Surrogate Thermochemical Kinetics for Nonequilibrium Hypersonic Flows

Claudio Rapisarda*, Justin Clarke †, Matthew McGilvray‡ and Luca di Mare §
Oxford Thermofluids Institute, University of Oxford, Oxford, Oxfordshire, OX1 2JD

Nonequilibrium thermochemical processes such as vibrational excitation, chemical dissociation, and ionization are critical to accurately characterize hypersonic flows. The numerical modelling of such mechanisms is computationally demanding due to the onerous number of degrees of freedom and kinetic parameters coupled with wide flow spatio-temporal scales, resulting in highly non-linear and stiff chemical kinetics. To accelerate computing of reacting flow simulations without losing fidelity, data-driven strategies, such as Principal Component Analysis (PCA) and physics-based models, such as Computational Singular Perturbation (CSP), are typically implemented. In the present work, a novel reduced-order model (ROM) framework is proposed to combine the two strategies into a surrogate model. First, it is shown that the equations governing the evolution in space and time of a small disturbance around an equilibrium state can be formulated as a generalised eigenvalue problem. The solution to this problem can be used to represent the evolution of disturbances relaxing towards equilibrium downstream of an initial perturbation, such as a shock. The eigenvectors of the system distinguish between groups of slow non-equilibrium processes and fast, near-equilibrium processes. The calculation of the system eigen-basis, however, is the most computationally expensive operation of the numerical algorithm. This paper proposes the substitution of this operation with non-linear regression fitting, namely Gaussian Process Regression (GPR), to map the system eigenvalues and eigenvectors. To reduce the training cost of GPR, a truncated basis of independent variables is identified using PCA. The method is applied to a two-temperature plasma evolving downstream of a shock, modelled using Park's two-temperature model with 11 species for air. The feasibility and fidelity of the novel ROM framework are presented, demonstrating the viable application of computational learning methods for eigen-basis prediction.

I. Nomenclature

CSP Computational Singular Perturbation

FDM Finite Difference Method

GPR Gaussian Process Regression

ILDM Intrinsic Low Dimensional Manifold

MMT Modified Marrone-Treanor

OCEAN Oxford Chemical Equilibrium Analysis

PCA Principal Component Analysis

PC Principal Component

RMSE Root Mean Squared Error

ROM Reduced Order Model

*DPhil Candidate, Department of Engineering Science, Oxford Thermofluids Institute, University of Oxford
†DPhil Candidate, Department of Engineering Science, Oxford Thermofluids Institute, University of Oxford
‡Professor, Department of Engineering Science, Oxford Thermofluids Institute, University of Oxford
§Associate Professor, Department of Engineering Science, Oxford Thermofluids Institute, University of Oxford

A_s = Matrix of truncated eigenvectors for PCA
B = Basis functions
f_ϕ = Non-linear regression function
$\boldsymbol{F}$ = Vector of inviscid flux
h_0 = Total specific enthalpy
$\hat{D}_s$ = Gained or lost specific vibrational energy
i = Index
$\hat{I}$ = First ionization energy
J = Jacobian matrix
k_b = Backwards reactant rate coefficient
k_f = Forward reactant rate coefficient
$\boldsymbol{K}$ = Vector of canonical variables
$\boldsymbol{L}$ = Left-eigenvector
M = Number of fast modes
M_w = Molar weight
n = Number of species
$\dot{n}_e$ = Molar rate of ionization
nr = Number of reactions
n_s = Number of samples for PCA
N = Number of state variables
N_η = Truncated number of variables using PCA
p = Pressure
p_e = Electron partial pressure
$\boldsymbol{Q}$ = Vector of primitive variables
$\bar{\boldsymbol{Q}}$ = Mean state vector of primitive variables
$\hat{\boldsymbol{Q}}$ = Vector of disturbances in primitive variables
$\boldsymbol{R}$ = Right-eigenvector
S = Scaling matrix for PCA
t = Time
T_{eff} = Effective temperature
u = Velocity
$\boldsymbol{U}$ = Vector of conservative variables
$\dot{w}$ = Mass production and destruction term
$\dot{w}_v$ = Electro-vibrational source term
x = Spatial dimension
X = Matrix of sampled data for PCA
$\bar{X}$ = Matrix of mean value of sampled data for PCA
$\tilde{X}$ = Matrix of scaled sampled data for PCA

λ = Eigenvalues
Λ = Diagonal matrix of eigenvalues
ρ = Density
σ = Mole-mass ratio
σ_e = Square root of error variance
τ = Chemical time-scales
τ_f = Fast chemical timescales
τ_s = Relaxation time
$\boldsymbol{\xi}$ = Local perturbation in physical space
$\boldsymbol{\zeta}$ = Local perturbation in eigen-space
$\Im$ = Scalar non-linear combination of thermochemical variables
ν = Wavenumber
ω = Angular frequency

II. Introduction

Non-equilibrium phenomena exert a significant influence on the heat transfer, radiative transport properties and aerodynamic performance of spacecraft (re-)entering planetary atmospheres. Vibrational excitation, chemical dissociation and ionization are chemical and kinetic processes characteristic of hypersonic flight that absorb energy, alter flow compressibility, lower temperature and increase density [1]. The chemical mechanisms described by such processes, however, are very complex and require the definition of numerous degrees of freedom and associated constants, such as reaction rates and relaxation times. While computing the large number of reactive scalar variables and solving the associated transport equations can be computationally demanding for multi-temperature models [2–5], the computational cost may become prohibitive for high-fidelity collisional models, also referred to as state-to-state, which treat each molecular energy state as a separate pseudo-species, thus requiring hundreds of internal energy states and thousands of kinetic processes [6–8].

In addition to the extensive number of variables and equations to compute, the variety of spatio-temporal scales inherently coupled through thermochemical and fluid dynamic interactions increases the numerical stiffness of reactive flow problems [9–11]. In fact, the complex nonlinear reaction source terms are imparted by the detailed chemical kinetic mechanisms which involve temporal scales spanning twelve orders of magnitude [12].

To deal with the wide temporal differences between nonequilibrium and transport processes, a common approach is that of employing an operator splitting strategy in which the chemistry effects are decoupled from the transport equations. A larger integration step can thus be used for the convective and diffusive scales while the stiff chemical source terms are solved using a dedicated time integrator. However, such strategies not only are subject to splitting errors [13] but also fail to remove the numerical stiffness from the chemical scale spectrum. As a consequence, implicit integration schemes based on backward differentiation formula [14] are often used, requiring a large computational overhead for large Jacobian matrix operations [15].

To alleviate the stiff kinetics ordinary differential equations integration cost, reduced order models (ROMs) or similar surrogates are typically proposed in the form of physics-based or data-based strategies [11]. These models leverage on the projection of the original flow dynamics onto a transformed basis that is of reduced dimensionality, less stiff or a combination of both. Among the most established data-based approaches, the principal component analysis (PCA) produces low-dimensional representation of the thermochemical state by creating linear projections of the input data that minimize statistical variance [16]. Such reduced representations can then be integrated by means of projected governing equations [16]. However, PCA does not guarantee stiffness reduction since it has no knowledge of the temporal dynamics.

Conversely, physics-based surrogates are based on the flow temporal dynamics and have the advantage of guaranteeing a reduction in the complexity of the chemical system and, most importantly, in its stiffness. Leveraging on the generalization of the quasi-steady state [17] and partial-equilibrium assumptions, the Intrinsic Low Dimensional Manifold (ILDM) [18–20], the Computational Singular Perturbation (CSP) [21–24] and its derivative, the G-Scheme [25, 26], have been successfully applied to complex reactive-flow systems. These techniques construct basis vectors of the Jacobian matrix of the local chemical source terms such that the ODEs associated to the fast time scales are replaced with algebraic relations and the reduced system is projected onto the slow invariant manifold. This leads to the development of an algebraically constrained non-stiff reduced order model that can be solved explicitly with a much larger time-step, typically of the same order as the fastest of the slow time-scales. Nevertheless, one major challenge of such physics-based models is that the computational cost associated with the evaluation of the eigenvalues and eigenvectors of the Jacobian matrix required for the projection basis increases exponentially with larger mechanisms [27], scaling with the cube of the matrix size as shown in Figure 13.

To overcome the drawback of the large computational cost for the eigenbasis projection, whilst addressing the need of reduced complexity and stiffness of chemical kinetics without losing fidelity, this paper aims to develop a new ROM framework that combines physics-based and data-driven techniques to establish non-linear regression fitting of the eigenspace as illustrated in Figure 1. The broader applicability of this work is envisioned in advancing numerical flow simulations for hypersonic vehicles, particularly those requiring increased complexity and fidelity. The eigen-decomposition of the Jacobian matrix is especially valuable in nonequilibrium reactive flows, where thermochemical kinetics introduce significant numerical stiffness. By leveraging the eigenbasis, it becomes possible to identify and separate slow and fast thermochemical processes, thereby reducing numerical stiffness and enabling the selection of an optimal integration time step. Additionally, the eigenbasis offers insights into the system's numerical stability and physical behaviour [29]. However, as the complexity and fidelity of flow simulations increase, the computational cost of eigen-decomposition also rises. Developing surrogate models for eigenbasis computation is expected to become more significant as the number of equations and transport variables is increased. Such models could also prove advantageous

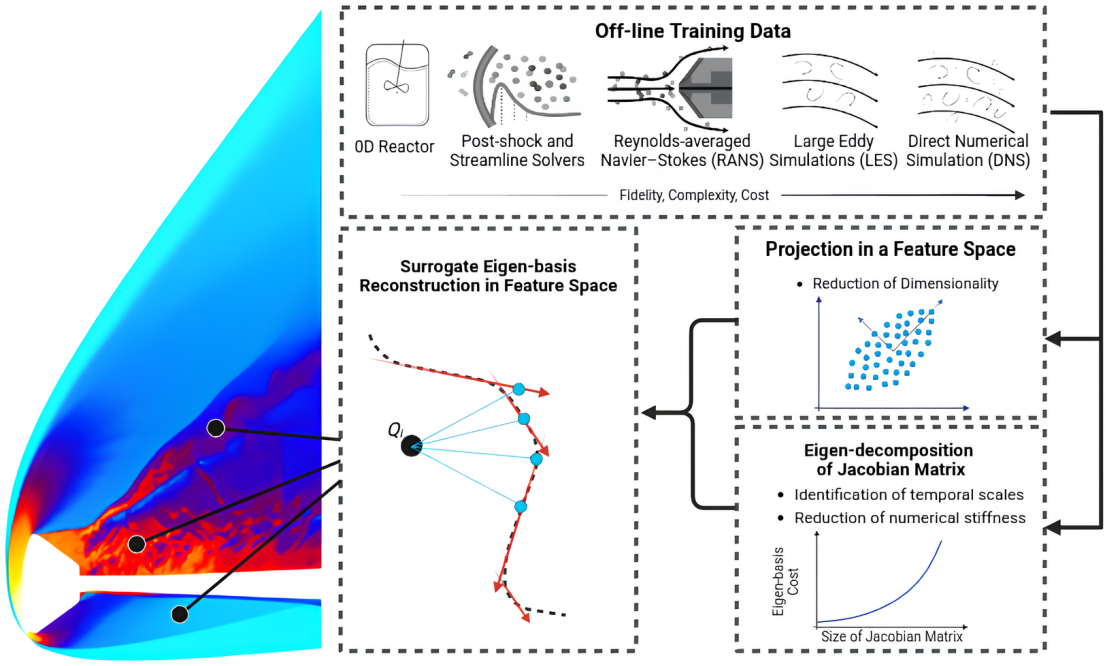


Fig. 1 Framework for training of surrogate eigen-basis models for hypersonic vehicle simulations. Contour plot of flow around re-entry capsule (left) adapted from [28]

in trajectory calculations, where interpolation between different flow solutions may be required. The creation of these surrogate models, however, relies on the availability of high-quality training data, which can be derived from diverse numerical datasets, each varying in the amount of information captured on the thermochemical processes [30].

To demonstrate the feasibility of the development of surrogate models, a 0D two-temperature post-shock relaxation problem is presented in this work according to the framework shown in Figure 2. A small-perturbation analysis is proposed, inspired by the ILDM and G-Scheme approaches. In this analysis, the eigendecomposition of the product of the inverse inviscid flux Jacobian and the reactive Jacobian is employed. This approach eliminates the need for isobaric or isochoric assumptions, as the system is fully constrained by the conservation equations. PCA is then employed to project the state vector onto a truncated basis of principal components that retains most of the statistical variance. This reduces the number of dependant variables needed for the non-linear regression fitting. While the framework may be adapted to other fitting routines, the Gaussian Process Regression (GPR) method is employed for the first time to create a non-linear kernel-based fitting of the eigenbasis onto the principal components. This paper demonstrates that the eigen-decomposition of the Jacobian matrix can be replaced by alternative fitting strategies to model and simplify chemical kinetics of hypersonic flow simulations in thermochemical nonequilibrium.

