Reduced-order kinetic plasma models using principal component analysis: Model formulation and manifold sensitivity

Bellemans, A.; Magin, T.; Coussement, A.; Parente, A.

Published in:
Physical Review Fluids

DOI:
10.1103/PhysRevFluids.2.073201

Publication date:
2017

Document Version:
Accepted author manuscript

Reduced-order kinetic plasma models using principal component analysis: model formulation and manifold sensitivity

Aurélie Bellemans* and Thierry Magin†

Aeronautics and Aerospace Department,

von Karman Institute for Fluid Dynamics

72 Chaussée de Waterloo, 1640 Rhode-Saint-Genèse, Belgium.

Axel Coussement‡ and Alessandro Parente§

Service d'Aéro-Thermo-Mécanique, Université libre de Bruxelles.

50 avenue F.D. Roosevelt, 1050 Bruxelles, Belgium.

(Dated: January 4, 2017)

Abstract

Simulating large reacting systems in non-equilibrium plasma mixtures remains a challenge with the currently available computational resources. Plasma flows involve hundreds of species and thousands of reactions at different time scales, resulting in a very large set of governing equations to solve. Principal Component Analysis (PCA) offers a general and rather simple method to reduce large kinetic mechanisms by principal variable selection. Although the technique has its origin in the combustion field, it has successfully been applied to plasma. This work shows how to adapt and apply the PCA-scores technique to a collisional-radiative argon plasma model.

* aurelie.bellemans@vki.ac.be; Also at ATM, Université libre de Bruxelles.

† thierry.magin@vki.ac.be

‡ axel.coussement@ulb.ac.be

§ alessandro.parente@ulb.ac.be

I. INTRODUCTION

Accurate predictive simulations are decisive to understand and solve many of today's technological challenges to non-equilibrium plasma applications, such as the optimization of electric propulsion thrusters operating on electronegative plasmas [1, 2], the reduction of the ignition delay time and ignition temperature in plasma-assisted combustion [3, 4], the design of light-weight carbon-composite materials used as thermal protection systems for spacecraft during atmospheric entries [5, 6], the demise of metallic space debris during their reentry into the Earth atmosphere [7], as well as to understand intricate phenomena such as the sunspot formation and magnetic reconnection in solar physics[8, 9]. Non-equilibrium effects in plasmas can be described by means of many physico-chemical models of various complexity and fidelity levels. The most accurate description lies in the State-To-State (STS) models which provides the populations of the internal energy levels of all molecules and atoms of a plasma. The inner states of each species are solved separately by means of a detailed kinetic mechanism for the electronic mode of atoms, and for the rotational, vibrational and electronic modes of molecules. Collisional-Radiative (CR) models describe both collisional and radiative elementary processes. Full STS models have been developed mostly for atomic plasmas [10–12] for which hundreds of inner states can be involved. One ends up with a massive system of governing equations which is very expensive to solve numerically. With the current computational resources, the detailed chemistry of non-equilibrium mixtures is often calculated through zero-dimensional or one-dimensional computational fluid dynamics solvers.

To alleviate the computational cost inherent to STS models, the chemistry must be simplified by degrading its level of detail and accuracy, for instance to obtain three-dimensional solutions. For molecular plasmas, thousands of inner states are present. Many reduction techniques have been developed in the literature while conserving some of the detailed information about the internal energy level populations. In Multi-Temperature (MT) models, a Maxwell-Boltzmann distribution, which is the most probable distribution function according to statistical mechanics, is assumed for the population of each internal energy mode of the species at their specific rotational, vibrational, and electronic temperatures, provided that the system is close to equilibrium. However, in strong non-equilibrium situations, this

assumption is wrong as the populations will deviate from the Maxwell-Boltzmann distribution. Hybrid models combine both STS and MT approaches [13–18]. Corse-grain models have also been proposed based on a so-called binning method: the inner energy states of the species are lumped into several bins after some suitable averaging. After lumping the levels, the macroscopic rates are extracted for each group[19–21]. However, many of these techniques require strong physical insight to finely tune the reduced model. For instance, additional energy equations are required to close the governing equations by providing the temperatures used in the distribution of the coarse-grain models [22–24].

Recently, a new generation of reduction techniques has been developed in the combustion community using an empirical method which is called Principal Component Analysis (PCA) [25]. The advantage of this method lies in its simplicity as the main parameters for the reduction are selected in an automated way after solving an eigenvalue problem [26, 27]. In this contribution, accurate physical models for plasmas are combined to PCA reduction techniques. In previous papers, we have developed the very first PCA models for plasma applications [28, 29]. Although the results reported were very encouraging, the reduced model remained expensive. The goal of the present paper is to show how an optimized PCA technique can overcome the relatively high dimensionality of a CR model for argon. The paper focuses on the development of a global reduction technique based on the PCA-score method to simplify detailed chemistry. The PCA-scores technique is applied to a 34-species argon model for simulating shock tube experiments. The first argon test case chosen has been validated by Kapper and Cambier [12] based on experiments performed at the University of Toronto's Institute of Aerospace Studies (UTIAS) [30]. Anticipating the application in a future work of the reduction technique to CR models for air, we propose to verify the robustness of the reduced model by exploring a broad range of free-stream conditions representative of a re-entry flight such as the Fire II mission: very high speeds up to 12 km/s in combination with low pressures of the order of 10 Pa [31].

The paper has the following structure: the physical model for collisional-radiative argon chemistry is presented in Section II. It describes the governing equations for modeling one-dimensional shocks to simulate the argon shock tube experiments. Section III shows how PCA-scores are derived from simple principal component analysis after some pre-processing

techniques, such as outlier removal, data centering and scaling. The PCA-score technique is applied to the argon shock tube simulations for a large span of free-stream conditions in Section IV, followed by the conclusion.

II. PHYSICAL MODELING

In this section, one briefly describes the collisional-radiative model for argon plasma before applying the PCA-score technique to reduce it. The same chemical mechanism has already been used in a previous work [29] using a reduction technique which is called Manifold Generated Principal Component Analysis (MG-PCA).

A. Detailed chemistry modeling

The electronic-specific CR model for argon used in the present work has been originally developed by Vlcek [32] and Bultel *et al.* [33], and later adapted and validated by Kapper and Cambier [11, 12] against experimental data acquired in the UTIAS shock tube facility. The gas mixture consists purely of argon and its ionization products: Ar, Ar^+, and e^-. The influence of additional components, such as Ar^{++} and Ar_2^+, has been neglected [33]. The number of electronic energy levels of Ar and Ar^+ are 31 and 2, respectively. The mixture contains in total 34 species when adding up the free electrons. Table I provides the excitation energy (E_i), degeneracy (g_i) and core angular momentum ($j_{c,i}$) for the 31 electronic energy levels Ar(i) of argon. From this table, one can observe there are two possible values for the angular momentum: 0.5 and 1.5. These correspond to two possible ionization potentials when considering the ionization reactions starting from the excited states. For example, Ar(2) will ionize to $Ar^+(1)$ and Ar(4) to $Ar^+(2)$. Table II shows the energy data for those two ionized states. To account for thermal non-equilibrium between heavy particles and free electrons, separate translational temperatures are denoted in what follows by symbols T_h and T_e, respectively.

