dynamics in the  $[-1, 1]$  scaled variables, and thus a post-processing step is required to obtain the true magnitudes of the variables.



**Fig. 10** Predictive results for velocity  $v_x$  at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0E + 04$ .



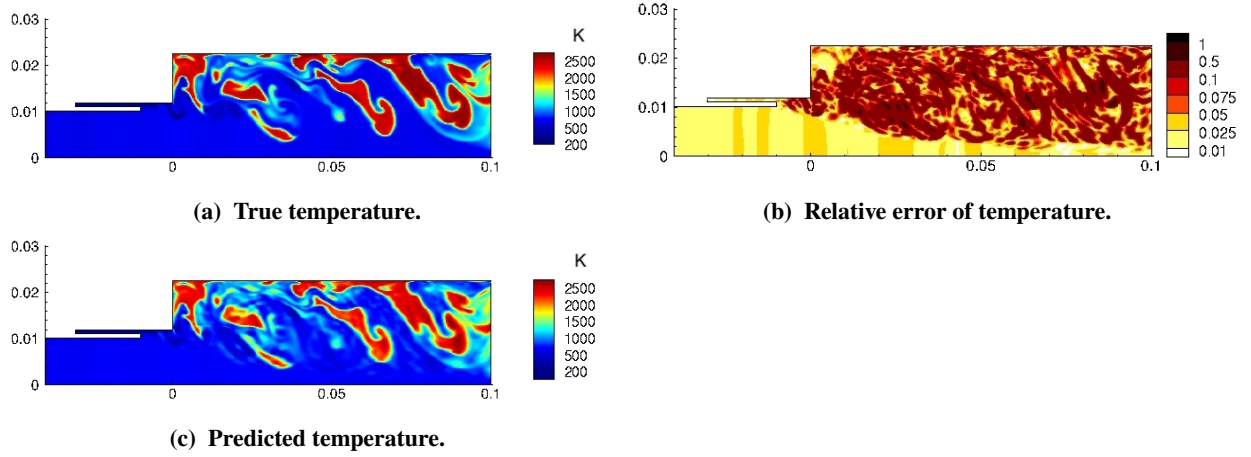
**Fig. 11** Predictive results for velocity  $v_y$  at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0E + 04$ .

In comparison with the approximately 200h of CPU time needed on a 40-core architecture (more details in Section IV.A) to compute the first 10000 snapshots of data, the ROMs provide five to six orders of magnitude in computational speedup.

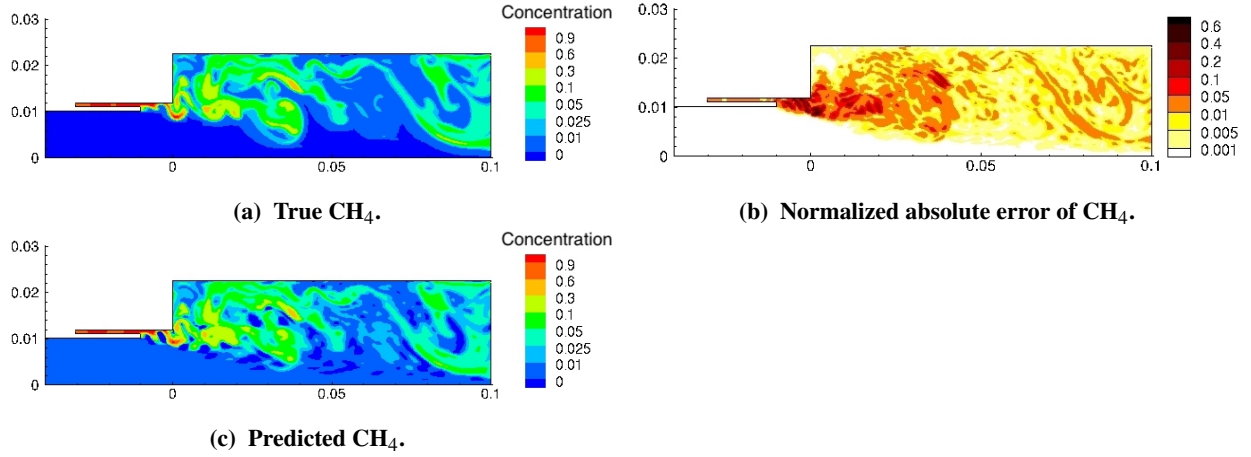
**Table 2** CPU times for two ROMs with time step size  $\Delta t = 1 \times 10^{-7}s$  and 30000 time steps.