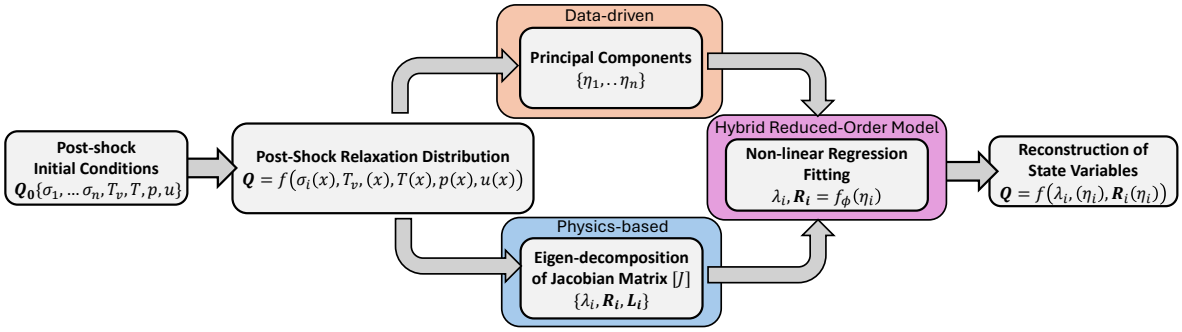


Fig. 2 Schematic of Proposed ROM framework for post-shock eigen-space surrogate

III. Methodology

A. Governing Equations

The governing equations for a chemically-reacting, inviscid quasi-one dimensional nonequilibrium flow can be written in conservation form as [31]:

$$\frac{\partial \boldsymbol{U}}{\partial t} + \frac{\partial \boldsymbol{F}}{\partial x} = \boldsymbol{W} \tag{1}$$

where the vector quantities $\boldsymbol{U}$, $\boldsymbol{F}$ and $\boldsymbol{W}$ are the conservative variable, inviscid flux and source term vectors, respectively. The detailed expressions for the vectors are given in Equation 2. In the notation adopted, ρ represents the density of the gas mixture, u the velocity, p the pressure, e_0 the total specific energy, h_0 the total specific enthalpy, e_v the specific vibrational energy, σ_i the individual mole-mass ratio, $\dot{w}_i$ the associated mass production and destruction for each chemical species and, finally, the source term of the electro-vibrational equation $\dot{w}_v$ [31, 32].

$$\boldsymbol{U} = \begin{pmatrix} \rho\sigma_1 \\ \vdots \\ \rho\sigma_n \\ \rho e_v \\ \rho e_0 \\ \rho \\ \rho u \end{pmatrix} \qquad \boldsymbol{F} = \begin{pmatrix} \rho\sigma_1 u \\ \vdots \\ \rho\sigma_n u \\ \rho e_v u \\ \rho h_0 u \\ \rho u \\ \rho u^2 + p \end{pmatrix} \qquad \boldsymbol{W} = \begin{pmatrix} \dot{w}_1 \\ \vdots \\ \dot{w}_n \\ \dot{w}_v \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{2}$$

B. Thermochemical Modelling

The governing equations employed in the present work assume that the translational mode of the heavy particles is fully equilibrated with the rotational mode of molecules such that both modes are represented by a singular temperature T [32]. At the same time, the distribution of vibrational, electronic and electron translational energies is described by the vibrational temperature T_v [31, 33]. The specific heats and enthalpies for the relevant gas species are evaluated at thermal equilibrium using McBride's polynomial expressions [34] and extended to thermal nonequilibrium by assuming full excitation of the translational-rotational energy modes [31].

The equilibrium thermodynamic composition of the high-enthalpy air mixture, expressed in mole-mass ratios in Equation 2, along with their corresponding thermodynamic equilibrium properties are evaluated using the in-house Oxford Chemical Equilibrium ANalysis (OCEAN) code, developed by The University of Oxford Hypersonics Group based on the minimisation of Gibbs free energy [35]. For air, a 11-species gas mixture is adopted (e^-, N^+, N_2^+, O^+, O_2^+, NO^+, NO, N, O, N_2 and O_2).

The stoichiometric relations between the reactants and products of a multi-component gas with n chemical species and $r \in nr$ number of reactions are given in generic form in Equation 3, where $\alpha_{i,r}$ and $\beta_{i,r}$ are the stoichiometric coefficients of the reactants and products for each species, χ_i is a chemical species in the fluid and $k_{f,r}$ and $k_{b,r}$ are the forward and backwards reactant rate coefficient for the r_{th} reaction respectively [36].

$$\sum_{i=1}^{n} \alpha_{i,r}\chi_i \underset{k_{b,r}}{\overset{k_{f,r}}{\rightleftharpoons}} \sum_{i=1}^{n} \beta_{i,r}\chi_i \tag{3}$$

The generation and consumption of the chemical species within the vector of production rates $\boldsymbol{W}$ in Equation 1 as a consequence of finite-rate chemical reactions may therefore be defined as in Equation 4 in terms of molar concentration:

$$\dot{w}_i = \frac{\partial \rho\sigma_i}{\partial t} = M_{w_i} \sum_{r=1}^{nr} (\beta_{i,r} - \alpha_{i,r}) \left[k_{f,r} \prod_{i=1}^{n} (\rho\sigma_i)^{\alpha_{i,r}} - k_{b,r} \prod_{i=1}^{n} (\rho\sigma_i)^{\beta_{i,r}} \right] \tag{4}$$

Since the reaction rates in Equation 3-4 are affected by the degree of thermal nonequilibrium in the flow, Park's model [37] is adopted such that the forward reaction rate coefficients are expressed in the modified Arrhenius form as a function of the effective temperature T_{eff}, defined in Equation 5 as the geometric mean of the rotational-translational and electron-vibrational temperatures. The backwards reaction rate, on the other hand, is a function of only the translational temperature [36].

$$T_{eff} = \sqrt{TT_v} \tag{5}$$

While the present model applies to any set of reaction rates and chemical reactions, the reactions in Table 4 and the corresponding rates are taken from Gupta et al. [36] with Park-93 model [2] for air.

The vibration energy source term in Equation 1 can be decomposed into several contributions as illustrated in Equation 6. The specific vibrational energy of the diatomic molecules gained or lost at the rate $\dot{w}_i$ is expressed by the variable $\hat{D}_s$ with no preferential dissociation. On the other hand, the electron energy loss due to ionization is here expressed as the product of the molar rate of ionization $\dot{n}_{e,s}$ and the first ionization energy $\hat{I}_s$. Additionally, the relaxation energy exchange between the translational-rotational modes and the electro-vibrational ones is included using Landau-Teller's model [38] between two energy states e_v^* and e_v. The relaxation time is taken from Millikan and White's semi-empirical correlation [39], and Park's correction [40] for temperatures above 8000 K is included. The work done by the electric field induced by the electron pressure gradient $p_e \nabla u$ is also encompassed [31].

$$\dot{w}_v = \sum_{s=mol} \dot{w}_s \hat{D}_s - \sum_{s=ions} \dot{n}_{e,s} \hat{I}_s + \sum_{s=mol} \rho_s \frac{e_v^* - e_v}{< \tau_s >} - p_e \nabla u \tag{6}$$

C. Finite Difference Method Solver

To investigate the steady nonequilibrium flow relaxation behind a normal shock wave, a finite-difference method is implemented. By studying the problem from the shock wave reference frame, the mathematical problem is reduced to an initial-value problem [41]. The governing equations in Equation 1 are therefore reformulated as a set of ordinary differential equations in canonical form:

$$\frac{d\boldsymbol{Q}}{dx} = \boldsymbol{K}(\boldsymbol{Q}); \quad \boldsymbol{Q}(x=0) = \boldsymbol{Q_0} \tag{7}$$

The vector of working variables $\boldsymbol{Q}$ in Equation 7 is conveniently set to equal to the primitive variables in Equation 8, where T and T_v are the translational-rotational and electro-vibrational temperatures respectively according to Park's model [33, 42]. The vector $\boldsymbol{K}(\boldsymbol{Q})$ is equal to $\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{Q}}\right)^{-1} \boldsymbol{W}$. To solve the first-order initial-value problem in Equation 7, the initial condition $\boldsymbol{Q}_0$ is derived from the Rankine-Hugoniot jump conditions across a shock by assuming that both the gas composition and the vibrational temperature remain frozen.

$$\boldsymbol{Q} = \left[\sigma_1, \quad \ldots, \quad \sigma_n, \quad T_v, \quad T, \quad p, \quad u\right]^T \tag{8}$$

Solving Equation 7 can be an expensive operation due to the high numerical stiffness caused by the wide range of time-scales spanned across the chemical processes captured in the source term vector $\boldsymbol{W}$. In fact, when an explicit integration scheme is adopted, the selection of the integration step is based on the local characteristic time scale. For the solution to be numerically stable, the slowest chemical time-scale must therefore be utilized to propagate the solution of all chemical processes. In a more efficient integration strategy, it is desirable to make the fast time-scales disappear from the problem such that the evolution of the system depends only on the slow modes. This idea is exploited in the CSP method [21–24] and its derivative, the G-Scheme method [25, 26], where the fast and slow time scales are separated. However, all the studies retrieved assume a simplified isobaric reactor system. While the assumption may be justifiable for specific combustion processes, it is not suitable for post-shock relaxation studies. The next section illustrates how this assumption is removed.

D. Small Perturbation Analysis

To study the nonequilibrium gas dynamics, the state vector of primitive variables is modelled as the sum of its mean value $\bar{\boldsymbol{Q}}$ and its fluctuating disturbance $\hat{\boldsymbol{Q}}$, where the latter is described by an harmonic wave with angular frequency ω and wave number ν according to Equation 9.

$$\boldsymbol{Q} = \bar{\boldsymbol{Q}} + \hat{\boldsymbol{Q}} e^{j(\omega t + \nu x)} \tag{9}$$

When small perturbations are assumed, the state variables deviate only slightly from equilibrium, therefore justifying the linearization of the governing equations around the equilibrium state. The mathematical system in Equation 1 is rewritten in terms of the partial derivatives with respect to the state variables in Equation 10, which is of analogous type

as a generalised homogeneous eigenvalue problem ($[A]\hat{\boldsymbol{Q}} = \lambda(\omega)[B]\hat{\boldsymbol{Q}}$) whose right eigenvectors are the perturbations introduced in the state variables Q.

$$\left(j\omega\frac{\partial \boldsymbol{U}}{\partial \boldsymbol{Q}} - \frac{\partial \boldsymbol{W}}{\partial \boldsymbol{Q}}\right)\hat{\boldsymbol{Q}} = -j\nu\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{Q}}\hat{\boldsymbol{Q}} \tag{10}$$

In the case of a stationary disturbance ($\omega = 0$), representative of a steady-state problem in which the wavelength becomes infinite, Equation 10 is reduced to the generalized eigenvalue problem in Equation 11 which is solved numerically in FORTRAN using the ZGGEVX subroutine within LAPACK. The analytical eigenvalues are shown in Equation 12 with $\lambda = \nu^{-1}$, where the terms $\mathfrak{I}_i$ correspond to non-linear combination of thermochemical variables. In agreement with the work of Maas and Pope [43], these eigenvalues correspond to the length-scales of the nonequilibrium processes that the flow is subject to, including both chemical and vibrational relaxation effects. Since the problem is describing a steady-state solution propagating along an open tube with constant cross-sectional area, the chemical relaxation distances identified by the diagonal elements of the matrix $[\Lambda] = \{\lambda_i\}_{i=1}^{N}$ provide information on the speed of the dynamic nonequilibrium processes.

$$\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{Q}}\hat{\boldsymbol{Q}} = j\nu\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{Q}}\hat{\boldsymbol{Q}} \tag{11}$$

$$[\Lambda] = \begin{bmatrix} 0 & & & \\ & \ddots & & \\ & & \frac{u}{\mathfrak{I}_1} & \\ & & & \ddots \\ & & & & \frac{u}{\mathfrak{I}_n} \end{bmatrix} \tag{12}$$

It is noted that Equation 11 is analogous to the linearization of the canonical expression in Equation 7. However, drawing from the physical interpretation of the eigenvalues, it is now clear that the eigen-decomposition of the combined Jacobian matrix $[J]$, as shown in Equation 13, enables the separation of relaxation scales to remove the stiffness from the ODE. By defining the chemical time-scales as $\tau^i = \frac{1}{|\lambda_i|}$, the associated mode contributions can be sorted from the smallest to the largest in magnitude of the real part of the eigenvalues $\boldsymbol{\lambda}$ [44]. In the notation adopted, the matrix of right-eigenvectors corresponds to the perturbations $[R] = [\hat{Q}] \equiv \{\hat{\boldsymbol{Q}}_{\boldsymbol{i}}\}_{i=1}^{N}$ and the left-eigenvectors are defined such that $[L] = [R]^{-1}$.

$$\frac{d\boldsymbol{Q}}{dx} = \left[\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{Q}}\right)^{-1}\frac{\partial \boldsymbol{W}}{\partial \boldsymbol{Q}}\right]\boldsymbol{Q} \equiv [J]\boldsymbol{Q} \rightarrow [J] = [\hat{Q}][\Lambda][\hat{Q}]^{-1} \tag{13}$$

The linearization of the state $\boldsymbol{Q}$ at a given local state $\boldsymbol{Q}^*$ such that $\boldsymbol{\xi} = \boldsymbol{Q} - \boldsymbol{Q}^*$ yields the following linear ODE:

$$\frac{d\xi_i}{d\tau} = \left(\frac{\partial \boldsymbol{F}(\boldsymbol{Q}^*)}{\partial \boldsymbol{Q}}\right)^{-1}\boldsymbol{W}(\boldsymbol{Q}^*) + [J(\boldsymbol{Q}^*)]\xi_i \tag{14}$$

By further mapping the local perturbation $\boldsymbol{\xi}$ onto the right-eigenvectors with $\xi_i = R_i^j\zeta_j$, Equation 14 is written in terms of the local perturbations in the eigen-space $\boldsymbol{\zeta}$ as shown in Equation 15, with initial condition $\zeta_i(\tau = 0) = 0$.

$$\frac{d\zeta_i}{d\tau} = L^i(\boldsymbol{Q}^*)\cdot\boldsymbol{K}(\boldsymbol{Q}^*) + \lambda_i(\boldsymbol{Q}^*)\zeta_i \tag{15}$$

Equation 15 reveals that the evolution of each latent variable, defined by the invariant space $\lambda_i = 0$, is given by $d\zeta_i^l = L^i(\boldsymbol{Q}^*)\cdot\boldsymbol{K}(\boldsymbol{Q}^*)d\tau$. These modes yield the maximum degree of time scale decoupling, and it is therefore intuitive to order the eigenvalues based on their magnitude. Conversely, the evolution of the active modes ($\lambda_i \neq 0$) is solved at third-order accuracy using Taylor's expansion for Equation 15. The solution, shown in Equation 16 in the eigen-space, can then be mapped back into physical space by using its definition $\zeta_i = L_i^j\xi_j$.

$$d{\zeta_i}^a \sim L^i\Delta x\left[1 + \frac{1}{2}\lambda_i\Delta x + \frac{1}{6}(\lambda_i\Delta x)^2\right] \tag{16}$$

The methodology presented allows for the identification of the different length-scales that the nonequilibrium flow is subject to. These can be further separated to solve the slow modes and approximate the fast ones. The number of fast modes M is determined as the largest M that satisfies Equation 17 [25, 45], where N is the number of state variables:

$$\left| \tau_{M+1} \left(\sum_{r=1}^{M} R_r^i \, (L^r \cdot \boldsymbol{K}) \right) \right| < \epsilon_r |Q^i| + \epsilon_a \qquad i = 1, ..., N \tag{17}$$

where ϵ_r and ϵ_a respectively quantify the relative and absolute tolerance on the time integration error introduced by neglecting the contribution of the fast modes. Hence, the fast modes are identified in $\lambda_i|_1^M$ whereas the slow modes are spanned by $\lambda_i|_{M+1}^N$ [45]. Not solving explicitly for the fast modes introduces the advantage of having a larger integration pace since $\tau_f >> \tau_1$ and no initialization overhead is required as opposed to multi-step implicit integration schemes. Nevertheless, the major drawback of the this methodology is the large computational cost associated with the calculation of the eigenvalues and eigenvectors of the jacobian matrix $[J]$, which scales approximately with N^3. In fact, given the non-linearity of the differential equation determined by $\boldsymbol{W}$, the eigen-basis rotates along the relaxation path thus requiring the recomputation of the basis at each state. This leads to the novelty of this work which is that of fitting the eigen-system onto the principal component space introduced in the next section.