The kinetic mechanism accounts for the following collisional and radiative processes: electron-impact excitation and ionization, atom-impact excitation and ionization, photo-ionization, spontaneous emission, and free-free or Bremsstrahlung radiation, as follows:

TABLE I: Energy data for the electronic inner states of Ar.

i	E_i $[eV]$	a_i	$j_{c,i}$	i	E_i $[eV]$	a_i	$j_{c,i}$
1	0	1	1.5	17	13.864	3	1.5
2	11.548	5	1.5	18	13.903	5	1.5
3	11.624	3	1.5	19	13.979	9	1.5
4	11.723	1	0.5	20	14.013	7	1.5
5	11.828	3	0.5	21	14.063	5	1.5
6	12.907	3	1.5	22	14.068	5	1.5
7	13.076	7	1.5	23	14.090	3	1.5
8	13.095	5	1.5	24	14.099	7	1.5
9	13.153	3	1.5	25	14.153	3	1.5
10	13.172	5	1.5	26	14.214	5	0.5
11	13.273	1	1.5	27	14.234	5	0.5
12	13.283	3	0.5	28	14.236	7	0.5
13	13.302	5	0.5	29	14.241	1	0.5
14	13.328	3	0.5	30	14.255	3	0.5
15	13.480	1	0.5	31	14.304	3	0.5
16	13.845	1	1.5				

TABLE II: Energy data for the electronic inner states of Ar.

i	E_i $[eV]$	a_i	$j_{c,i}$
1	15.760	4	1.5
2	15.937	2	0.5

1. Excitation / deexcitation by electron and atom impact:

$$\mathrm{Ar(i)} + \mathrm{e^-} \underset{Fji}{\overset{Cij}{\rightleftharpoons}} \mathrm{Ar(j)} + \mathrm{e^-}, \tag{1}$$

$$\mathrm{Ar(i)} + \mathrm{Ar(_1)} \underset{Lji}{\overset{Kij}{\rightleftharpoons}} \mathrm{Ar(j)} + \mathrm{Ar(_1)}, \tag{2}$$

2. Ionization /recombination by electron and atom impact:

$$\mathrm{Ar(i)} + \mathrm{e^-} \underset{Oi}{\overset{Si}{\rightleftharpoons}} \mathrm{Ar^+} + 2\,\mathrm{e^-}, \tag{3}$$

$$\mathrm{Ar(i)} + \mathrm{Ar(_1)} \underset{Wi}{\overset{Vi}{\rightleftharpoons}} \mathrm{Ar^+} + \mathrm{e^-} + \mathrm{Ar(_1)}, \tag{4}$$

3. Spontaneous emission and absorption (*bound-bound*):

$$\mathrm{Ar(i)} + \mathrm{h_P}\nu \underset{\Lambda Aji}{\overset{(1-\Lambda)Aji}{\rightleftharpoons}} \mathrm{Ar(j)}, \tag{5}$$

4. Photo-ionization and radiative recombination (*bound-free* and *free-bound*)

$$\mathrm{Ar(i)} + \mathrm{h_P}\nu \underset{\Lambda Ri}{\overset{(1-\Lambda)Ri}{\rightleftharpoons}} \mathrm{Ar^+} + \mathrm{e^-}, \tag{6}$$

5. Bremsstrahlung radiation (*free-free*)

$$\Omega_{\mathrm{ff}}^{\mathrm{R}} = -\frac{16\pi^2}{3\sqrt{3}} \frac{\bar{\nu_e} Z_{eff}^2 e^6 \bar{g}}{m_e h (4\pi\epsilon_0 c)^3} n_{Ar+} n_e \tag{7}$$

where the indices $i = \{1, \ldots, 31\}$ and $j > i$ denote the electronic energy levels of Ar involved in the transitions. The indices for the two levels of and $\mathrm{Ar^+}$ are omitted here for the sake of brevity of the notation. Quantities $\mathrm{h_P}$ and ν in Eqs. (5-6) stand for Planck's constant and the radiation frequency, respectively. Table III gives an overview of all the rate coefficients used for each chemical process. In this model, radiation absorption is taken into account via the use of escape factors Λ [11]. The former can take values between 0 and 1, which correspond, respectively, to an optically thick and thin plasma. An optically thin plasma is considered here. For Bremsstrahlung emission, $\bar{g}$ is the gaunt factor (in this case equal to 1), and Z_{eff}^2 the effective charge (a value of 1.67 has been taken here). The total number of elementary processes (both collisional and radiative) taken into account in the adopted CR model is equal to 962. Each of these processes intervenes at a specific location after the shock wave. The initial production of electrons after the shock front starts with atom-impact collisions. The lowest excited states of argon are excited and then ionized to produce the very first electrons. Once enough electrons are produced, the chemistry dynamics changes, and electron-impact processes take over to excite additional internal energy states and produce more electrons.

TABLE III: Collisional and radiative elementary processes for argon: forward rate coefficient k_f, backward rate coefficient k_b, collisional/radiative process.

k_f	k_b	Process
Cij	*Fij*	excitation by electron impact
Kij	*Lij*	excitation by impact with the ground electronic state
Si	*Oi*	ionization by electron impact
Vi	*Wi*	ionization by impact with the ground electronic state
(1-Λ)Aji	*Aji*	radiative excitation/deexcitation
(1-Λ)Ri	*Ri*	radiative ionization/recombination

B. Governing equations

The Euler equations for two-temperature plasmas represent the conservation of mass for the N species, mixture momentum, mixture energy, and electron energy. For a 1D steady flow in the x-direction, they take the general differential form:

$$\frac{d}{dx}F = \dot{\Omega}. \tag{8}$$

In this expression, the vector F contains the fluxes, and the vector $\dot{\Omega}$ the source terms:

$$F = \begin{pmatrix} \rho u y_1 \\ \vdots \\ \rho u y_N \\ \rho u^2 + p \\ \rho u (h + \frac{1}{2}u^2) \\ \rho u y_e e^e \end{pmatrix}, \quad \dot{\Omega} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_N \\ 0 \\ -Q_{radI} \\ \Omega_e - p_e \frac{d}{dx} u - Q_{radII} \end{pmatrix}.$$

In these expressions ρ is the total mass density, u, the hydrodynamic velocity, y_i, the mass fraction of species i, ω_i, its chemical production term, p, the mixture pressure, and h, the mixture enthalpy. The variables related to the free electrons are the electron specific energy e^e , the electron energy production Ω_e due to thermal relaxation with heavy particles and electron-impact chemical reactions, and the electron partial pressure p_e.

TABLE IV: Shock tube test cases considered in the present work: from UTIAS shock tube test cases to re-entry free-stream conditions.