ROM order	LS solve of (16)	ROM simulation	Reconstructing high-dim. field
$r = 24$	2.80s	6.06s	0.04s
$r = 29$	6.05s	6.22s	0.05s

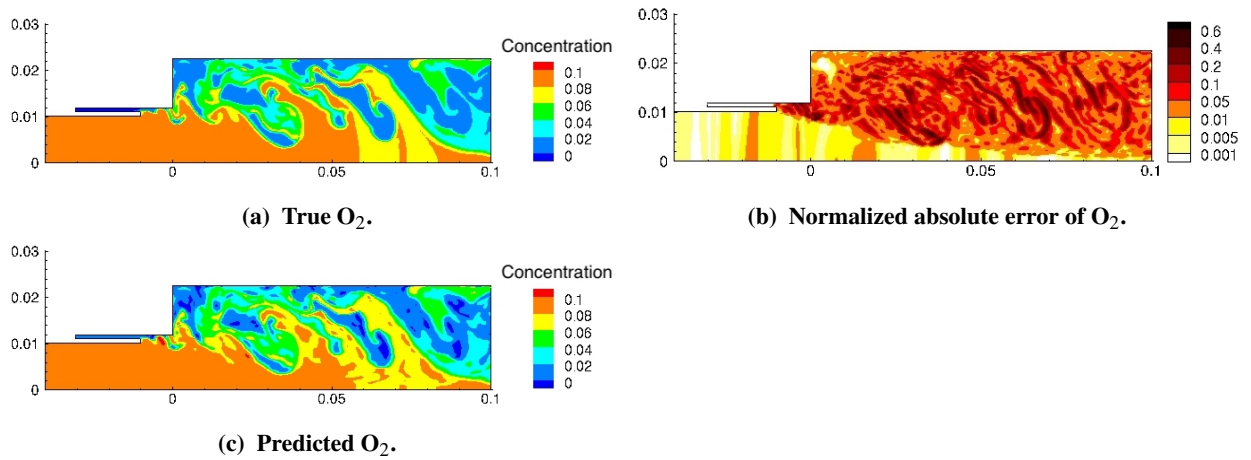




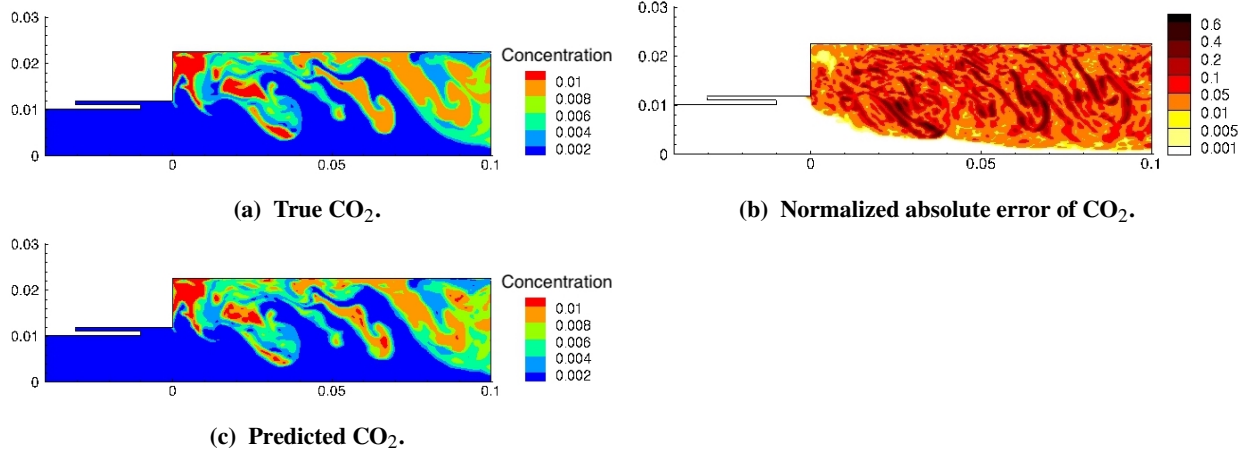
**Fig. 12** Predictive results for temperature at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0E + 04$ .