E. Principal Component Analysis

In line with the objective of this work, which aims to develop a fitted mapping of the eigenvector space, PCA offers an efficient method for dimensionality reduction in large kinetic mechanisms. By projecting the detailed system onto a smaller, truncated basis that retains most of the original variance [16, 46, 47], PCA simplifies the representation of the system. Starting with the thermodynamic primitive variable basis $\boldsymbol{Q}$, PCA provides a linear transformation into a new basis defined by the principal components (PCs) $\boldsymbol{\eta}$. This reduction in the number of independent variables lowers the complexity of the regression needed to parametrize the eigenspace. Since PCA is a data-driven statistical technique aimed at identifying correlations among the state variables that linearize the data variability, a set of initial training data is required. In this work, the initial dataset $[X]$ is extracted from the numerical solution of the space-marching differential equation in Equation 7:

$$[X] = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \dots & T_{v1} & T_1 & p_1 & u_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{1_{n_s}} & \sigma_{2_{n_s}} & \dots & T_{v_{n_s}} & T_{n_s} & p_{n_s} & u_{n_s} \end{bmatrix} \tag{18}$$

where $[X]$ has dimensions $(n_s \times N)$, containing n_s samples, corresponding to the solution at a given spatial-step in the post-shock region of N state variables $\boldsymbol{Q}$. The PCA transformation then defines the PCs as [46]:

$$[\eta] = [A^T][S]\left([X] - [\bar{X}]\right) \tag{19}$$

where $[\bar{X}]$ is the mean value matrix $(n \times N)$ of each variable Q_i in its corresponding row, $[S]$ is a $(N \times N)$ diagonal matrix containing the scaling factors for each Q_i and $[A]$ are the orthonormal ($[A^{-1}] = [A^T]$) eigenvectors of the covariance matrix of the centered and scaled data $[\tilde{X}] = [S]\left([X] - [\bar{X}]\right)$. The least number of principal components required to retain most amount of variance is investigated using the four scaling criteria outlined in Table 1 as discussed in the next subsection. Further details on different scaling factors can be retrieved from [46].

Table 1 Scaling criteria for multivariate datasets

Scaling Criterion	Scaling Factor
Auto	$\sigma(X)$
Pareto	$\sqrt{\sigma(X)}$
Range	max(X)-min(X)
Vast	$\sigma(X)\frac{\sigma(X)}{mean(X)}$

While Equation 19 does not reduce the dimensionality of the chemical kinetics, it does transform the original basis to the PC basis such that the eigenvalues of $[\tilde{X}]$ indicate which eigenvector contain most of the information. By retaining N_η eigenvectors $[A_s]$ associated to the largest eigenvalues, the dimensionality of the system is reduced [16, 46]:

$$[\eta] \simeq [\eta_s] = [A_s^T][S]\left([X] - [\bar{X}]\right) \tag{20}$$

Equation 20 shows that $[\eta_s]$ contains the PC vectors of size N_η. For a strongly attractive manifold in state space it is found that $N_\eta << N$ [48]. It is highlighted that in this work the utilization of PCA is not advocated to reconstruct or transport the state variables downstream of the strong shock, but to identify a reduced number of independent variables sufficient to train the regression model, as discussed in the next subsection.

F. Regression Modelling of Eigen-basis

A non-linear regression function f_ϕ is employed to map the eigen-space, encompassing eigenvalues and right-eigenvectors, of N primitive variables onto a reduced basis N_η such that $\boldsymbol{\lambda}, \boldsymbol{R} = f_\phi(\boldsymbol{\eta})$. While the utilization of nonlinear regression methods is commonly used to map the source terms onto the PCs through a variety of techniques [49–54], the mapping of the eigenvector basis has only been recently addressed by Galassi et al. [27] using neural networks for the isobaric ignition of an H_2/air mixture.

In the present work, the mapping of the eigenvector basis is extended for the first time using GPR. This is a non-parametric kernel-based stochastic process that predicts the value of a response variable by modelling a multi-variate Gaussian [55]. The eigenvalues and eigenvectors obtained from the small perturbation analysis are hence fitted using models of the form shown in Equation 21, where B_i are the basis functions of the Gaussian processes and $\epsilon \sim N(0, \sigma_e^2)$, with σ_e^2 being the error variance. In this work, the *fitrgp* function in MATLAB is used to estimate the basis functions and noise variance. The utilization of PCA for reduced dimensionality is desirable in GRP for large amounts of covariant regressions to ease the computational cost and improve the accuracy of the predictions. Moreover because the model assumes zero mean, GPR is characterized by its kernel function, which greatly influences the model's behaviour. The Automatic Relevance Determination Squared exponential kernel is chosen as it automatically identifies the relevant features to map.

$$\lambda_i, R_i = B_i(\boldsymbol{\eta}) + \epsilon \tag{21}$$

The number of eigenvalues that need to be mapped can be identified from Equation 12 since the conserved laws associated to $\lambda_i = 0$ remain constant throughout the solution. Each eigenvalue is associated to an right-eigenvector. However, the right-eigenvectors associated to the zero eigenvalues are trivial invariant directions associated to the atomic conservation and linear combination of the conserved system variables and can be identified by retaining the invariant eigen-basis at any point in the solution. Therefore, one GPR model is needed for each n_λ non-zero eigenvalues along with a total of corresponding n_λ GPR models for the associated eigenvectors, each containing N elements. However, the size of the latter can be decreased to $N - 1$ by enforcing the normalization condition of the right-eigenvector. The left-eigenvectors can then be determined by taking the inverse of the right-eigenvector matrix according to their definition $[L] = [R]^{-1}$.

IV. Results

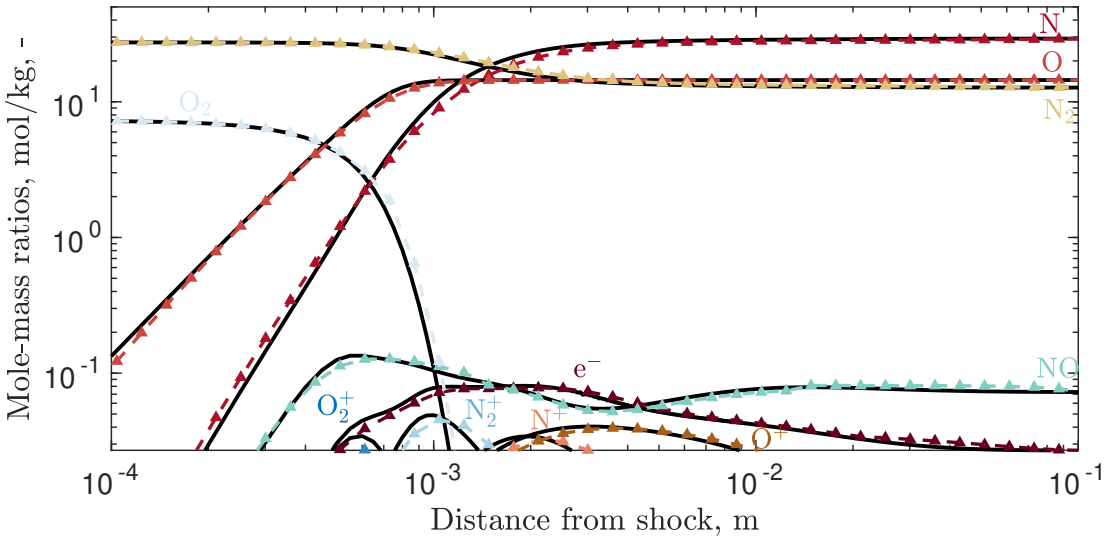
A. Verification of Finite Difference Method Solver

To study the relaxing flow behind a strong normal shock wave, the post-shock conditions in Table 2 are used as the initial conditions to solve Equation 7. The post-shock pressure, temperature and velocity have been determined using the Rankine-Hugoniot relations assuming frozen vibrational temperature and composition across the shock from the free-stream conditions corresponding to a re-entry flight at an altitude of 71 km using synthetic air [56].

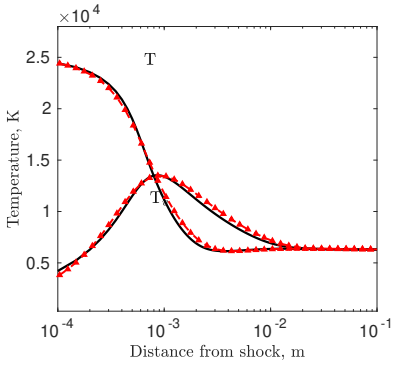
To verify the correct implementation of the FDM solver, the post-shock solution is compared to the 1D NESS algorithm, based on the parabolised Navier-Stokes equations [35], where the viscous contributions are disabled. Both codes adopt the same thermochemistry library within OCEAN which has been validated in previous work by Clarke et al. [35]. Minor discrepancies are noted in the immediate post-shock production of species due to the different integration schemes implemented in the two algorithm, which are propagated downstream of the strong shock. However, the evident agreement between the two codes is deemed sufficient to establish the correct trends captured by the present solver such that the FDM solution is used as a reference for the remainder of this work.

Table 2 Free-stream and frozen post-shock thermodynamic properties

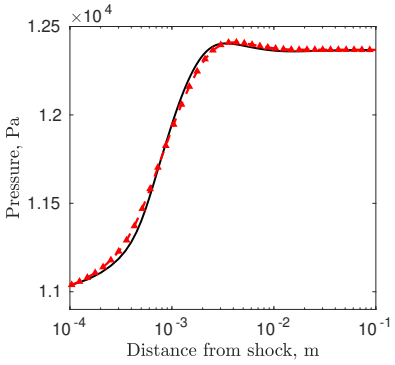
Flow State	Free-stream	Post-Shock
T [K]	205.0	25162.0
T_v [K]	205.0	205.0
p [Pa]	15.0	10960.0
u [m/s]	7198.0	1209.0



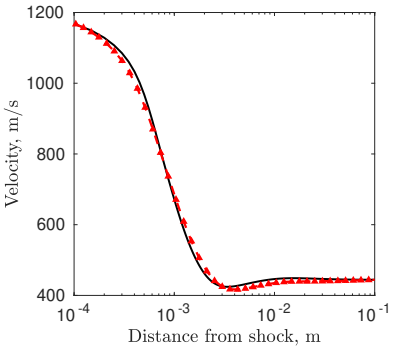
(a) Composition



(b) Temperatures



(c) Pressure



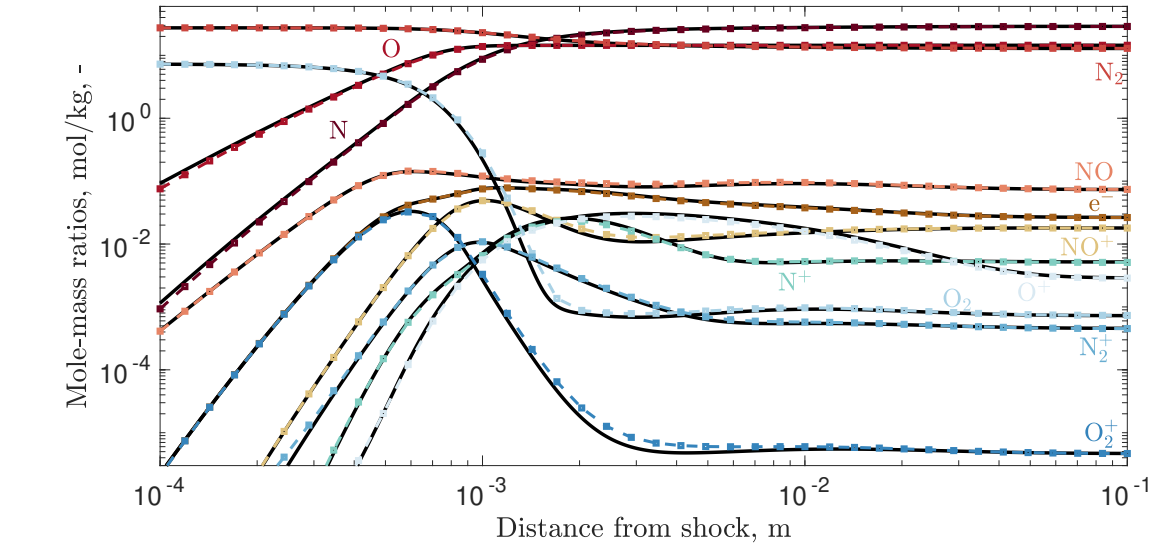
(d) Velocity

Fig. 3 Post-shock flow distribution
—, FDM; △- -△- -△, NESS

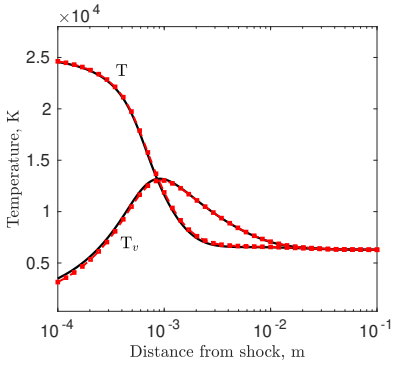
B. Small Perturbation Analysis and Eigen-basis solution

Having established a reference post-shock profile, the small-perturbation method proposed in this work is compared with the FDM solution in Figure 4. As expected, the approximate solution of the minor species slightly deviates from the exact one, as evident from the O_2^+ and O_2 distribution, but is overall found to be highly accurate by retaining all eigen-modes.