Case	velocity [m/s]	pressure [Pa]	temperature [K]
1	5075.23	685.32	293.6
2	5075.23	300	293.6
3	8000	50	293.6

Radiative source terms are added to the mixture and electron energy conservation equations:

$$Q_{radI} = \sum (E_j - E_i) n_j A_{ji}, \tag{9}$$

$$Q_{radII} = \sum (E_{Ar^+} - E_i) n_j R_{ji} + Q_{Brem}, \tag{10}$$

$$Q_{Brem} = -1.42 \times 10^{-40} Z_{eff}^2 \sqrt{T_e} n_+ n_e. \tag{11}$$

Quantity Q_{radI} represents the radiative power due to bound-bound transitions. These reactions are driven by the Einstein transition coefficients, Aji [s^{-1}]. Another radiative power term, Q_{radII}, has to be added for the conservation of electron energy. It regroups the free-bound and free-free transitions. In the Eqs. (9) and (10), quantity n_i stands for the number density of species i, E_{Ar+} stands for the energy of the ionized argon levels.

C. Detailing the shock structure

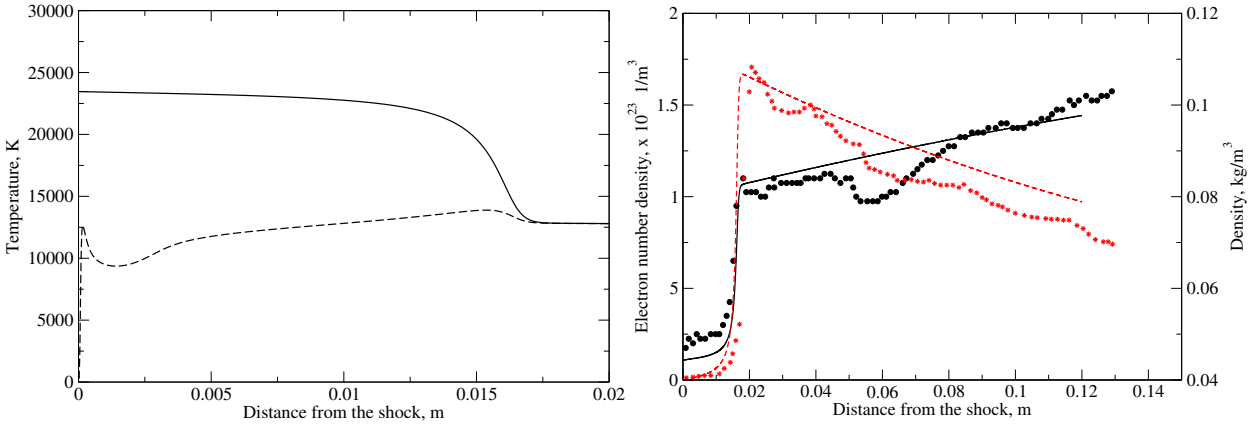
In the present work, the shock is treated as a discontinuity and placed at the location $x = 0$. Starting from free stream parameters such as the velocity, temperature and pressure, the post-shock conditions are obtained based on the 1D Rankine-Hugoniot jump relations under the assumption of frozen kinetics within the shock. The free electron temperature is also assumed to be frozen across the shock. Precursor ionization is disregarded. The plasma starts to react after the shock and the very first electrons are created at that specific moment.

A one-dimensional Eulerian solver is used to describe the flow-field behind a normal shock wave using the LSODE package [34–36]. The code has been verified and validated against the UTIAS shock tube experiments [30]. Figure 1a shows the temperature evolution of the

heavy species and the free electrons after the shock. The free electron temperature describes a peak right after the shock before decreasing an rising again until thermal equilibrium is reached. This electron peak is triggered by heavy impact ionization as it is the reaction that will create the first electrons. Chemistry changes afterwards, and electron impact processes take over the dynamics of the system. Experimental data has been plotted against the calculated electron number density and total density in Figure 1b to validate the code.

Table IV contains the free-stream parameters for three main test cases. The first test case validates the shock relaxation code against an argon shock tube experiment from the University of Toronto (UTIAS) [30]. The free-stream pressure and velocity have gradually been changed to severe conditions to show the method can catch complex flow features in a highly non-equilibrium regime. For Case 2 the pressure has been decreased while conserving the same pre-shock speed and temperature. Case 3 has severe free-stream conditions as it matches re-entry speeds and pressures: 8 km/s and 50 Pa respectively. Figure 2a shows the three cases on the testing grid. The temperature evolution is shown for these three test cases in Figure 2b.

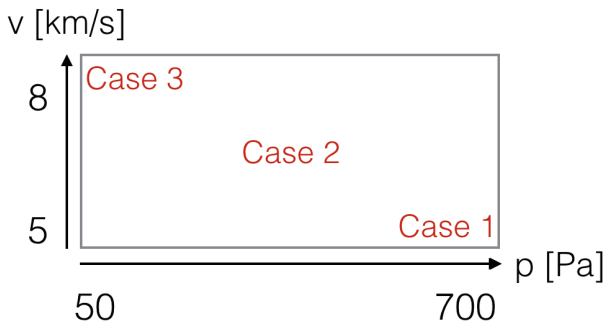
The populations of the inner states of argon have been modeled through a collisional-radiative model as explained previously. Such a model allows us to calculate the population in an accurate way through the full simulation. When plotting these populations against



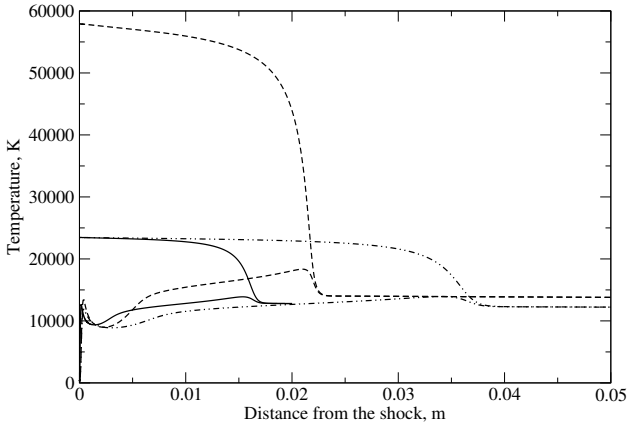
(a) Temperature evolution after the shock. Full: heavy species translational temperature. Dashed: free electron temperature.

(b) Electron number density and total density after the shock. Dashed:n_e. Stars : experimental n_e. Full: ρ. Dots: experimental ρ.

FIG. 1: Verification and validation of the Scores reduction technique on Case 1.



(a) Testing grid: Pressure varies from 50 to 700 Pa. Speed ranges from 5 to 8 km/s.