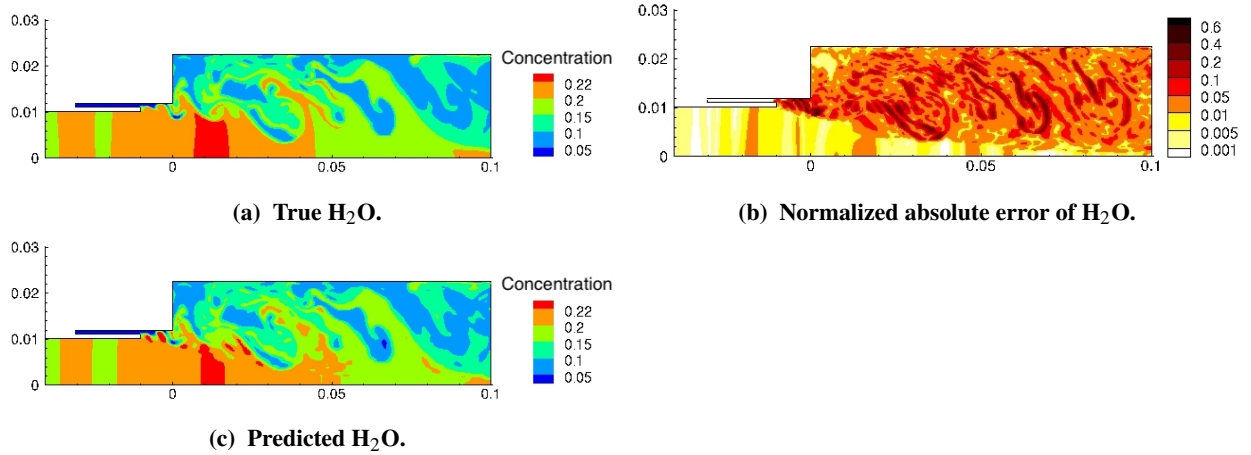
**Fig. 13** Predictive results for CH<sub>4</sub> molar concentration at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0E + 04$ .



**Fig. 14** Predictive results for O<sub>2</sub> molar concentration at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0E + 04$ .



**Fig. 15** Predictive results for  $\text{CO}_2$  molar concentration at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0\text{E} + 04$ .



**Fig. 16** Predictive results for  $\text{H}_2\text{O}$  molar concentration at the last time step of training data. Training with 10000 snapshots, a basis size of  $r = 29$ , and regularization set to  $\lambda = 3.0\text{E} + 04$ .

## V. Conclusion

Operator inference is a data-driven method for learning reduced-order models (ROMs) of dynamical systems with polynomial structure. This paper demonstrates how variable transformations can expose quadratic structure in the nonlinear system of partial differential equations describing a complex combustion model. This quadratic structure is preserved under projection, providing the mathematical justification for learning a quadratic ROM using operator inference. An important feature of the approach is that the learning of the ROM is non-intrusive—it requires state solutions generated by running the high-fidelity combustion model, but it does not require access to the discretized operators of the governing equations. This is important because it means that the variable transformations can be applied as a post-processing step to the simulation data set, but no intrusive modifications are needed to the high-fidelity code. While the quadratic model form is an approximation for this particular application problem, the numerical results show that the learned quadratic ROM can predict relevant quantities of interest and can also conserve species accurately. Many nonlinear equations in scientific and engineering applications admit variable transformations that expose polynomial structure. This combined with the non-intrusive nature of the approach make it a viable option for deriving ROMs for complex nonlinear applications where traditional intrusive model reduction is impractical and/or unreliable. While we improved ROM stability through the presented regularization of the least-squares problem, some of the ROMs were unstable. Future work thus includes devising alternative approaches (see e.g., [23, 24]) and developing theory to guarantee stable learned ROMs.

## Appendix A: Lifting chemical source terms

The chemical source terms in  $\vec{S}$  in Eq. (4) are  $\dot{\omega}_l = M_l \dot{c}_l^{\text{reaction}}$ , see Eq. (6). The dynamics for the source terms  $\dot{c}_l^{\text{reaction}}$  are given by

$$\dot{c}_1^{\text{reaction}} = -A \exp\left(-\frac{E_a}{R_u T}\right) c_1^{0.2} c_2^{1.3}, \quad (23a)$$

$$\dot{c}_2^{\text{reaction}} = 2\dot{c}_1^{\text{reaction}}, \quad (23b)$$

$$\dot{c}_3^{\text{reaction}} = -\dot{c}_1^{\text{reaction}}, \quad (23c)$$

$$\dot{c}_4^{\text{reaction}} = -2\dot{c}_1^{\text{reaction}}. \quad (23d)$$

To lift this system to polynomial form, we introduce the auxiliary variables

$$w_1 = c_1^{0.2}, \quad w_2 = c_1^{-1}, \quad w_3 = c_2^{1.3}, \quad w_4 = c_2^{-1}, \quad w_5 = \exp\left(-\frac{E_a}{R_u T}\right), \quad w_6 = \frac{1}{T^2}.$$