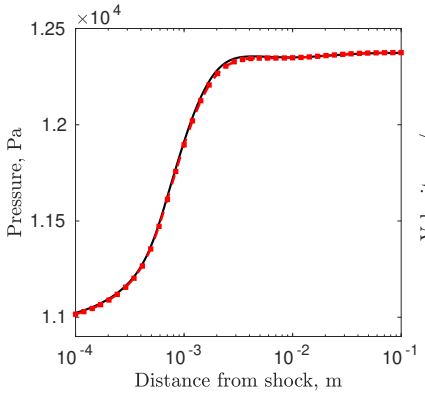
The small-perturbation method yields 9 non-zero eigenvalues out of a 15-equation system, which correspond to the thermochemical and vibrational relaxation processes. The remaining 6 zero eigenvalues correspond to 3 atomic conservation laws for O, N, and e^-, along with conservation of momentum, total energy and vibrational energy. The



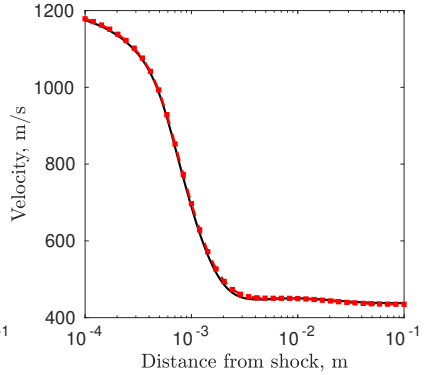
(a) Composition



(b) Temperatures



(c) Pressure



(d) Velocity

Fig. 4 Post-shock flow distribution
—, FDM; □- -□- -□, Small-perturbation Analysis

latter span the conserved subspace as previously shown in Equation 12. In fact, the eigenvectors associated to the zero-eigenvalues are invariant directions associated with the conservation laws. Since their complete basis must be conserved across the solution, the eigenvectors associated to the zero eigenvalues are evaluated at a single spatial location and maintained fixed across the iterations as to avoid rotations and swapping of the individual components within the complete basis. On the contrary, the 9 remaining non-zero directions span the tangent eigen-space representing variations in the composition mixture allowed within the conservation laws. The spatial evolution of the real component of the non-zero eigenvalues is shown in Figure 5 on a bi-symmetric logarithmic scale. The number of exhausted modes is also indicated to separate the effective number of slow and fast modes that may be retained based on the quasi-steady state assumption.

The most important observation is that the eigenvalues are continuous functions that vary smoothly. This is a necessary condition for the hybrid ROM framework proposed in this work since it shows that the evolution of the eigenvalues can be mapped or parametrized into a more convenient representation. Interestingly, two positive eigenvalues are observed immediately after the shock, as the vibrational-relaxation effects become significant, which merge into an eigenvalue with the same real component and a complex conjugate term. The existence of such explosive modes is well known in the field of combustion and is often attributable to auto-ignition and chain-branching reaction sequences [57, 58]. However, the study of these modes in hypersonic flows is mostly unexplored. In the present study, the two

positive eigenvalues, which then turn into dissipative modes, are not due to independent chemical processes but are in fact influenced by non-chemical processes such as vibrational relaxation. Mathematically, special care must be taken when treating complex eigenvalue-pairs since the numerical algorithm used to solve the eigenvalue-problem may swap the conjugate terms and lead to numerical noise in the associated eigenvector directions. A discussion is thus presented on the ordering of eigenvalues.

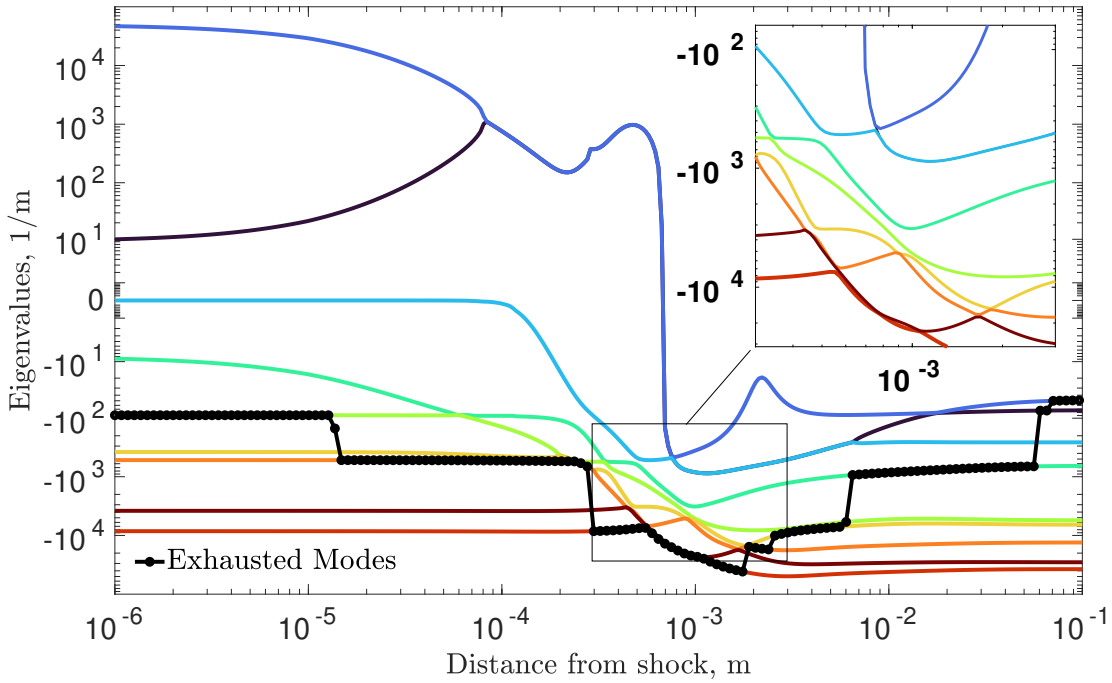
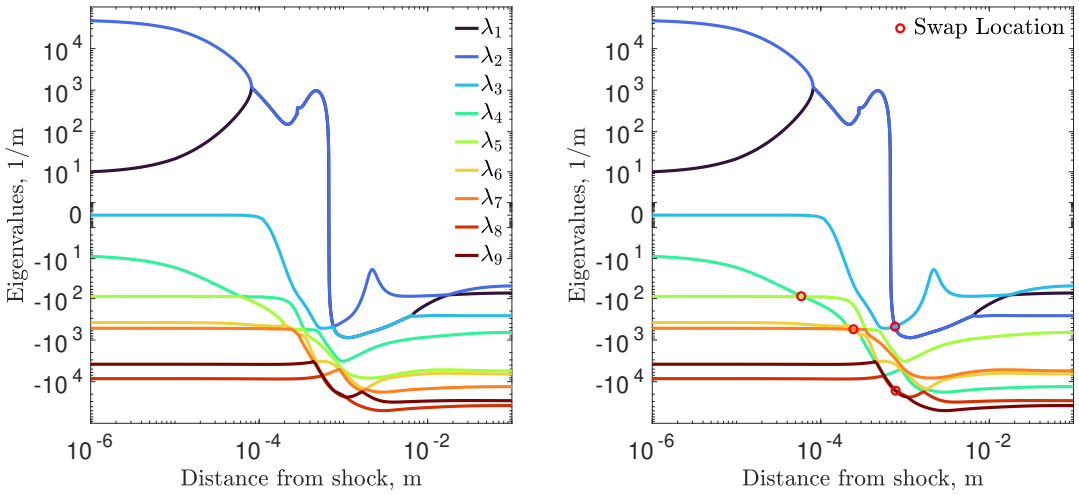


Fig. 5 Eigenvalue spectrum across the post-shock relaxation distance ○- -○- -○, Exhausted Modes for length-scale separation ($\epsilon_r = 1 \cdot 10^{-4}$, $\epsilon_a = 1 \cdot 10^{-6}$)

According to the ILDM [43] and CSP [23, 24] methodologies, the ordering strategy of the eigenvalues is done according to the decreasing magnitude of their modulus. However, the condition for continuous function aforementioned is not uniquely defined to a single order. The strong nonlinearities inherent to the system influence the magnitude, and thus the sorting, of the eigenvalues which may cause pairs of eigenvalues to appear to "switch order" over space. The identification of the original direction spanned by the eigenvalue as smoothly evolving quantities is carried out as a post-processing step once the complete spatial-history of the eigenvalues is available. This is shown in Figure 6, where the crossing locations of the eigenvalue-pairs are highlighted with red circles, demonstrating the correct sorting of the eigenvalues to preserve the original thermochemical processes.

The spatial evolution of the eigenvalues unveils a four phenomena which are relevant to providing insight into the fundamental thermochemical dynamics as the eigenvalues represent the relaxation length-scales of the nonequilibrium processes [27]:

1) **Merging**: two distinct real eigenvalues converge and coalesce into a single pair of complex conjugate eigenvalues. The transformation is a result to a variation in the stability of the system dynamics with an oscillatory behavior. A merging event is enlarged in Figure 7 (a) for λ_4 and λ_6.
2) **Bifurcation**: as opposed to the merging condition, this phenomenon occurs when a pair of complex conjugate eigenvalues split into two distinct real values. The imaginary part of the original eigenvalues vanishes and the dynamics return to a non-oscillatory behavior, reflecting a stabilization of the system. This is also illustrated in Figure 7 (a) for λ_4 and λ_6, following a merging event.
3) **Crossing**: two real eigenvalues exchange their relative magnitudes without merging into complex conjugates. The eigenvalues pass each other and retain their original value. When the eigenvalues are sorted by magnitude, this condition may lead to their incorrect ordering based on modulus as shown in Figure 6. An example of

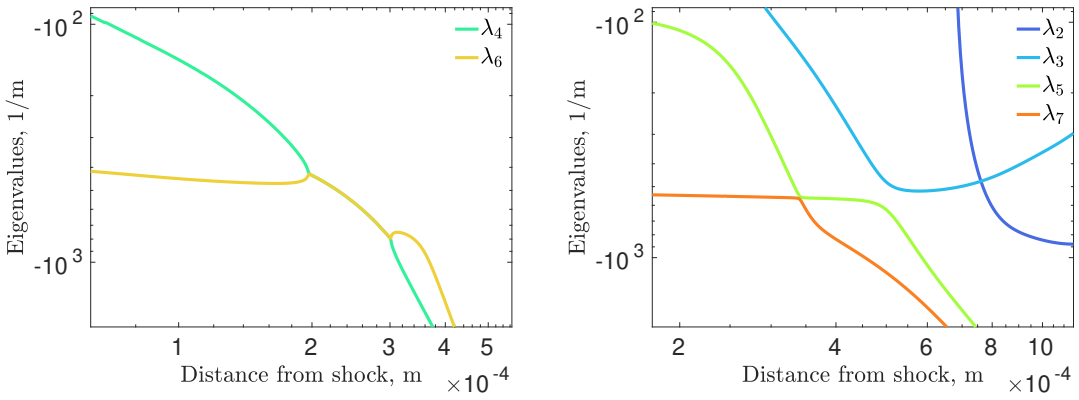


(a) Decreasing order **(b) Continuous order**

Fig. 6 Sorting strategy of eigenvalues as continuous and smooth thermochemical processes

crossing is provided in Figure 7 (b) for λ_2 and λ_3, after the appropriate order is established.

4) **Mis-crossing**: similarly to the crossing event, two eigenvalues approach each other closely but never merge or cross. After reaching a point of proximity, they diverge immediately. No swapping of eigenvalues ordering is thus caused by such events. For instance, λ_5 and λ_7 in Figure 7 (b) mis-cross each other. More proximity events are visible in the magnified region of Figure 5.