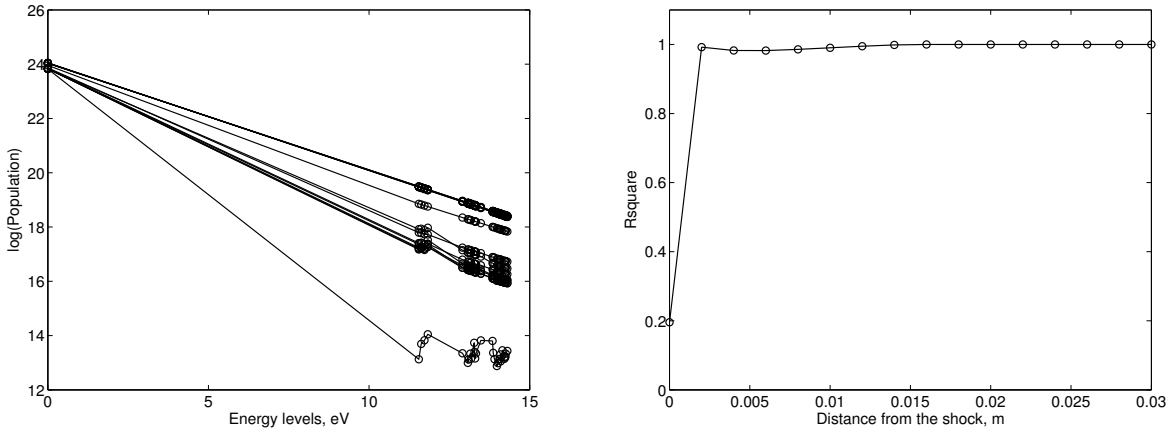
(b) Presentation of the 3 test cases. Full: Case 1. Dot-Dashed: Case 2. Dashed: Case 3.

FIG. 2: Verification and validation of the Scores reduction technique on Case 1.

their corresponding inner energy Figure 3a is obtained. For an equilibrium mixture, the populations fit to a straight line. By calculating the R^2 error between the actual populations and the straight line which represents the Maxwell-Boltzmann distribution at equilibrium, one can quantify its degree of equilibrium. At a distance from the shock higher than 0.017 m, the equilibrium condition is verified.

III. PRINCIPAL COMPONENT ANALYSIS FOR CHEMISTRY REDUCTION

Principal component analysis offers a way to reduce the dimensionality of the reacting system by projecting it on a truncated base made up out of its principal components. These principal components are uncorrelated and retain most of the variance of the system. The



(a) Logarithm of the populations of the inner states against their energy for various distances after the shock: x=0 to x= 0.02

(b) Degree of equilibrium: R^2 value for the populations in function of the distance from the shock.

FIG. 3: The populations of the inner argon states against their corresponding energy for Case 1. The R^2 value quantifies how far the mixture is from equilibrium.

computational cost decreases considerably as only a smaller number of variables, the principal components, are taken into account to solve the set of governing equations.

An eigenvalue problem is solved on a training data set made up of the mass fractions and temperatures of the system. This eigenvalue problem gives the base of principal components. By truncating the base with the component corresponding to the highest eigenvalue one can define a new, reduced, state-space. Two main methods have been developed based on principal component analysis: PCA-score by Sutherland *et al.* [37] and Manifold Generated Principal Component Analysis (MG-PCA) by Coussement *et al.* [38]. The PCA-score method allows transporting the principal components directly, as PCA is a linear transformation of the original state space. If a set of governing equations exists in the original space, a set of equation can be found in the space given by the scores. The MG-PCA method on the contrary, transports a set of variables of the state space and reconstructs the others at every iteration. The main difference between both methods lies in working space. When working with scores the governing equations are rewritten accordingly, while MG-PCA uses the original equations in the original state space. Two formulations can be adopted for the

MG-PCA method: (i) local and (ii) global. In the local approach the data set is divided in several clusters or bins, and MG-PCA is applied on each of them to obtain an optimal reduction as described in the work of Coussement *et al.* [38], where MG-local-PCA has been applied on a DNS code for simulating flame-vortex interactions. In case of global MG-PCA, the full dataset is kept as one principal cluster. A local variant of the method does not exist for PCA-scores, as a subdivision in bins would modify the definition of the principal components, leading to potential discontinuity issues.

As explained before, one can choose to work in the eigenvector space, solving transport equations for the principal components using the PC-scores approach. This section of the paper describes the PCA-score technique in more detail and shows how it can be coupled to a rotation method, such as the VARIMAX method, for retrieving a more stable formulation of the source terms and increasing the robustness of the code.

A. PCA-scores

PCA starts with a training data set containing the value of all conserved variables for several observations of the system state-space. These conserved variables correspond to mass fractions, temperatures and the velocity. However, previous work [39] has shown the reduction works best when using only mass fractions when carrying out PCA. For instance, the mass fractions are retrieved at every observation point using a simple one-dimensional CFD code in a matrix $\mathbf{Y}$, which has the size $[N_{obs} \times N_s]$:

$$\mathbf{Y} = \begin{bmatrix} y_{11} & \cdots & y_{1N_s} \\ \vdots & \ddots & \vdots \\ y_{N_{obs}1} & \cdots & {}_{N_{obs}N_s} \end{bmatrix}. \tag{12}$$

Some pre-processing techniques are applied to prepare the data for PCA. After centering the variables, they are scaled by dividing them by a scaling factor which is determined by a suitable scaling technique. Choosing a good scaling method is essential as it can affect the size and the accuracy of the reconstruction of the manifold after PCA reduction. An overview of different scaling techniques (variable stability, Pareto, max, ...) are given in the work of Parente and Sutherland [40]. In previous work on the reduction of collisional-

radiative chemistry, Pareto scaling appeared to be the most convenient method.

To correlate information of this system in terms of variance, an eigenvalue problem is solved on the covariance matrix, given by $\mathbf{S}$, to obtain the eigenvalues, $\mathbf{L}$, and eigenvectors, $\mathbf{A}$:

$$\mathbf{S} = \frac{1}{N_{obs} - 1}\mathbf{Y}^T\mathbf{Y} = \mathbf{A}\mathbf{L}\mathbf{A}^T. \tag{13}$$

As our main interest lies in the eigenvectors containing most of the variance of the system, we only select those with the highest eigenvalue. The matrix of eigenvectors can be truncated to a matrix $\mathbf{A}_q$ containing only the $q < N_s$ eigenvectors with the highest variance:

$$\mathbf{Z}_q = \mathbf{Y}\mathbf{A}_q, \tag{14}$$

$$\tilde{\mathbf{Y}}_q = \mathbf{Z}_q\mathbf{A}_q^T. \tag{15}$$

When projecting the original data set on this truncated matrix, one obtains the principal components or scores which correspond to the most important directions of the reduced system. The scores are a linear combination of the original species as they contain each variable weighted by a PCA-defined loading. When inverting relation 14 one can find back the original sample as shown in Eq. 15.