The source term dynamics (23) are then cubic in  $w_1, w_3, w_5$ :

$$\dot{c}_1^{\text{reaction}} = -A w_1 w_3 w_5, \quad (24a)$$

$$\dot{c}_2^{\text{reaction}} = 2A w_1 w_3 w_5, \quad (24b)$$

$$\dot{c}_3^{\text{reaction}} = -A w_1 w_3 w_5, \quad (24c)$$

$$\dot{c}_4^{\text{reaction}} = -2A w_1 w_3 w_5. \quad (24d)$$

We next derive the dynamics for the auxiliary variables  $w_1, \dots, w_6$ . For instance, we have  $\dot{w}_1 = 0.2 c_1^{0.2-1} \dot{c}_1 = 0.2 w_1 w_2 \dot{c}_1$ . Similarly, we obtain the system of auxiliary dynamics:

$$\dot{w}_1 = 0.2 w_1 w_2 \dot{c}_1,$$

$$\dot{w}_2 = -w_1^2 \dot{c}_1,$$

$$\dot{w}_3 = 1.3 w_3 w_4 \dot{c}_2,$$

$$\dot{w}_4 = -w_4^2 \dot{c}_2,$$

$$\dot{w}_5 = \frac{E_a}{R_u} \frac{1}{T^2} w_5 \dot{T} = \frac{E_a}{R_u} w_5 w_6 \dot{T},$$

$$\dot{w}_6 = -2 \frac{1}{T^3} = -2T w_6^2.$$

The dynamics of the lifted variables are quintic in the variables  $w_1, \dots, w_6$ . If we further include an additional auxiliary variable  $w_7 = w_1 w_3 w_5$ , then we obtain the system of equations

$$\begin{aligned}\dot{w}_1 &= 0.2Aw_1w_2w_7, & \dot{w}_2 &= -Aw_1^2w_7, \\ \dot{w}_3 &= 2.6Aw_3w_4w_7, & \dot{w}_4 &= -2Aw_4^2w_7, \\ \dot{w}_5 &= \frac{E_a}{R_u}w_5w_6\dot{T}, & \dot{w}_6 &= -2Tw_6^2, \\ 0 &= w_7 - w_1w_3w_5.\end{aligned}$$

The temperature  $T = \frac{p\xi}{R(Y_l)}$  can be obtained from the states  $\xi, p, Y_l$  via the ideal gas relationship in Eq. (8).

## Appendix B: Equations for pressure and chemical species

Here, we give the complete governing equation for the species molar concentrations  $c_l$  and pressure  $p$ . Recall from Section III.C the notation used to denote the conservative variables  $g_i$ :  $g_1 = \rho$ ,  $g_2 = \rho v_x$ ,  $g_3 = \rho v_y$ ,  $g_4 = \rho e$ ,  $g_5 = \rho Y_1$ ,  $g_6 = \rho Y_2$ ,  $g_7 = \rho Y_3$ ,  $g_8 = \rho Y_4$ .

**Species molar concentrations  $c_l$ :** For the species molar concentration dynamics, we use the relationship  $c_l = \frac{\rho Y_l}{M_l}$  from Eq. (5), where the constants  $M_1, \dots, M_4$  are molar masses. We obtain for  $l = 1, 2, \dots, n_{\text{sp}}$ :

$$\begin{aligned}\frac{\partial c_l}{\partial t} &= \frac{1}{M_l} \frac{\partial \rho Y_l}{\partial t} \\ &= \frac{1}{M_l} \left( \dot{\omega}_l + \nabla \cdot \left( -v_x \rho Y_l \vec{i} - v_y \rho Y_l \vec{j} + \vec{j}_l^m \right) \right) \\ &= \frac{1}{M_l} \left( M_l \dot{c}_l^{\text{reaction}} - M_l \left( \frac{\partial v_x c_l}{\partial x} + \frac{\partial v_y c_l}{\partial y} \right) + \nabla \cdot \vec{j}_l^m \right) \\ &= \dot{c}_l^{\text{reaction}} - \left( \frac{\partial v_x c_l}{\partial x} + \frac{\partial v_y c_l}{\partial y} + \frac{1}{M_l} \nabla \cdot \vec{j}_l^m \right).\end{aligned}$$