(a) Merging and Bifurcation **(b) Crossing and Mis-crossing**

Fig. 7 Observed Eigenvalue phenomena

To demonstrate the physical meaning of the eigenvectors, the eigen-objects of the three slowest thermochemical processes are plotted in Figure 8. It is noted that only the first 11 components are plotted, corresponding to the chemical species, such that the normalization arising from the thermodynamic variables does not influence the scaling of the figures. From the slowest eigenvalue λ_3 in the region spanned by $10^{-3} < x < 10^{-1}$ m, the two dominating components of the eigenvector along the chosen distance reveal the pseudo-relationship of type $N_2 \rightarrow N_2$, where the exact stoichiometric coefficients correspond to the value of the eigenvector at a given distance from the shock x, given that the scaling does not influence the overall reaction balance as long as the atomic mass of the reactants and products is conserved. The pseudo-reaction is in agreement with the relaxation of the flow distribution resulting into the dissociation of nitrogen molecules post-shock. The influence of the other minor species may not necessarily be negligible; for

instance, at $x = 10^{-2}$ m using the stoichiometric coefficients to 3 decimal places, the pseudo-reaction in R_3 is: 0.420 N_2 + 0.225 O + 0.038 NO + 0.103 NO^+ + 0.002 N_2^+ + 0.015 $N^+ \rightarrow N$ + 0.246 e^- + 0.368 O^+ which conserves charge and elemental mass on the two sides of the reactions. The discussion can be extended to the two other eigenvectors in Figure 8, where the effect of competing length-scales is more relevant. For instance, the pseudo-reaction at the same location for R_1 is given as 0.494 N_2 + 0.009 NO + 0.001 NO^+ + 0.002 N^+ + 0.002 $e^- \rightarrow N$ + 0.011 O which reveals that the linear combination of reactions in R_3 clearly corresponding to the ionization of atomic oxygen and dissociation of molecular nitrogen is "slower" than the pseudo-reaction in R_1, mainly representing dissociation of nitrogen and nitric oxide. These combinations are of course a function of distance and the scales given by the competing eigenvalues. Interestingly, for R_2 the presence of the mis-crossing at approximately $x = 0.0195$ m leads to a sharp change in the eigenvectors associated with a change in construction and destruction of species. This dynamic behaviour highlights a further numerical difficulty that is encountered when fitting the eigen-basis in the presence of such mis-crossing events.

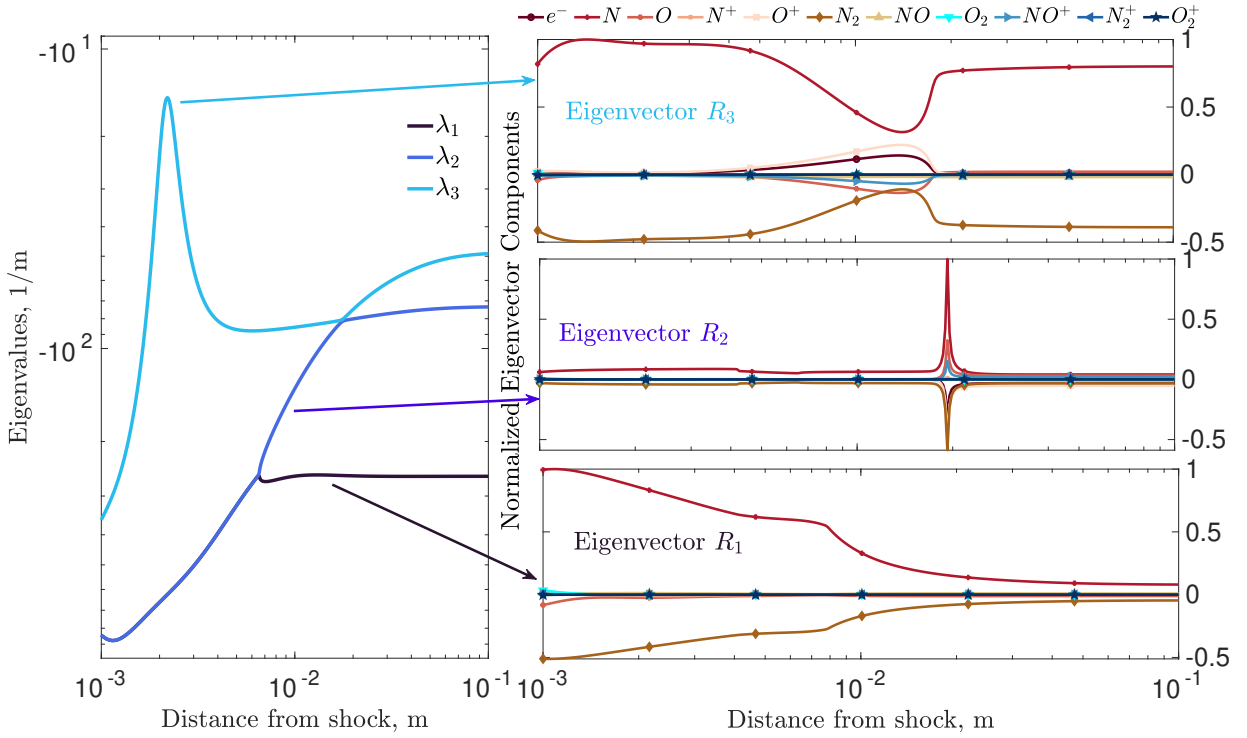


Fig. 8 Eigenvector dynamics for the three slowest thermochemical eigenvalues

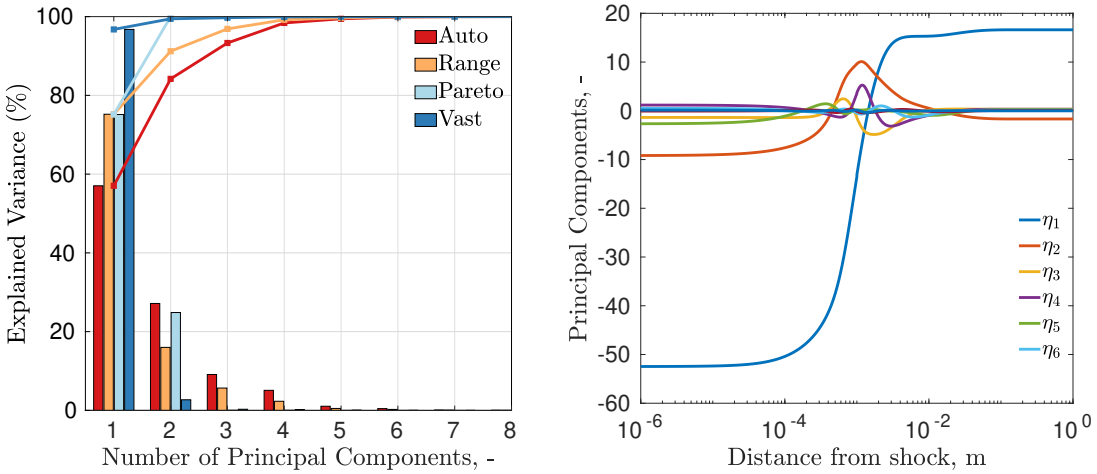
The primary focus of the proposed methodology lies in ensuring the smooth behavior of eigenvectors associated with the non-zero eigenvalues. Achieving a smooth evolution of these eigenvectors is essential for the regression fitting routine to converge more quickly and yield a more accurate solution. In contrast to the work by Galassi et al. [27], which uses a reduced Jacobian matrix limited to species conservation and incorporates temperature as a separate constraint to smooth the eigenvector behavior, this study retains the full Jacobian matrix, comprising 15 terms: 11 species conservation laws, momentum, global continuity, energy, and vibrational energy. Although the raw right-eigenvector data from LAPACK initially appears noisy, postprocessing ensures smoothness. Specifically, the sign of a normalization variable—often T, T_v, p or u due to the larger values—is used to set the eigenvector's direction. Additionally, by examining the angle variation between successive eigenvectors along the relaxation path, the solution is further refined to remove or minimize discontinuities in the eigenvector behaviour, resulting in the desired outcome. The eigenvectors are shown in Figure 11 along with the predicted fitted solution.

C. Principal Components

In this work, the advantage of the reduced dimensionality attainable through PCA is that the fitting routine of eigen-basis may be trained on a reduced number of independent variables, thus reducing computational cost and over-fitting with variables that carry negligible variance. In fact, the state vector $\boldsymbol{Q}$ is composed of 15 variables that

would need to be propagated along the post-shock relaxation path. On the contrary, PCA reduces the dimensionality of the original system from the 15 primitive variables to the truncated number of principal components. To identify the number of PCs that shall be retained in the truncated PCA-basis whilst maintaining high-fidelity, the explained and cumulative variance is plotted in Figure 9 with increasing number of PCs.

The explained variance in Figure 9 (a) shows that it is sufficient to retain two PCs using the Vast scaling method whilst achieving a cumulative variance greater than 99%. This signifies a reduction of 86.6% in variables used to train the regression fitting model. The Pareto criterion would also lead to a similar conclusion, while range and auto would require between 4 and 5 PCs to capture the desired level of variance. Vast is thus chosen as the scaling criterion, though choosing Pareto would also be justifiable. The distribution of PCs across the post-shock distance is shown in Figure 9 (b) using the Vast scaling method. The selection of two PCs is even more clear as η_1 and η_2 are significantly larger than all other PCs throughout the relaxation. Hence, the two PCs η_1 and η_2 are selected using the Vast scaling method as independent variables to train the eigen-basis fitting method.



(a) Explained and Cumulative (solid lines) variance **(b) Spatial distribution of PCs using Vast scaling criterion**

Fig. 9 Selection strategy for number of principal components and scaling criterion

D. Reconstruction of Eigen-basis using PCA-GPR

The eigen-basis obtained through the small-perturbation method is used to train the GPR models. Specifically, the eigenvalue profiles in Figure 5 and associated eigenvectors are utilized as inputs to the training routine. In the *fitrgp* routine implemented, the entire data set is by default set for training. Cross-validation data is recommended for future implementations aimed at interpolating the eigen-basis between different trajectory points or initial conditions.

To assess the accuracy of the reconstructed eigen-basis, the Root Mean Square Error (RMSE) is utilized. Table 3 presents the RMSE for each eigenvalue and the corresponding components of the eigenvectors. It is important to note that zero-eigenvalues and their associated eigenvectors do not need to be predicted, as they represent invariant subspaces that remain constant throughout the simulation. The RMSE values for each eigenvector component are also plotted in Figure 10, revealing notable variability across the components. Some eigenvectors exhibit consistently low RMSE values, such as R_1 which in this case is more desirable as this eigenvector corresponds to one of the explosive modes in the post-shock region. In contrast, components associated with chemical species typically show smaller RMSE values compared to the thermodynamic variables (T_v, T, p, u). This discrepancy arises because thermodynamic variables exhibit more abrupt changes, particularly in regions where eigen-modes merge or cross. Furthermore, components linked to dominant species such as N_2 and O_2 generally have lower RMSE, reflecting a more accurate reconstruction. Conversely, components corresponding to less dominant species like NO^+, N_2^+, and O_2^+ tend to show higher RMSE, indicating greater uncertainty. This aligns with the PCA truncation, as these less dominant species contribute less information and have a smaller impact on the overall fidelity of the model, making them more challenging to capture accurately. Overall, RMSE values in the range of 10^{-2} to 10^{-6} are obtained, demonstrating a strong correlation between the training data and the reconstructed data using predictive estimates from the basis functions of the form $f_\phi(\eta_1, \eta_2)$.

Table 3 RMSE For Eigenvalues and Eigenvectors

Component i	λ_i	R_1^i	R_2^i	R_3^i	R_4^i	R_5^i	R_6^i	R_7^i	R_8^i	R_9^i
1	5.898×10^{0}	7.200×10^{-6}	3.343×10^{-4}	6.114×10^{-5}	1.685×10^{-4}	6.046×10^{-5}	1.081×10^{-4}	4.945×10^{-5}	1.736×10^{-4}	1.386×10^{-4}
2	1.578×10^{1}	1.514×10^{-4}	3.058×10^{-4}	9.476×10^{-5}	3.708×10^{-4}	1.448×10^{-4}	2.625×10^{-4}	9.684×10^{-5}	2.787×10^{-4}	1.482×10^{-4}
3	3.243×10^{-1}	7.420×10^{-5}	3.829×10^{-4}	6.085×10^{-5}	6.429×10^{-4}	1.003×10^{-4}	4.794×10^{-4}	1.107×10^{-4}	2.245×10^{-4}	1.125×10^{-4}
4	1.145×10^{1}	3.065×10^{-6}	4.588×10^{-5}	5.119×10^{-6}	1.396×10^{-4}	6.691×10^{-5}	3.598×10^{-5}	2.788×10^{-6}	3.223×10^{-5}	1.047×10^{-4}
5	2.119×10^{0}	3.448×10^{-6}	3.697×10^{-4}	6.078×10^{-5}	1.474×10^{-5}	8.576×10^{-5}	3.713×10^{-6}	1.233×10^{-4}	1.934×10^{-5}	9.979×10^{-5}
6	6.022×10^{0}	1.151×10^{-4}	3.352×10^{-4}	7.315×10^{-5}	3.768×10^{-4}	2.512×10^{-4}	2.453×10^{-4}	1.261×10^{-4}	1.563×10^{-4}	1.123×10^{-4}
7	3.924×10^{0}	1.071×10^{-4}	7.413×10^{-5}	1.009×10^{-4}	3.605×10^{-4}	1.929×10^{-4}	2.793×10^{-4}	1.053×10^{-4}	2.435×10^{-4}	1.325×10^{-4}
8	2.064×10^{1}	1.015×10^{-4}	9.244×10^{-5}	8.607×10^{-5}	2.324×10^{-4}	2.245×10^{-4}	1.900×10^{-4}	1.184×10^{-4}	1.893×10^{-4}	1.093×10^{-4}
9	1.538×10^{1}	3.681×10^{-6}	1.799×10^{-4}	3.208×10^{-5}	1.295×10^{-4}	4.924×10^{-5}	6.516×10^{-5}	2.368×10^{-5}	1.311×10^{-4}	1.097×10^{-4}
10	-	1.194×10^{-6}	4.061×10^{-6}	6.441×10^{-5}	1.506×10^{-4}	2.442×10^{-4}	1.233×10^{-4}	1.315×10^{-4}	6.270×10^{-5}	1.299×10^{-4}
11	-	5.642×10^{-6}	8.013×10^{-6}	5.008×10^{-5}	1.309×10^{-4}	2.501×10^{-4}	1.112×10^{-4}	6.126×10^{-5}	1.768×10^{-4}	1.032×10^{-4}
12	-	1.055×10^{-2}	4.126×10^{-3}	7.166×10^{-4}	6.525×10^{-3}	2.033×10^{-4}	1.216×10^{-4}	1.049×10^{-3}	1.027×10^{-2}	6.663×10^{-4}
13	-	1.710×10^{-3}	1.828×10^{-3}	6.511×10^{-4}	6.311×10^{-3}	1.565×10^{-2}	2.351×10^{-2}	1.442×10^{-3}	4.987×10^{-3}	7.226×10^{-4}
14	-	7.647×10^{-4}	2.138×10^{-4}	1.500×10^{-4}	5.523×10^{-3}	5.838×10^{-3}	2.786×10^{-3}	4.686×10^{-4}	1.439×10^{-3}	9.003×10^{-5}
15	-	4.593×10^{-4}	2.927×10^{-4}	1.056×10^{-4}	9.082×10^{-4}	3.292×10^{-3}	1.490×10^{-3}	3.945×10^{-4}	6.657×10^{-4}	9.906×10^{-5}