The implementation of the PC-score approach requires the solution of transport equations which are formally very similar to the ones of classic non-conserved scalars, i.e. reacting species. More generally, if a set of transport equations exists under the following conservative form:

$$\frac{\partial}{\partial t}\rho\mathbf{Y} + \nabla \cdot (\rho\mathbf{u} \otimes \mathbf{Y}) = \nabla \cdot \rho D_Y \nabla\mathbf{Y} + \omega_Y, \tag{16}$$

than it can be rewritten in the score space as follows [40]:

$$\frac{\partial}{\partial t}\rho\mathbf{Z} + \nabla \cdot (\rho\mathbf{u} \otimes \mathbf{Z}) = \nabla \cdot \rho D_Z \nabla\mathbf{Z} + \omega_Z. \tag{17}$$

The diffusion matrix and the species source terms should be transformed by using the truncated matrix of eigenvectors $\mathbf{A}_q$:

$$\omega_Z = \omega_Y\mathbf{A}_q, \tag{18}$$

$$D_Z = \mathbf{A}_q^T D_Y \mathbf{A}_q. \tag{19}$$

B. Rotated scores

The objective of using a rotation method with PCA, is to maximize the variance expressed by the eigenvectors. Using such a rotation method simplifies the interpretation of the PCA results as the score loadings after rotation will be either very large or close to zero. Consequently, it clearly shows which variables are expressed the most by the scores, and thus which are the dominating ones, for a given size. It is important to note that the total expressed variance doesn't change before and after rotation. Besides a better interpretation of the PCA results, a rotation methods also ensures additional stability within the CFD code. After rotation, the data flattens out and contains less peaks, which is beneficial for the computation of the source terms for example, as stated by Coussement *et al.* [41]. It ensures robustness within the code.

The VARIMAX rotation developed by Kaiser [42] has generally been accepted as the most accurate orthogonal rotation and has been used widely in combination with PCA [43]. The VARIMAX rotation criterion maximizes the sum of the variances of the squared coefficients within each eigenvector. The axes in the new system are rotated to maximize the rotation criterion given by the following expression:

$$V = \frac{n \sum_{i=1}^{n} (b_{ij}^2)^2 - \gamma (\sum_{i=1}^{n} b_{ij}^2)^2}{n^2}. \tag{20}$$

In this expression b stands for the principal component loadings, and n for the number of variables, with $i \in \{1, ..., n\}$ and $j \in \{1, ..., N_{pc}\}$. When γ equals 1, a classical VARIMAX criterion is applied. For $\gamma = 0$ we obtain the QUARTIMAX criterion. Values between those two extremes correspond to the EQUAMAX and PARSIMAX rotation criterions. In this work, γ has been chosen to 1 to represent the VARIMAX criterion.

A normalized variant of this criterion exists. The b_{ij} terms in expression (20) should be normalized by the square root of the communalities h_i. The communalities are defined as the sum of the squares of the i^{th} row of the loading matrix. This normalized variant has been used in the present work:

$$V_{scaled} = \frac{n \sum_{i=1}^{n} (b_{ij}^2/h_i^2)^2 - \gamma (\sum_{i=1}^{n} b_{ij}^2/h_i^2)^2}{n^2}. \tag{21}$$

For working with rotated scores, the rotation matrix $\mathbf{T}$ should be taken into account when transforming the conserved variables, diffusion matrix and source terms to the new score space:

$$\mathbf{Z}_q = \mathbf{YTA}_q, \tag{22}$$

$$\omega_Z = \omega_Y \mathbf{TA}_q. \tag{23}$$

IV. RESULTS

The one-dimensional SHOCKING code, developed by Munafò [36] has been used to simulate the relaxation of shocks in argon plasma. The code is coupled with the LSODE solver [34]. A large frame of free-stream conditions will be covered to demonstrate the applicability of the PCA-score technique to plasma flows. The free-stream parameters for the different test cases are presented in Table IV.

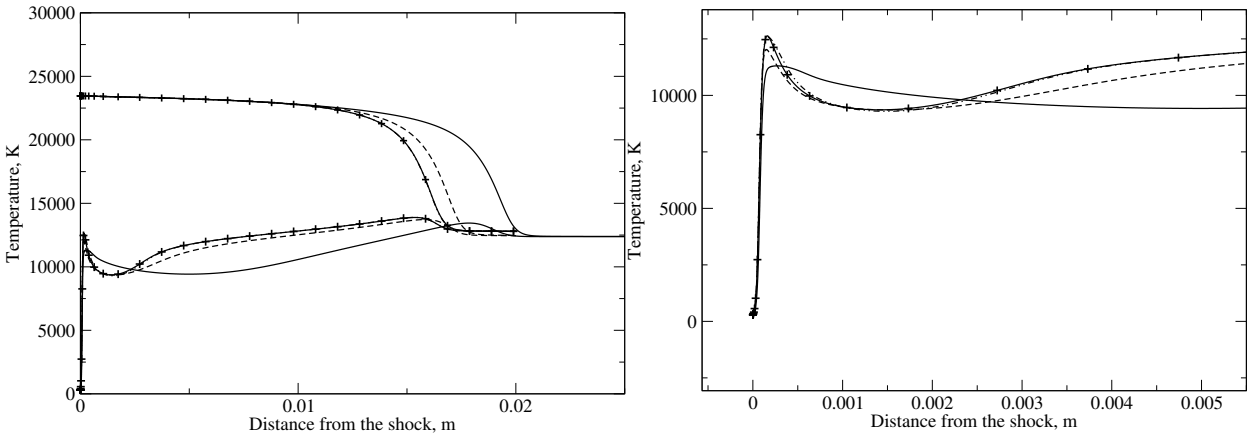
A. PCA-scores for re-entry conditions

The PCA-score method has been applied on the test cases described in Table IV. To do so, the CFD code has been modified to solve the equations in the score space. As only mass fractions are considered in the present work, only the species transport equations need to be projected onto the eigenvector space:

$$\frac{\partial}{\partial x}[\rho u \mathbf{Z}_i] = \omega_{Zi}, \tag{24}$$

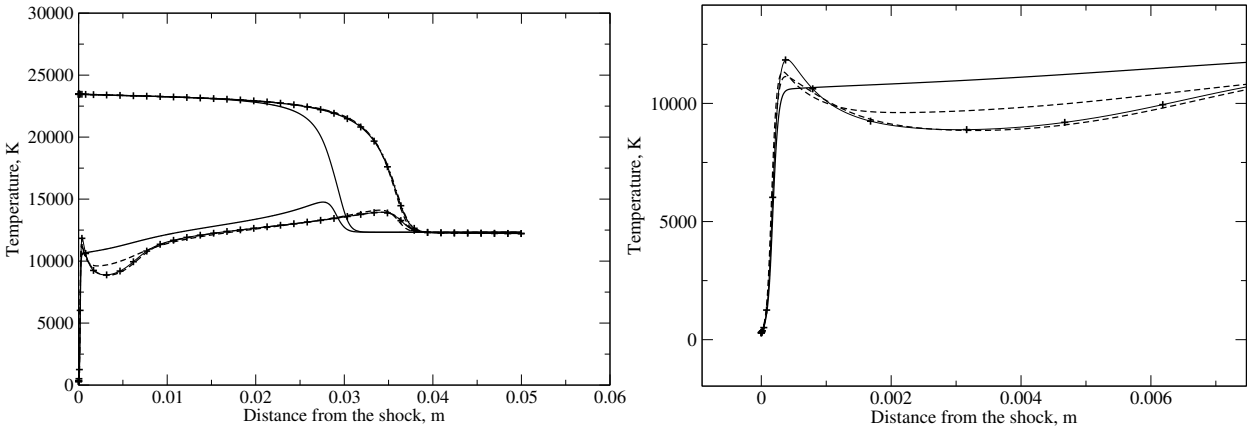
with $i \in \{1, \cdots, N_{scores}\}$. It is important to notice we need all species mass fractions to calculate the thermodynamic and transport properties of the plasma. After each iteration of the solver, the scores are reverted to mass fractions by using Equation 15.