Note that the chemical source terms  $\dot{c}_l^{\text{reaction}}$  were given in Appendix A, and from Eq. (24) we see that the  $\dot{c}_l$  are cubic in the auxiliary lifted states  $w_1, \dots, w_6$ . The divergence term is

$$\nabla \cdot \vec{j}_l^m = D_l \left( \frac{\partial}{\partial x} \left( \rho \frac{\partial Y_l}{\partial x} \right) + \frac{\partial}{\partial y} \left( \rho \frac{\partial Y_l}{\partial y} \right) \right),$$

and since  $Y_l = M_l c_l \xi$ , we have that

$$\rho \frac{\partial Y_l}{\partial x} = \rho M_l \frac{\partial c_l \xi}{\partial x} = M_l \left( \frac{\partial c_l}{\partial x} + \rho c_l \frac{\partial \xi}{\partial x} \right).$$

Overall, we have that

$$\frac{\partial c_l}{\partial t} = \dot{c}_l^{\text{reaction}} - \left( \frac{\partial v_x c_l}{\partial x} + \frac{\partial v_y c_l}{\partial y} + \frac{\partial c_l}{\partial x} + \rho c_l \frac{\partial \xi}{\partial x} + \frac{\partial c_l}{\partial y} + \rho c_l \frac{\partial \xi}{\partial y} \right),$$

which is quadratic in the learning variables  $v_x, v_y, c_l$  with the exception of the terms  $\rho c_l \frac{\partial \xi}{\partial x}$  and  $\rho c_l \frac{\partial \xi}{\partial y}$ . We note that if  $\rho$  were included as a lifted variable (in addition to  $\xi$ ), these terms would become quadratic in the lifted state.

**Pressure  $p$ :** We start with the energy equation (2). By multiplying with density  $\rho$ , we have  $\rho e = \rho h^0 - p$ , so from the conservation equation (1) for  $\rho e$  we obtain

$$\frac{\partial (\rho h^0 - p)}{\partial t} + \frac{\partial \rho v_x h^0}{\partial x} + \frac{\partial \rho v_y h^0}{\partial y} + \frac{\partial}{\partial x} (v_x \tau_{xx} + v_y \tau_{yx} - j_x^q) + \frac{\partial}{\partial y} (v_x \tau_{xy} + v_y \tau_{yy} - j_y^q) = 0.$$

This directly gives an equation for the time evolution of pressure:

$$\frac{\partial p}{\partial t} = \frac{\partial \rho h^0}{\partial t} + \frac{\partial \rho v_x h^0}{\partial x} + \frac{\partial \rho v_y h^0}{\partial y} + \frac{\partial}{\partial x} (v_x \tau_{xx} + v_y \tau_{yx} - j_x^q) + \frac{\partial}{\partial y} (v_x \tau_{xy} + v_y \tau_{yy} - j_y^q).$$

Moreover, per definition of  $h^0$  in Eq. (2) and with  $c_l = \frac{\rho Y_l}{M_l}$  we have that

$$\frac{\partial \rho h^0}{\partial t} = \sum_{i=1}^{n_{sp}} \frac{\partial \rho h_l Y_l}{\partial t} + \frac{\partial \rho \frac{1}{2} (v_x^2 + v_y^2)}{\partial t} = \sum_{i=1}^{n_{sp}} M_l \frac{\partial h_l c_l}{\partial t} + \frac{1}{2} (v_x^2 + v_y^2) \frac{\partial \rho}{\partial t} + \rho \frac{\partial (v_x + v_y)}{\partial t}.$$

Overall, we have that

$$\begin{aligned} \frac{\partial p}{\partial t} = & \sum_{i=1}^{n_{sp}} M_l \left( h_l \frac{\partial c_l}{\partial t} + c_l \frac{\partial h_l}{\partial t} \right) + \frac{1}{2} (v_x^2 + v_y^2) \frac{\partial \rho}{\partial t} + \rho \frac{\partial (v_x + v_y)}{\partial t} + \frac{\partial \rho v_x h^0}{\partial x} + \frac{\partial \rho v_y h^0}{\partial y} \\ & + \frac{\partial}{\partial x} (v_x \tau_{xx} + v_y \tau_{yx} - j_x^q) + \frac{\partial}{\partial y} (v_x \tau_{xy} + v_y \tau_{yy} - j_y^q). \end{aligned}$$