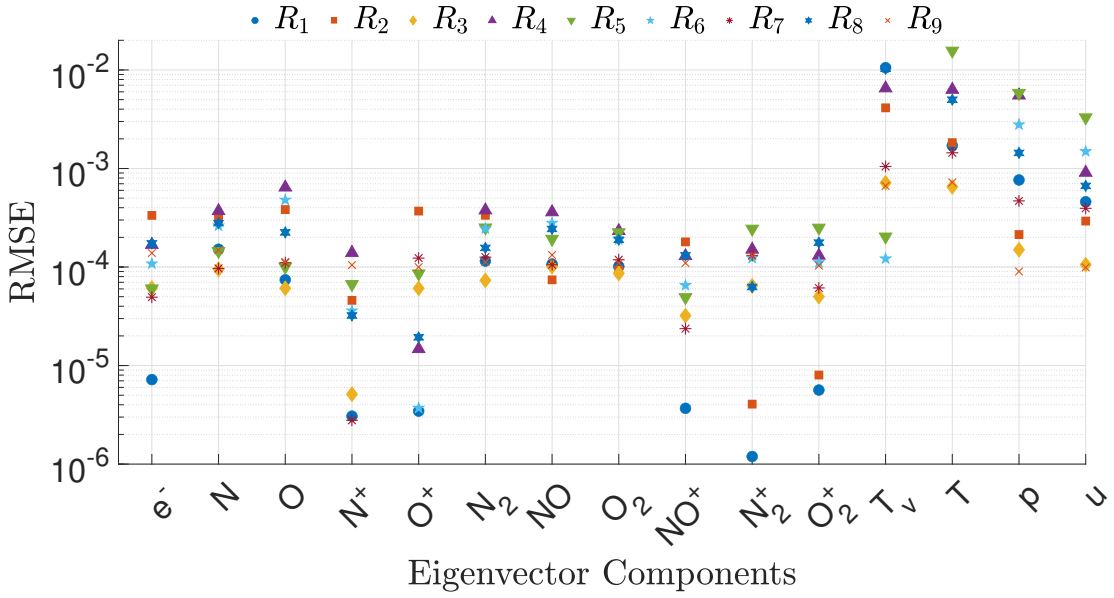
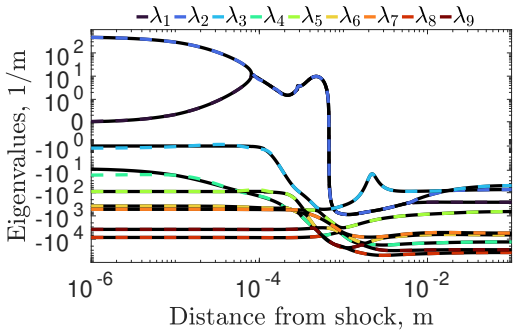


Fig. 10 RMSE for Reconstructed Eigenvectors

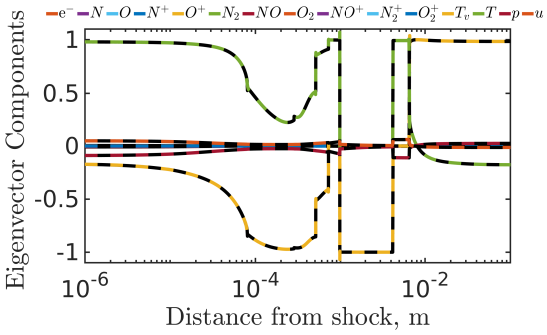
The training data, corresponding to the correct solution as computed by solving the eigenvalue problem in LAPACK, and the predicted eigenvalues and eigenvectors are shown in Figure 11. It is clear that the reconstruction is very accurate and that the presence of minor errors is only noticeable in the presence of abrupt eigenvector changes in the dominating variables, such as pressure and temperature. These are also in agreement with the higher RMSE found for these variables, attributable to the sharp variations coinciding with the eigenvalue phenomena previously discussed. The presence of merging, splitting, crossing and mis-crossing events is therefore undesirable for the eigenvectors reconstruction.

E. Joint PCA-GPR and Small-Perturbation Analysis ROM

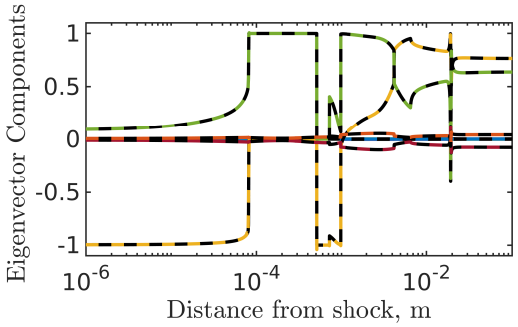
The numerical reconstruction of the solution under the conditions specified in Table 2 is depicted in Figure 12 using the GRP models. This represents the best-case scenario since the surrogate model is reconstructing the solution that was previously used in the training. Notable inconsistencies in the minor species are observed, highlighting the solution's sensitivity to the eigenvectors. The immediate post-shock region also presents some discrepancies in the production of nitric oxide and the charged species. This sensitivity arises because each eigenvector represents a linear combination of reactions, meaning that even a slight error in one eigenvector affects the combined production or destruction of species. This underscores the critical importance of accurately reconstructing the eigen-basis.



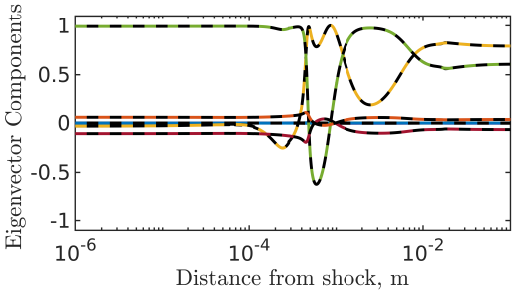
(a) Eigenvalues



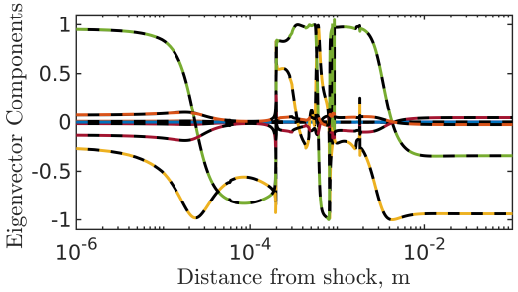
(b) Eigenvector 1



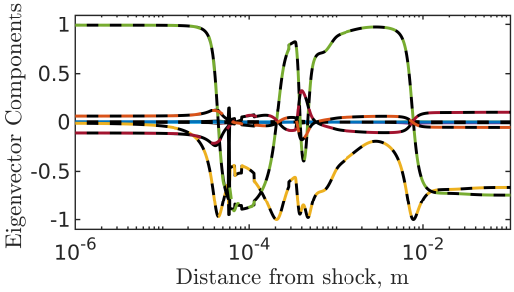
(c) Eigenvector 2



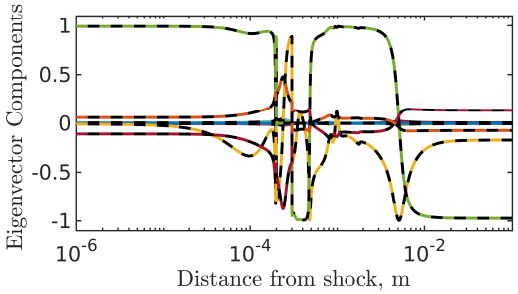
(d) Eigenvector 3



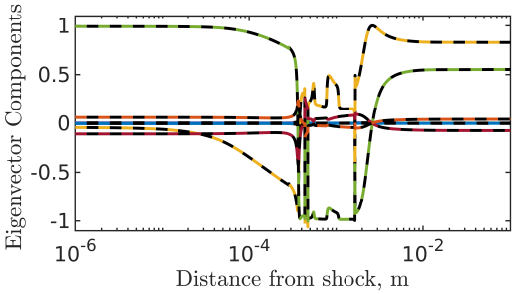
(e) Eigenvector 4



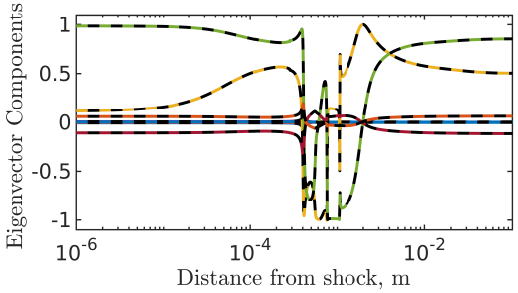
(f) Eigenvector 5



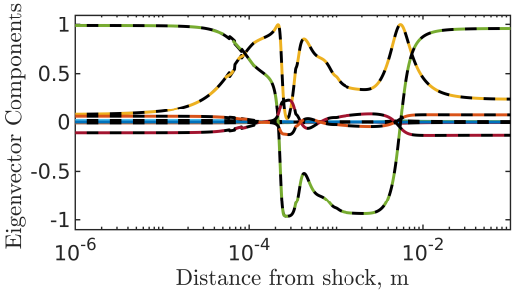
(g) Eigenvector 6



(h) Eigenvector 7



(i) Eigenvector 8

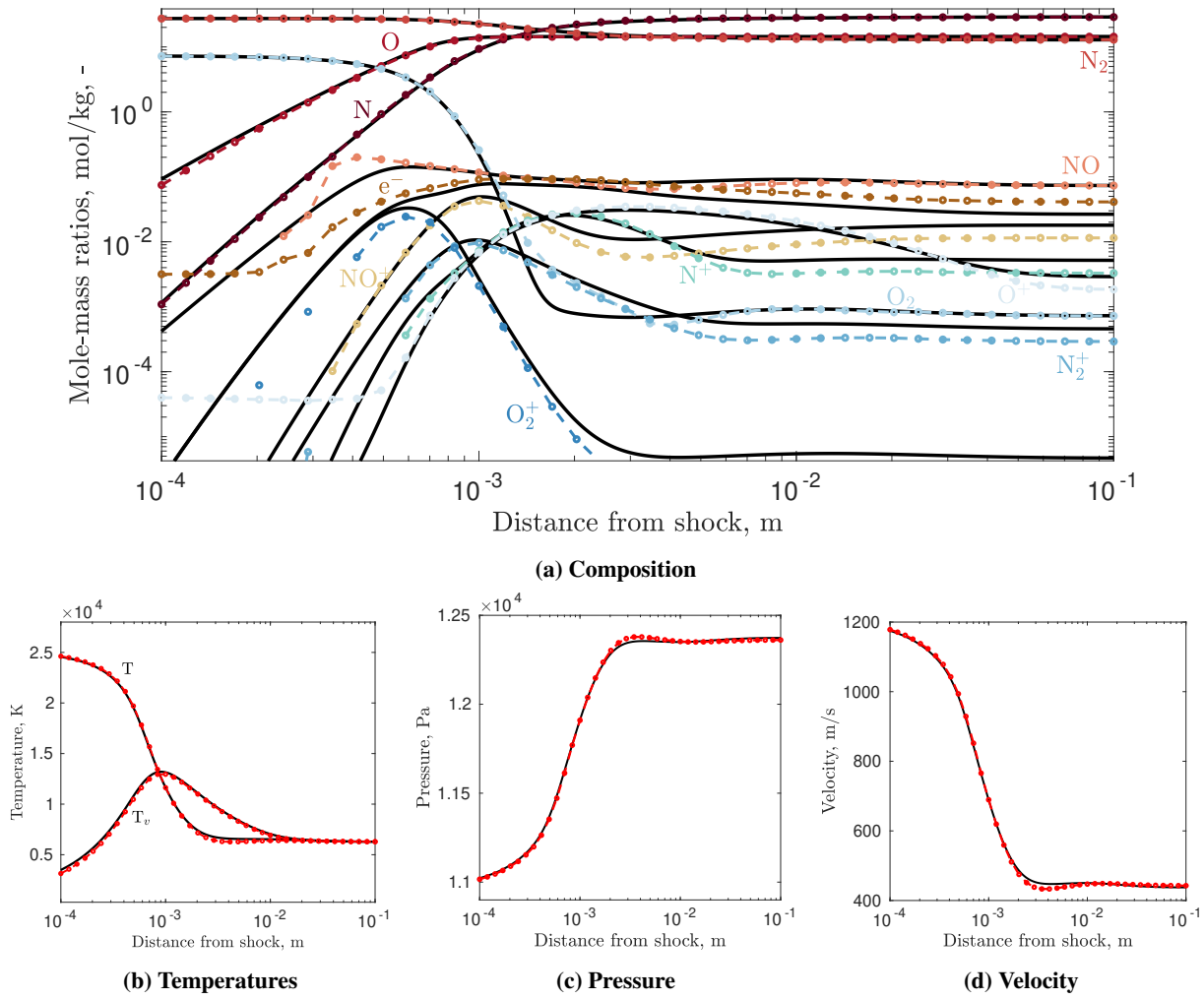


(j) Eigenvector 9

Fig. 11 Eigen-objects of Jacobian matrix decomposition
—, Small Perturbation Method; - - -, Reconstructed PCA-GPR

Additional errors in the present data can be attributed to the PCA, which retained only 99% of the variance using the first two principal components (PCs). Since minor species contribute minimally to the total variance, they are likely not captured by the first two PCs. Including more PCs during the training stage would be necessary to account for such variations. However, this approach would increase the training cost as more independent variables are used.