In a next step, the necessary amount of scores is determined for each test case. Ideally, the case should run using only one score. The main criterion for retaining the right amount of scores is the following: they should represent the original manifold in a perfect way. The ideal number of scores required to reconstruct with sufficient accuracy the original state-space, can be determined with an *a priori* study comparing the reconstruction error of each variable after PCA. Of course this is not a sufficient condition the actual solution of the score transport equations in a CFD code lead to error propagation, resulting in a



(a) Simulation of Case 1 with 3 to 5 scores. (b) Zoom on the zone right after the shock.

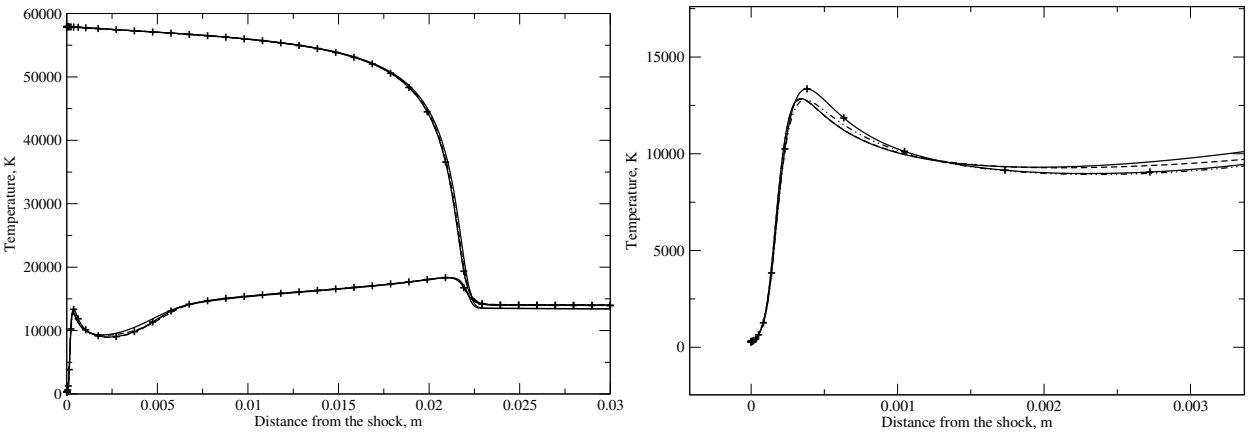
FIG. 4: PCA-score simulation for Case 1.
+ : CR. Full: 3 scores. Dashed: 4 scores. Dot-Dashed: 5 scores.



(a) Simulation of Case 2 with 3 to 5 scores. (b) Zoom on the zone right after the shock.

FIG. 5: PCA-score simulation for Case 2.
+ : CR. Full: 3 scores. Dashed: 4 scores. Dot-Dashed: 5 scores.

higher number of scores needed. Figure 4 presents the results of the PCA-score reduction on Case 1. The left picture shows the temperature evolution after the shock (located at the y-axis) for both heavy species (upper curves) and the electrons (lower curves). With 5 scores the solution is perfect compared to the full CR model. With 4 scores some discrepancies are observed, but with 3 scores there are major differences, in the shape of the electron temperature evolution and in the relaxation time before obtaining equilibrium. Figure 5b, shows a zoom of the electron temperature. The dimensionality reduction is impressive as



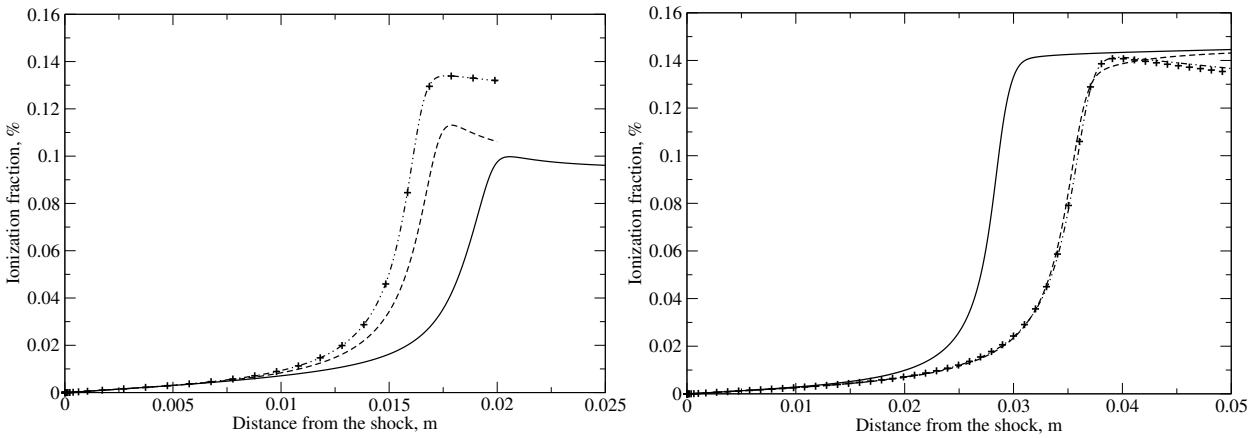
(a) Simulation of Case 3 with 4 to 6 scores. (b) Zoom on the zone right after the shock.

FIG. 6: PCA-score simulation for Case 3.
+ : CR. Full: 4 scores. Dashed: 5 scores. Dot-Dashed: 6 scores.

the species equations have been reduced from 34 to only 5.

The reduction of Case 2 is very similar to the previous one as only the pressure has been lowered. The lower the pressure, the more non-equilibrium effects can be observed. Figure 5 shows that 5 scores are needed to obtain a satisfactory results. The results for case 3 are shown in Figure 6. Case 3 has special features, as the speed has been increased and the pressure lowered to match atmospheric re-entry situations. As the free-stream conditions are severe, the solver needs more scores to retrieve an accurate solution. No more than 4 scores were needed to retrieve a converged solution. The PCA-scores solution is virtually identical compared to the full CR model, with only 6 scores.

Next, the ionization degree obtained using the full and reduced models has been compared. Figure 7 shows the ionization degree for both Cases 1 and 2. When using 5 scores for Case 1, the ionization degree is perfectly reproduced. For Case 2, a very good representation of the ionization degree is obtained with 5 scores. The ionization degree for Case 3 is shown in Figure 8. As it can be observed, the ionization degree cannot be accurately predicted in the equilibrium zone with 6 scores, despite the fact that temperature was perfectly predicted. A perfect approximation of the ionization degree could be obtained using a reduction based on 13 scores. A model reduction from 34 to 13 variables still represents a good reduction



(a) Case 1 with 3, 4 and 5 scores.

(b) Case 2 with 3, 4 and 5 scores.

FIG. 7: Comparison of the ionization degree between the full CR model and the reduced model with PCA-score. + : CR. Full: 3 scores. Dashed: 4 scores. Dot-Dashed: 5 scores.