This equation remains nonlinear in our chosen learning variables  $\vec{q}_L$ . In particular, the enthalpies  $h_l = h_l(T)$  and their time derivatives are nonlinear functions of temperature. The other terms show some polynomial structure; for example, in Section III.C we showed that  $\frac{\partial (v_x + v_y)}{\partial t}$  is quadratic in the learning state variables  $p, v_x, v_y, \xi$ . However, to write this equation exactly in a polynomial form would require introducing a large number of auxiliary variables along with their corresponding dynamics. We have instead chosen to introduce an approximation by learning a ROM in the variables  $\vec{q}_L$  with quadratic form.

## Funding Sources

This work has been supported in part by the Air Force Center of Excellence on Multi-Fidelity Modeling of Rocket Combustor Dynamics award FA9550-17-1-0195, and the Air Force Office of Scientific Research (AFOSR) MURI on managing multiple information sources of multi-physics systems award numbers FA9550-15-1-0038 and FA9550-18-1-0023.

## References

- [1] Bui-Thanh, T., Damodaran, M., and Willcox, K. E., “Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition,” *AIAA Journal*, Vol. 42, No. 8, 2004, pp. 1505–1516.
- [2] Tadmor, G., Centuori, M., Lehmann, O., Noack, B., Luchtenburg, M., and Morzynski, M., “Low Order Galerkin Models for the Actuated Flow Around 2-D Airfoils,” *45th AIAA Aerospace Sciences Meeting and Exhibit*, 2007, p. 1313.
- [3] Lieu, T., and Farhat, C., “Adaptation of Aeroelastic Reduced-Order Models and Application to an F-16 Configuration,” *AIAA Journal*, Vol. 45, No. 6, 2007, pp. 1244–1257.
- [4] Amsellem, D., Cortial, J., and Farhat, C., “Towards Real-Time Computational-Fluid-Dynamics-Based Aeroelastic Computations using a Database of Reduced-Order Information,” *AIAA Journal*, Vol. 48, No. 9, 2010, pp. 2029–2037.
- [5] Berger, Z., Low, K., Berry, M., Glauser, M., Kostka, S., Gogineni, S., Cordier, L., and Noack, B., “Reduced Order Models for a High Speed Jet with Time-Resolved PIV,” *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2013, p. 11.
- [6] Brunton, S. L., Rowley, C. W., and Williams, D. R., “Reduced-Order Unsteady Aerodynamic Models at Low Reynolds Numbers,” *Journal of Fluid Mechanics*, Vol. 724, 2013, p. 203–233.
- [7] Kramer, B., Grover, P., Boufounos, P., Nabi, S., and Benosman, M., “Sparse Sensing and DMD-Based Identification of Flow Regimes and Bifurcations in Complex Flows,” *SIAM Journal on Applied Dynamical Systems*, Vol. 16, No. 2, 2017, pp. 1164–1196.
- [8] Nguyen, V., Buffoni, M., Willcox, K., and Khoo, B., “Model reduction for reacting flow applications,” *International Journal of Computational Fluid Dynamics*, Vol. 28, No. 3-4, 2014, pp. 91–105.
- [9] Nguyen, V. B., Dou, H.-S., Willcox, K., and Khoo, B.-C., “Model Order Reduction for Reacting Flows: Laminar Gaussian Flame Applications,” *30th International Symposium on Shock Waves 1*, Springer, 2017, pp. 337–343.
- [10] Buffoni, M., and Willcox, K., “Projection-based model reduction for reacting flows,” *40th Fluid Dynamics Conference and Exhibit, Chicago, IL, June 28-July 1*, 2010, p. 5008.