Despite these challenges, the macro-properties of the flow—such as temperature, pressure, and velocity—are predicted with high accuracy, showing only minor deviations due to error propagation in the GRP models.



(a) Composition

(b) Temperatures **(c) Pressure** **(d) Velocity**

Fig. 12 Post-shock flow distribution
—, FDM; ○- -○- -○, Reconstructed PCA-GPR Surrogate

A final consideration of the computational cost of the proposed framework is necessary since this paper aims to identify a local-basis surrogate that enhances computational efficiency. The computational cost for lower-upper (LU) factorization and eigenvalue computation scales with $O(N^3)$ while the resolution of linear equations increases with $O(N^2)$ as shown in Figure 13. The FDM method solves linear systems using LU-decomposed matrices, which involves matrix-vector operations and relies on routines such as DGETRF and DGETRS. While the computational cost is lower, it requires a smaller integration step due to the lack of knowledge regarding time scales, which may lead to more iterations. The small-perturbation method, on the other hand, requires eigen-decomposition to reduce numerical stiffness, but this step incurs a substantially higher computational cost compared to the other operations. This observation supports the aim of this work, which is to develop surrogate models for eigen-decomposition.

A direct comparison of the computational cost with the surrogate PCA-GRP models is not presented as it would

be improper, as the GRP model is trained in MATLAB while the main code runs in FORTRAN. Additionally, the cost of the GRP model is highly dependent on the size of the training data. It is important to note that the fitted basis does not provide a clear computational advantage in the current setup given the simplicity of the two-temperature 0D problem solved. Nonetheless, while the computational cost of evaluating eigen-objects scales with the cube of the Jacobian matrix size, the surrogate model's cost scales with the cube of the training data but increases only linearly with the number of eigenvectors to be predicted, each containing N variables. Therefore, a significant computational benefit is expected when the surrogate model is applied to larger kinetic mechanisms. It is also important to emphasize that the number of eigen-objects required is always smaller than the number of equations due to the conservation laws that yield invariant eigenvectors. Moreover, leveraging the main strength of time-scale separation enabled by the small-perturbation method, the criterion in Equation 17 can be used to identify which eigenvectors need to be reconstructed at a given iteration. While directly solving the eigen-system for the desired number of "slow" eigen-objects may not be feasible, the reconstruction of these eigen-objects through predictive fitting models is easily attainable. This approach would further reduce computational time by a factor linearly proportional to the number of excluded modes, thus improving efficiency.

V. Conclusion

This work was motivated by the increasing need for high-fidelity and computationally efficient numerical modelling of nonlinear and stiff thermochemical kinetics, particularly in the context of nonequilibrium hypersonic flow simulations. By exploiting the advantages of time-scale separation in thermochemical nonequilibrium processes and optimizing time-step selection for explicit integration schemes, eigen-decomposition of the Jacobian matrix emerged as a compelling approach for solving the governing ODEs. A small-perturbation method inspired by the ILDM and CSP frameworks was developed to address the generalized eigenvalue problem for the reactive Euler equations. The proposed method involves pre-multiplying the source term Jacobian by the inverse Jacobian of the inviscid fluxes, enabling the separation of spatial scales for flow convection and chemical processes without imposing assumptions on the conservation of pressure or volume. The zero-eigenvalues characterize invariant subspaces associated with conservation laws, while the nonzero eigenvalues define the spatial or temporal scales of distinct physical processes. The associated eigenvectors, consisting of linear combinations of stoichiometric coefficients and thermochemical relaxation, describe the characteristic directions of these processes. Post-shock dynamics were shown to exhibit complex phenomena such as eigenvalue splitting, merging, and crossing, which present challenges for eigenvector analysis and post-processing ordering.

Despite the physical insights provided by eigen-decomposition, the computational cost of directly calculating eigenvalues and eigenvectors is significant where operations such as LU factorization are more efficient, especially for large systems. To address this limitation, a novel ROM framework was proposed. This framework employs the GPR method to predict the eigen-basis at interpolated conditions based on training data. Dimensionality reduction using PCA was introduced, achieving an 86% reduction in the number of independent variables required for training, thereby minimizing training cost. The methodology was applied to a 2-temperature, 11-species air mixture relaxing downstream of a strong shock at a freestream velocity of 7.198 km/s. The surrogate eigenvector models achieved RMSE figures in the range of 10^{-2} to 10^{-6} and accurately reproduced the reference FDM solution, demonstrating the feasibility of replacing the computationally expensive local eigen-basis calculation in high-speed flow simulations.

While the computational benefits of the framework are not clear for the simple post-shock relaxation scenario presented, its scalability makes it particularly advantageous for problems involving large kinetic mechanisms, such as state-to-state collisional models, where eigen-basis computation scales with $O(N^3)$ and can dominate overall CPU time. Future work should focus on expanding the training datasets to encompass a broader range of initial conditions, including variations in Mach number, flow composition, and density. This would enhance the benefit of the surrogate models as they may be used to interpolate between different sets of training data. However, a balance must be struck between model generalization and computational efficiency, as increasing the size of the training dataset may elevate prediction costs. Tailored training datasets for specific applications are therefore recommended to optimize performance.

A major challenge identified in this work is the accuracy of basis reconstruction, which depends on the ability to resolve eigenvalue phenomena such as merging, bifurcations, and sharp discontinuities in eigenvectors. These challenges are expected to intensify for larger kinetic mechanisms due to increased spectral density. Nevertheless, this study demonstrates the potential of local-basis surrogates, enhanced by advanced computational learning techniques, for efficient modelling of nonequilibrium flows. As the demand for accurate and efficient hypersonic simulations continues to grow, integrating machine learning methods with highly parallelized GPU architectures offers a promising avenue for accelerating both training and reconstruction processes, paving the way for scalable and accurate predictive modelling.

Appendix

A. Detailed Kinetics Model

Table 4 Chemical Reactions for Air from Gupta et al. [36]

Reaction Number	Reactants		Products	Third Body M
1 – 6	$N_2 + M$	$\rightleftharpoons$	$N + N + M$	N_2, O_2, NO, N_2+, O_2+, NO+
7 – 10	$N_2 + M$	$\rightleftharpoons$	$N + N + M$	N, O, N+, O+
11	$N_2 + e^-$	$\rightleftharpoons$	$N + N + e^-$	
12 – 17	$O_2 + M$	$\rightleftharpoons$	$O + O + M$	N_2, O_2, NO, N_2+, O_2+, NO+
18 – 21	$O_2 + M$	$\rightleftharpoons$	$O + O + M$	N O, N+, O+
22 – 26	$NO + M$	$\rightleftharpoons$	$N + O + M$	N_2 O_2, N_2+, O_2+, NO+
27 – 31	$NO + M$	$\rightleftharpoons$	$N + O + M$	NO, N, O, N+, O+
32	$NO + O$	$\rightleftharpoons$	$N + O_2$	
33	$N_2 + O$	$\rightleftharpoons$	$NO + N$	
34	$N + N$	$\rightleftharpoons$	$N_2^+ + e^-$	
35	$O + O$	$\rightleftharpoons$	$O_2^+ + e^-$	
36	$N + O$	$\rightleftharpoons$	$NO^+ + e^-$	
37	$N + e^-$	$\rightleftharpoons$	$N^+ + e^- + e^-$	
38	$O + e^-$	$\rightleftharpoons$	$O^+ + e^- + e^-$	
39	$NO^+ + O$	$\rightleftharpoons$	$N^+ + O_2$	
40	$O_2^+ + N$	$\rightleftharpoons$	$N^+ + O_2$	
41	$O^+ + NO$	$\rightleftharpoons$	$N^+ + O_2$	
42	$O_2^+ + N_2$	$\rightleftharpoons$	$N_2^+ + O_2$	
43	$O_2^+ + O$	$\rightleftharpoons$	$O^+ + O_2$	
44	$NO^+ + N$	$\rightleftharpoons$	$O^+ + N_2$	
45	$NO^+ + O_2$	$\rightleftharpoons$	$O_2^+ + NO$	
46	$NO^+ + O$	$\rightleftharpoons$	$O_2^+ + N$	
47	$O^+ + N_2$	$\rightleftharpoons$	$N_2^+ + O$	
48	$NO^+ + N$	$\rightleftharpoons$	$N_2^+ + O$	
49	$N_2 + N^+$	$\rightleftharpoons$	$N_2^+ + N$	

B. CPU Cost scaling of FORTRAN operations

The computational time per operation for the FORTRAN subroutines DGETRF, DGEEV, and DGETRS was calculated using a single 11th Gen Intel® Core™ i7-11800H @ 2.30GHz CPU.

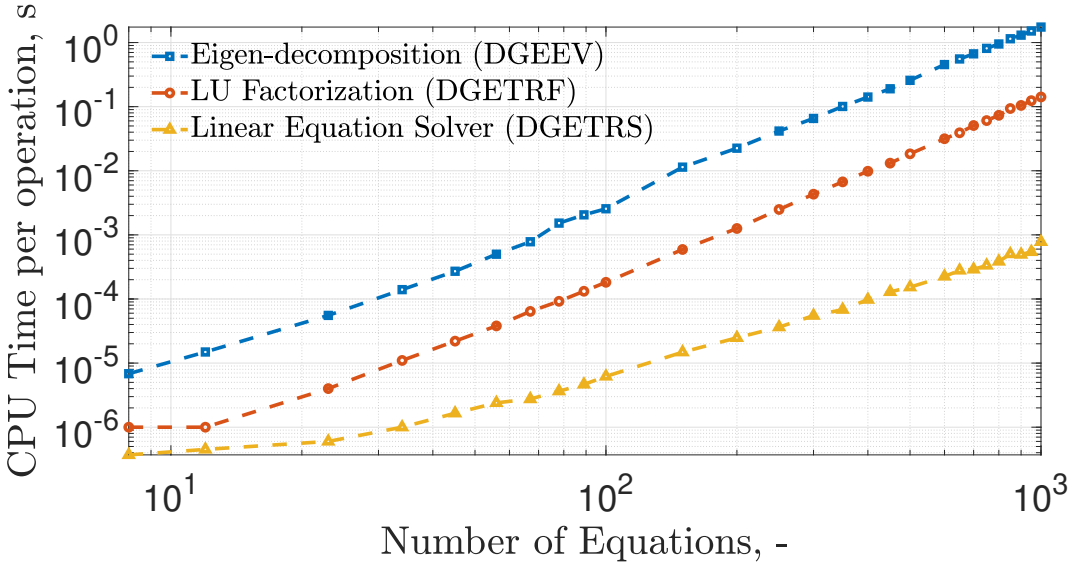


Fig. 13 CPU time scaling of eigen-decomposition, factorization and linear equation solver in FORTRAN

C. Jacobian Matrices

$$\frac{\partial U}{\partial Q} = \begin{bmatrix} \rho + \frac{\partial \rho}{\partial \sigma_1}\sigma_1 & \dots & \frac{\partial \rho}{\partial \sigma_n}\sigma_1 & \frac{\partial \rho}{\partial T_e}\sigma_1 & \frac{\partial \rho}{\partial T}\sigma_1 & \frac{\partial \rho}{\partial P}\sigma_1 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \rho}{\partial \sigma_1}\sigma_n & \dots & \rho + \frac{\partial \rho}{\partial \sigma_n}\sigma_n & \frac{\partial \rho}{\partial T_e}\sigma_n & \frac{\partial \rho}{\partial T}\sigma_n & \frac{\partial \rho}{\partial P}\sigma_n & 0 \\ \frac{\partial \rho}{\partial \sigma_1}e_v + \frac{\partial e_v}{\partial \sigma_1}\rho & \dots & \frac{\partial \rho}{\partial \sigma_n}e_v & \frac{\partial \rho}{\partial T_v}e_v + \frac{\partial e_v}{\partial T_v}\rho & \frac{\partial \rho}{\partial T}e_v & \frac{\partial \rho}{\partial P}e_v & 0 \\ \frac{\partial \rho}{\partial \sigma_1}e^0 + \frac{\partial e}{\partial \sigma_1}\rho & \dots & \frac{\partial \rho}{\partial \sigma_n}e^0 + \frac{\partial e}{\partial \sigma_n}\rho & \frac{\partial \rho}{\partial T_v}e^0 + \frac{\partial e}{\partial T_v}\rho & \frac{\partial \rho}{\partial T}e^0 + \frac{\partial e}{\partial T}\rho & \frac{\partial \rho}{\partial P}e^0 & \rho u \\ \frac{\partial \rho}{\partial \sigma_1} & \dots & \frac{\partial \rho}{\partial \sigma_n} & \frac{\partial \rho}{\partial T_v} & \frac{\partial \rho}{\partial T} & \frac{\partial \rho}{\partial P} & 0 \\ \frac{\partial \rho}{\partial \sigma_1}u & \dots & \frac{\partial \rho}{\partial \sigma_n}u & \frac{\partial \rho}{\partial T_v}u & \frac{\partial \rho}{\partial T}u & \frac{\partial \rho}{\partial P}u & \rho \end{bmatrix}$$