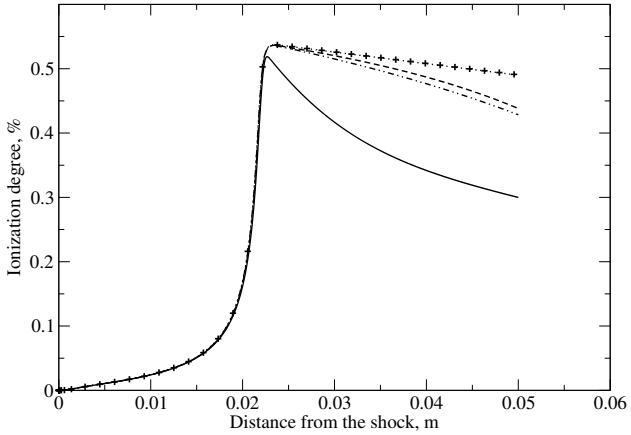


FIG. 8: Comparison of the ionization degree between the full CR model and the reduced model with PCA-score for Case 3. + : CR. Full: 4 scores. Dashed: 5 scores. Dot-Dashed: 6 scores. Dotted: 13 scores.

considering the difficult free-stream condition and high degree of non-equilibrium.

B. PCA-scores with VARIMAX rotation

A VARIMAX rotation on the eigenvectors of the data has been carried on all the cases as it increases the robustness of the code. Moreover, rotated eigenvectors and loadings are easier to interpret as the variance expressed by them has been maximized. To illustrate

this, the loadings for the first two scores have been compared against their rotated ones for Case 2 in Figure 9. As a reminder: the loadings correspond to the weights of the original variables expressed by the principal components or scores. Figure 9 shows that Score 1 is mainly composed of the ground state of argon and the ionized species. This information was concealed within the unrotated scores.

As mentioned before, the rotation operation flattens the new source terms for calculating the mass conservation of the scores. Figure 10 represents the source terms for the ground state of argon before and after rotation. The red curve, corresponding to the rotated scores, shows a significantly less noisy behavior than the non-rotated ones. This is an important results as it shows how VARIMAX rotation can improve the stability of the method within CFD codes.

C. Manifold sensitivity study

The potential reduction of the PC-score approach has been demonstrated in the previous Sections. We want to investigate now the applicability of the reduced models in conditions which differ from the ones used to generate the model. This implies imposing different free stream conditions for models to the reduced models obtained for cases 1-3. The test domain can be visualized in Figure 2a. The pressure varies from 50 to 700 Pa, and the velocity from 5 to 8 km/s. The objective of this study is to investigate if a reduced model based on a

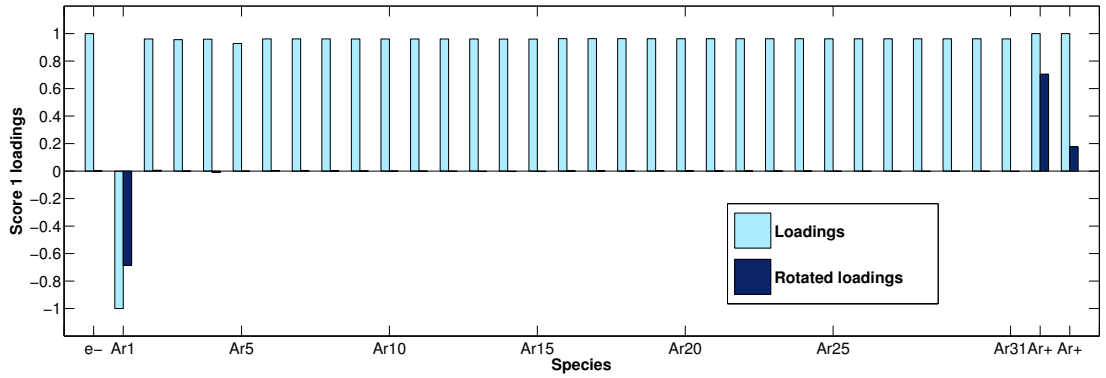


FIG. 9: Loadings for the original variables before and after the VARIMAX rotation. Case 2: 300 Pa. Light blue bars: original loadings. Dark blue bars: rotated loadings.

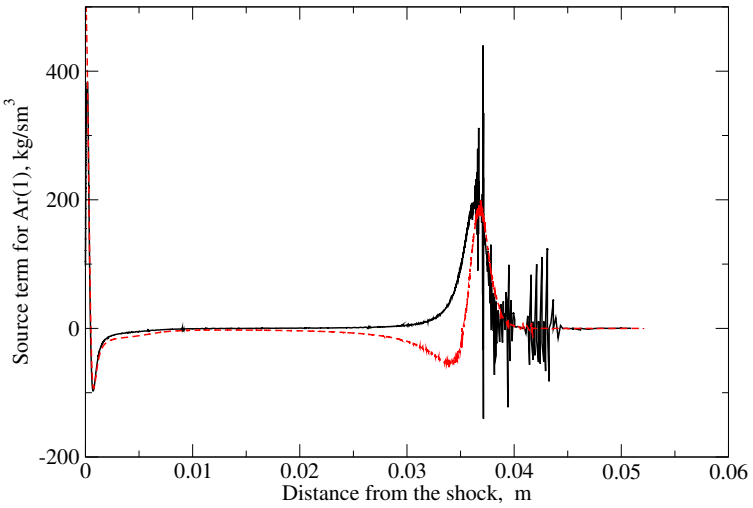
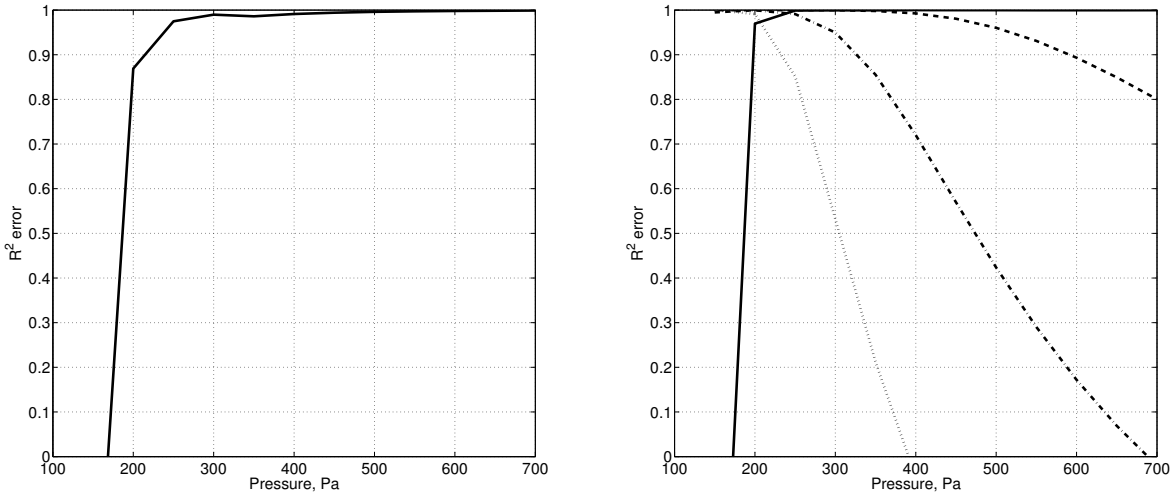


FIG. 10: Comparison of mass production source term for the ground state of argon with and without VARIMAX rotation.
Full black: original loadings. Dashed red: rotated loadings.

particular training data set given by case 1, 2 or 3, is still accurate in the other conditions given by the test grid. To this purpose, a least square error (R^2) will be evaluated between the full and reduced model. With a R^2=1, the reduced model gives a perfect representation of the original one.