- [11] Huang, C., Duraisamy, K., and Merkle, C. L., “Investigations and Improvement of Robustness of Reduced-Order Models of Reacting Flow,” *AIAA Journal*, Vol. 0, No. 0, 0, pp. 1–13. doi:10.2514/1.J058392, URL <https://doi.org/10.2514/1.J058392>.
- [12] Huang, C., Xu, J., Duraisamy, K., and Merkle, C., “Exploration of reduced-order models for rocket combustion applications,” *2018 AIAA Aerospace Sciences Meeting*, 2018, p. 1183.
- [13] Wang, Q., Hesthaven, J. S., and Ray, D., “Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem,” *Journal of computational physics*, Vol. 384, 2019, pp. 289–307.
- [14] Lumley, J. L., “The Structure of Inhomogeneous Turbulent Flows,” *Atmospheric Turbulence and Radio Wave Propagation*, 1967.
- [15] Sirovich, L., “Turbulence and the Dynamics of Coherent Structures. I-Coherent Structures. II-Symmetries and Transformations. III-Dynamics and Scaling,” *Quarterly of Applied Mathematics*, Vol. 45, 1987, pp. 561–571.
- [16] Veroy, K., Rovas, D. V., and Patera, A. T., “A posteriori error estimation for reduced-basis approximation of parametrized elliptic coercive partial differential equations: ‘convex inverse’ bound conditioners,” *ESAIM: Control, Optimisation and Calculus of Variations*, Vol. 8, 2002, pp. 1007–1028.
- [17] Grepl, M. A., Maday, Y., Nguyen, N. C., and Patera, A. T., “Efficient Reduced-Basis Treatment of Nonaffine and Nonlinear Partial Differential Equations,” *ESAIM: Mathematical Modelling and Numerical Analysis*, Vol. 41, No. 3, 2007, pp. 575–605.
- [18] Tröltzsch, F., and Volkwein, S., “POD a-posteriori error estimates for linear-quadratic optimal control problems,” *Computational Optimization and Applications*, Vol. 44, No. 1, 2009, p. 83.
- [19] Hesthaven, J. S., Rozza, G., and Stamm, B., *Certified reduced basis methods for parametrized partial differential equations*, Springer, 2016.
- [20] Rowley, C. W., Colonius, T., and Murray, R. M., “Model reduction for compressible flows using POD and Galerkin projection,” *Physica D: Nonlinear Phenomena*, Vol. 189, No. 1-2, 2004, pp. 115–129.
- [21] Barone, M. F., Kalashnikova, I., Segalman, D. J., and Thornquist, H. K., “Stable Galerkin reduced order models for linearized compressible flow,” *Journal of Computational Physics*, Vol. 228, No. 6, 2009, pp. 1932–1946.
- [22] Serre, G., Lafon, P., Gloerfelt, X., and Bailly, C., “Reliable reduced-order models for time-dependent linearized Euler equations,” *Journal of Computational Physics*, Vol. 231, No. 15, 2012, pp. 5176–5194.
- [23] Kalashnikova, I., and Barone, M., “Stable and efficient Galerkin reduced order models for non-linear fluid flow,” *6th AIAA Theoretical Fluid Mechanics Conference*, 2011, p. 3110.
- [24] Balajewicz, M., Tezaur, I., and Dowell, E., “Minimal subspace rotation on the Stiefel manifold for stabilization and enhancement of projection-based reduced order models for the compressible Navier–Stokes equations,” *Journal of Computational Physics*, Vol. 321, 2016, pp. 224–241.
- [25] Carlberg, K., Barone, M., and Antil, H., “Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction,” *Journal of Computational Physics*, Vol. 330, 2017, pp. 693–734.
- [26] Huang, C., Duraisamy, K., and Merkle, C., “Challenges in Reduced Order Modeling of Reacting Flows,” *2018 Joint Propulsion Conference*, 2018, p. 4675.
- [27] Hornik, K., Stinchcombe, M., and White, H., “Multilayer feedforward networks are universal approximators,” *Neural networks*, Vol. 2, No. 5, 1989, pp. 359–366.
- [28] Swischuk, R., Mainini, L., Peherstorfer, B., and Willcox, K., “Projection-based model reduction: formulations for physics-based machine learning,” *Computers and Fluids*, Vol. 179, 2019, pp. 704–717.
- [29] Brunton, S. L., Proctor, J. L., and Kutz, J. N., “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” Vol. 113, No. 15, 2016, pp. 3932–3937. doi:10.1073/pnas.1517384113.
- [30] Rudy, S. H., Brunton, S. L., Proctor, J. L., and Kutz, J. N., “Data-driven discovery of partial differential equations,” *Science Advances*, Vol. 3, No. 4, 2017, p. e1602614.
- [31] Schaeffer, H., Tran, G., and Ward, R., “Extracting sparse dynamics from limited data,” *SIAM Journal on Applied Mathematics*, Vol. 78, No. 6, 2018, pp. 3279–3295.