$$\frac{\partial F}{\partial Q} = \begin{bmatrix} \left(\rho + \frac{\partial \rho}{\partial \sigma_1}\sigma_1\right)u & \dots & \frac{\partial \rho}{\partial \sigma_n}\sigma_1 u & \frac{\partial \rho}{\partial T_v}\sigma_1 u & \frac{\partial \rho}{\partial T}\sigma_1 u & \frac{\partial \rho}{\partial P}\sigma_1 u & \rho\sigma_1 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \rho}{\partial \sigma_1}\sigma_n u & \dots & \left(\rho + \frac{\partial \rho}{\partial \sigma_n}\sigma_n\right)u & \frac{\partial \rho}{\partial T_v}\sigma_n u & \frac{\partial \rho}{\partial T}\sigma_n u & \frac{\partial \rho}{\partial P}\sigma_n u & \rho\sigma_n \\ \left(\frac{\partial \rho}{\partial \sigma_1}e_v + \frac{\partial e_v}{\partial \sigma_1}\rho\right)u & \dots & \frac{\partial \rho}{\partial \sigma_n}e_v u & \left(\frac{\partial \rho}{\partial T_v}e_v + \frac{\partial e_v}{\partial T_v}\rho\right)u & \frac{\partial \rho}{\partial T}e_v u & \frac{\partial \rho}{\partial P}e_v u & \rho e_v \\ \left(\frac{\partial \rho}{\partial \sigma_1}h^0 + \frac{\partial e}{\partial \sigma_1}\rho\right)u & \dots & \left(\frac{\partial \rho}{\partial \sigma_n}h^0 + \frac{\partial e}{\partial \sigma_n}\rho\right)u & \left(\frac{\partial \rho}{\partial T_v}h^0 + \frac{\partial e}{\partial T_v}\rho\right)u & \left(\frac{\partial \rho}{\partial T}h^0 + \frac{\partial e}{\partial T}\rho\right)u & \frac{\partial \rho}{\partial P}h^0 u & \rho h^0 + \rho u^2 \\ \frac{\partial \rho}{\partial \sigma_1}u & \dots & \frac{\partial \rho}{\partial \sigma_n}u & \frac{\partial \rho}{\partial T_v}u & \frac{\partial \rho}{\partial T}u & \frac{\partial \rho}{\partial P}u & \rho \\ \frac{\partial \rho}{\partial \sigma_1}u^2 & \dots & \frac{\partial \rho}{\partial \sigma_n}u^2 & \frac{\partial \rho}{\partial T_v}u^2 & \frac{\partial \rho}{\partial T}u^2 & \frac{\partial \rho}{\partial P}u^2 + 1 & 2\rho u \end{bmatrix}$$

References

[1] Sharma, S., and Park, C., "A survey of simulation and diagnostic techniques for hypersonic nonequilibrium flows," *25th AIAA Aerospace Sciences Meeting*, 1987. https://doi.org/10.2514/6.1987-406.

[2] Park, C., "Review of chemical-kinetic problems of future NASA missions. I - earth entries," *Journal of Thermophysics and Heat Transfer*, Vol. 7, No. 3, 1993, p. 385–398. https://doi.org/10.2514/3.431.

[3] Candler, G. V., Subbareddy, P. K., and Brock, J. M., "Advances in computational fluid dynamics methods for hypersonic flows," *Journal of Spacecraft and Rockets*, Vol. 52, No. 1, 2015, pp. 17–28.

[4] Chaudhry, R. S., and Candler, G. V., "Statistical analyses of quasiclassical trajectory data for air dissociation," *AIAA Scitech 2019 Forum*, 2019, p. 0789.

[5] Chaudhry, R. S., Boyd, I. D., Torres, E., Schwartzentruber, T. E., and Candler, G. V., "Implementation of a chemical kinetics model for hypersonic flows in air for high-performance CFD," *AIAA Scitech 2020 Forum*, 2020, p. 2191.

[6] Munafò, A., Liu, Y., and Panesi, M., "Modeling of dissociation and energy transfer in shock-heated nitrogen flows," *Physics of Fluids*, Vol. 27, No. 12, 2015.

[7] Wang, X., Hong, Q., Hu, Y., and Sun, Q., "On the accuracy of two-temperature models for hypersonic nonequilibrium flow," *Acta Mechanica Sinica*, Vol. 39, No. 2, 2023, p. 122193.

[8] Dubuet, U., Pannier, E., and Laux, C., "Uncertainties in multi-temperature nonequilibrium partition functions and application to CO2," *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 290, 2022, p. 108314.

[9] Oran, E. S., and Boris, J. P., "Numerical approaches to combustion modeling," 1991.

[10] Poinsot, T., "Theoretical and Numerical Combustion," *RT Edwards*, 2005.

[11] Raman, V., and Hassanaly, M., "Emerging trends in numerical simulations of combustion systems," *Proceedings of the Combustion Institute*, Vol. 37, No. 2, 2019, pp. 2073–2089.

[12] Frassoldati, A., Faravelli, T., and Ranzi, E., "Kinetic modeling of the interactions between NO and hydrocarbons at high temperature," *Combustion and Flame*, Vol. 135, No. 1-2, 2003, pp. 97–112.

[13] Lanser, D., and Verwer, J. G., "Analysis of operator splitting for advection–diffusion–reaction problems from air pollution modelling," *Journal of computational and applied mathematics*, Vol. 111, No. 1-2, 1999, pp. 201–216.

[14] Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., and Woodward, C. S., "SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers," *ACM Transactions on Mathematical Software (TOMS)*, Vol. 31, No. 3, 2005, pp. 363–396.

[15] Vijayarangan, V., Uranakara, H. A., Barwey, S., Galassi, R. M., Malik, M. R., Valorani, M., Raman, V., and Im, H. G., "A data-driven reduced-order model for stiff chemical kinetics using dynamics-informed training," *Energy and AI*, Vol. 15, 2024, p. 100325.

[16] 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.

[17] Goussis, D. A., "Quasi steady state and partial equilibrium approximations: their relation and their validity," *Combustion Theory and Modelling*, Vol. 16, No. 5, 2012, pp. 869–926.

[18] Pope, S., and Maas, U., "Simplifying chemical kinetics: Trajectory-generated low-dimensional manifolds," *Mechanical and Aerospace Engineering Report: FDA*, Vol. 11, 1993.

[19] Maas, U., and Pope, S. B., "Implementation of simplified chemical kinetics based on intrinsic low-dimensional manifolds," *Symposium (international) on combustion*, Vol. 24, Elsevier, 1992, pp. 103–112.

[20] Maas, U., and Pope, S., "Laminar flame calculations using simplified chemical kinetics based on intrinsic low-dimensional manifolds," *Symposium (International) on Combustion*, Vol. 25, Elsevier, 1994, pp. 1349–1356.

[21] Lam, S.-H., and Coussis, D., "Conventional asymptotics and computational singular perturbation for simplified kinetics modelling," *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames: A Topical Volume*, 2005, pp. 227–242.

[22] Lam, S., "Using CSP to understand complex chemical kinetics," *Combustion Science and Technology*, Vol. 89, No. 5-6, 1993, pp. 375–404.

[23] Lam, S., and Goussis, D., "The CSP method for simplifying kinetics," *International journal of chemical kinetics*, Vol. 26, No. 4, 1994, pp. 461–486.

[24] Valorani, M., and Goussis, D. A., "Explicit time-scale splitting algorithm for stiff problems: auto-ignition of gaseous mixtures behind a steady shock," *Journal of Computational Physics*, Vol. 169, No. 1, 2001, pp. 44–79.

[25] Valorani, M., and Paolucci, S., "The G-Scheme: A framework for multi-scale adaptive model reduction," *Journal of Computational Physics*, Vol. 228, No. 13, 2009, pp. 4665–4701.

[26] Valorani, M., Ciottoli, P. P., Malpica Galassi, R., Paolucci, S., Grenga, T., and Martelli, E., "Enhancements of the G-scheme framework," *Flow, Turbulence and Combustion*, Vol. 101, 2018, pp. 1023–1033.

[27] Galassi, R. M., Ciottoli, P. P., Valorani, M., and Im, H. G., "An adaptive time-integration scheme for stiff chemistry based on computational singular perturbation and artificial neural networks," *Journal of Computational Physics*, Vol. 451, 2022, p. 110875.

[28] Salazar, G., and Edwards, J. R., "Mach 6 wake flow simulations using a large-eddy simulation/Reynolds-averaged Navier–Stokes model," *Journal of Spacecraft and Rockets*, Vol. 51, No. 4, 2014, pp. 1329–1348.

[29] Hansen, M. A., Armstrong, E., and Sutherland, J. C., "State space parameterization of explosive eigenvalues during autoignition," *Combustion and Flame*, Vol. 196, 2018, pp. 182–196.

[30] Zdybał, K., D'Alessio, G., Aversano, G., Malik, M. R., Coussement, A., Sutherland, J. C., and Parente, A., "Advancing reacting flow simulations with data-driven models," *arXiv preprint arXiv:2209.02051*, 2022.

[31] Gnoffo, P. A., *Conservation equations and physical models for hypersonic air flows in thermal and chemical nonequilibrium*, Vol. 2867, National Aeronautics and Space Administration, Office of Management . . . , 1989.

[32] Lee, J.-H., "Basic governing equations for the flight regimes of aeroassisted Orbital Transfer vehicles," *19th Thermophysics Conference*, 1984. https://doi.org/10.2514/6.1984-1729.

[33] Park, C., "Assessment of two-temperature kinetic model for ionizing air," *22nd Thermophysics Conference*, 1987. https://doi.org/10.2514/6.1987-1574.

[34] McBride, B. J., *NASA Glenn coefficients for calculating thermodynamic properties of individual species*, National Aeronautics and Space Administration, John H. Glenn Research Center . . . , 2002.

[35] Clarke, J., Brody, S., Steer, J., McGilvray, M., and Di Mare, L., "Quasi-one-dimensional non-equilibrium method for shock tube and stagnation line flows," *Physics of Fluids*, Vol. 36, No. 9, 2024.

[36] Gupta, R. N., Yos, J. M., Thompson, R. A., and Lee, K.-P., "A review of reaction rates and thermodynamic and transport properties for an 11-species air model for chemical and thermal nonequilibrium calculations to 30000 K," 1990.

[37] Park, C., "Nonequilibrium hypersonic aerothermodynamics," 1989.

[38] Landau, L., "Theory of sound dispersion," *Physikalische zeitschrift der Sowjetunion*, Vol. 10, 1936, pp. 34–43.

[39] Millikan, R. C., and White, D. R., "Systematics of vibrational relaxation," *The Journal of Chemical Physics*, Vol. 39, No. 12, 1963, p. 3209–3213. https://doi.org/10.1063/1.1734182.

[40] Park, C., "Problems of rate chemistry in the flight regimes of aeroassisted Orbital Transfer vehicles," *Thermal Design of Aeroassisted Orbital Transfer Vehicles*, 1985, p. 511–537. https://doi.org/10.2514/5.9781600865718.0511.0537.

[41] Munafò, A., Panesi, M., and Magin, T., "Boltzmann rovibrational collisional coarse-grained model for internal energy excitation and dissociation in hypersonic flows," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, Vol. 89, No. 2, 2014, p. 023001.

[42] Park, C., "Two-temperature interpretation of dissociation rate data for N2 and O2," *26th Aerospace Sciences Meeting*, 1988. https://doi.org/10.2514/6.1988-458.

[43] Maas, U., and Pope, S. B., "Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space," *Combustion and flame*, Vol. 88, No. 3-4, 1992, pp. 239–264.

[44] Valorani, M., Goussis, D. A., Creta, F., and Najm, H. N., "Higher order corrections in the approximation of low-dimensional manifolds and the construction of simplified problems with the CSP method," *Journal of Computational Physics*, Vol. 209, No. 2, 2005, pp. 754–786.

[45] Debusschere, B. J., Marzouk, Y. M., Najm, H. N., Rhoads, B., Goussis, D. A., and Valorani, M., "Computational singular perturbation with non-parametric tabulation of slow manifolds for time integration of stiff chemical kinetics," *Combustion Theory and Modelling*, Vol. 16, No. 1, 2012, pp. 173–198.

[46] Parente, A., and Sutherland, J. C., "Principal component analysis of turbulent combustion data: Data pre-processing and manifold sensitivity," *Combustion and flame*, Vol. 160, No. 2, 2013, pp. 340–350.

[47] 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.

[48] Biglari, A., and Sutherland, J. C., "An a-posteriori evaluation of principal component analysis-based models for turbulent combustion simulations," *Combustion and Flame*, Vol. 162, No. 10, 2015, pp. 4025–4035.

[49] Biglari, A., and Sutherland, J. C., "A filter-independent model identification technique for turbulent combustion modeling," *Combustion and Flame*, Vol. 159, No. 5, 2012, pp. 1960–1970.

[50] Pope, S. B., "Small scales, many species and the manifold challenges of turbulent combustion," *Proceedings of the Combustion Institute*, Vol. 34, No. 1, 2013, pp. 1–31.

[51] Yang, Y., Pope, S. B., and Chen, J. H., "Empirical low-dimensional manifolds in composition space," *Combustion and Flame*, Vol. 160, No. 10, 2013, pp. 1967–1980.

[52] 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.

[53] 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.

[54] Isaac, B., Parente, A., Smith, P., Fru, G., and Thevenin, D., "Source term parameterization for PCA combustion modelling," *Proceedings of the European Combustion Meeting, Lund University Press, Sweden*, 2013.

[55] Wilson, A. G., Knowles, D. A., and Ghahramani, Z., "Gaussian process regression networks," *arXiv preprint arXiv:1110.4411*, 2011.

[56] Knab, O., Fruehauf, H.-H., and Messerschmid, E., "Theory and validation of the physically consistent coupled vibration-chemistry-vibration model," *Journal of Thermophysics and Heat Transfer*, Vol. 9, No. 2, 1995, pp. 219–226.

[57] Grenga, T., Paolucci, S., and Valorani, M., "Sensitivity analysis and mechanism simplification using the G-Scheme framework," *Combustion and Flame*, Vol. 189, 2018, pp. 275–287.

[58] Shan, R., Yoo, C. S., Chen, J. H., and Lu, T., "Computational diagnostics for n-heptane flames with chemical explosive mode analysis," *Combustion and Flame*, Vol. 159, No. 10, 2012, pp. 3119–3127.