- [32] Schmid, P. J., “Dynamic mode decomposition of numerical and experimental data,” *Journal of Fluid Mechanics*, Vol. 656, 2010, pp. 5–28.
- [33] Rowley, C. W., Mezić, I., Bagheri, S., Schlatter, P., and Henningson, D. S., “Spectral analysis of nonlinear flows,” *Journal of Fluid Mechanics*, Vol. 641, 2009, pp. 115–127.
- [34] Peherstorfer, B., and Willcox, K., “Data-driven operator inference for nonintrusive projection-based model reduction,” *Computer Methods in Applied Mechanics and Engineering*, Vol. 306, 2016, pp. 196–215.
- [35] Qian, E., Kramer, B., Marques, A., and Willcox, K., “Transform & Learn: A data-driven approach to nonlinear model reduction,” *AIAA Aviation and Aeronautics Forum and Exposition*, Dallas, TX, 2019. doi:10.2514/6.2019-3707.
- [36] Gu, C., “QLMOR: A projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 30, No. 9, 2011, pp. 1307–1320.
- [37] Kramer, B., and Willcox, K., “Nonlinear Model Order Reduction via Lifting Transformations and Proper Orthogonal Decomposition,” *AIAA Journal*, Vol. 57, No. 6, 2019, pp. 2297–2307. doi:10.2514/1.J057791.
- [38] Kramer, B., and Willcox, K., “Balanced Truncation Model Reduction for Lifted Nonlinear Systems,” Tech. rep., 2019. ArXiv:1907.12084.
- [39] Harvazinski, M. E., Huang, C., Sankaran, V., Feldman, T. W., Anderson, W. E., Merkle, C. L., and Talley, D. G., “Coupling between hydrodynamics, acoustics, and heat release in a self-excited unstable combustor,” *Physics of Fluids*, Vol. 27, No. 4, 2015, p. 045102.
- [40] Harvazinski, M. E., “Modeling self-excited combustion instabilities using a combination of two-and three-dimensional simulations,” Ph.D. thesis, Purdue University, 2012.
- [41] Huang, C., Gejji, R., Anderson, W., Yoon, C., and Sankaran, V., “Combustion Dynamics in a Single-Element Lean Direct Injection Gas Turbine Combustor,” *Combustion Science and Technology*, 2019, pp. 1–28.
- [42] Yu, Y., O’Hara, L., Sisco, J., and Anderson, W., “Experimental study of high-frequency combustion instability in a continuously variable resonance combustor (CVRC),” *47th AIAA Aerospace Sciences Meeting*, Orlando, FL, 2009.
- [43] Westbrook, C. K., and Dryer, F. L., “Simplified reaction mechanisms for the oxidation of hydrocarbon fuels in flames,” *Combustion Science and Technology*, Vol. 27, No. 1-2, 1981, pp. 31–43.
- [44] Martins, J. R., and Hwang, J. T., “Review and unification of methods for computing derivatives of multidisciplinary computational models,” *AIAA journal*, Vol. 51, No. 11, 2013, pp. 2582–2599.
- [45] Knowles, I., and Renka, R. J., “Methods for numerical differentiation of noisy data,” *Electron. J. Differ. Equ.*, Vol. 21, 2014, pp. 235–246.
- [46] Chartrand, R., “Numerical differentiation of noisy, nonsmooth, multidimensional data,” *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, IEEE, 2017, pp. 244–248.
- [47] Golub, G. H., and Van Loan, C. F., “Matrix computations. 1996,” *Johns Hopkins University, Press, Baltimore, MD, USA*, 1996, pp. 374–426.
- [48] Swischuk, R. C., “Physics-based machine learning and data-driven reduced-order modeling,” Master’s thesis, Massachusetts Institute of Technology, 2019.
- [49] Martinsson, P.-G., Rokhlin, V., and Tygert, M., “A randomized algorithm for the decomposition of matrices,” *Applied and Computational Harmonic Analysis*, Vol. 30, No. 1, 2011, pp. 47–68.
- [50] Hansen, P., “The L-curve and its use in the numerical treatment of inverse problems,” *Computational Inverse Problems in Electrocardiology*, Vol. 5, 2001, pp. 119–142.