The first investigation uses case 1, 2 and 3 separately, and looks how the reduction behaves throughout the domain. Figure 11 represents the least square error associated to the reduced models for cases 1 and 2. Using the reduction for case 1 at lower pressures and higher speeds does not provide satisfactory results outside the conditions used to train the model. However, the reduced model for case 2 provides good results at 5km/s at at relatively low pressures for higher speeds. The same trend can be observed for the reduced model based on case 3, in Figure 12. The results of two reduced models for case 3 are compared in Figure 12: one based on 13 scores (Figure 12a), as it correctly reproduced the ionization degree, and one based on 7 scores (Figure 12b), as it correctly reproduced the post-shock temperature. Figure 12 clearly shows that the 13 scores model has a wider range of applicability than the one based on 7 scores. This is expected since increasing the number of scores also increases the amount of information included in the reduced model. To conclude this investigation, we can state that a reduced model is accurate for free-stream



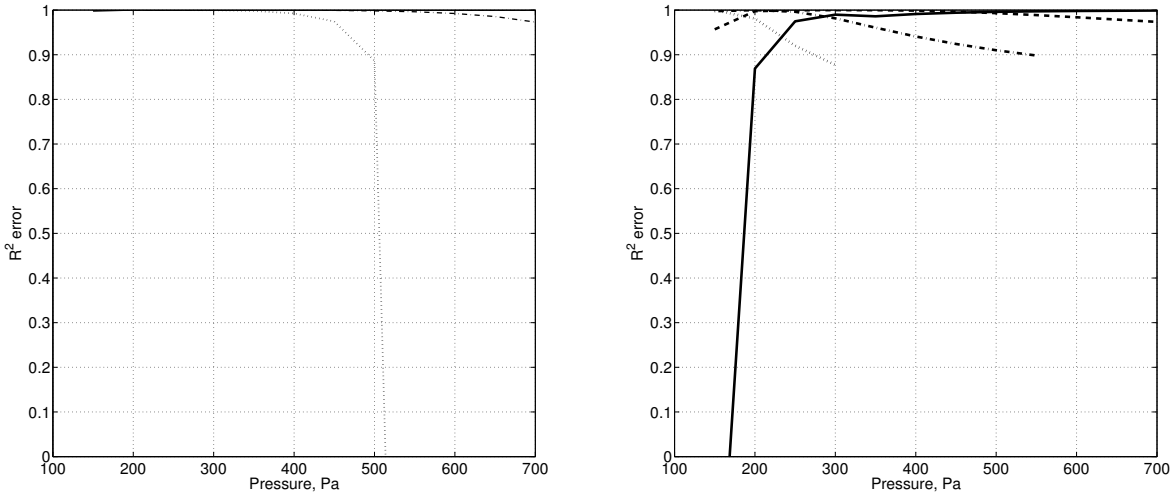
(a) R^2 error for the 5-scores reduced model for Case 1.

(b) R^2 error for the 5-scores reduced model for Case 2.

FIG. 11: Manifold sensitivity study by comparing the R^2 error for different test cases in different pressure and temperature conditions. Full: 5 km/s. Dashed: 6 km/s. Dot-Dashed: 7 km/s. Dotted: 8 km/s.

parameters which are less severe than the one of the training data set.

Next, we investigate the possibility of generating reduced models from a finite number of cases spanning the complete domain of interest, in terms of speed and pressure, and to use them for conditions not originally included in the training dataset. Cases 1 and 3 are used to generate a reduced model, that is later used to simulate the conditions of case 2. The PCA-score method has been applied to the combined data sets for cases 1 and 3 to find a global reduced model. The approach is found to be very effective, as it provides virtually perfect results for case 2. This is a good strategy for generating reduced models for CFD applications, since the operating conditions might vary in terms of pressure and velocity within a single simulation and therefore, a case built from a span of conditions is needed.



(a) R^2 error for the 13-score reduced model for Case 3.

(b) R^2 error for the 7-score reduced model for Case 3.

FIG. 12: Application of reduced models developed for specific flow conditions to different temperatures and pressures. Full: 5 km/s. Dashed: 6 km/s. Dot-Dashed: 7 km/s. Dotted: 8 km/s.

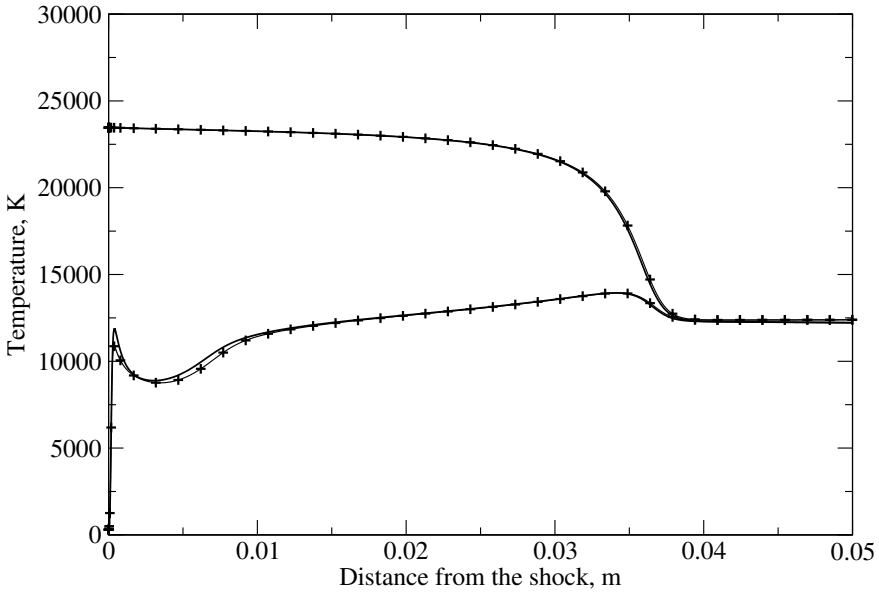


FIG. 13: The score-models for Case 1 and 3 have been used in an interpolation to simulate Case 2. + : CR data for Case 2. Full : interpolated model.

V. CONCLUSION

The present work shows how principal component analysis can be used in the PCA-score technique to reduce large and complex chemistry models. Moreover, it shows how reduction methods developed for combustion applications can easily be transferred to plasma flows. PCA-score is a very powerful tool to reduce, in this particular case, collisional-radiative chemistry for argon plasma. The main results of this work are the following:

- For a model using 34 species and 2 temperatures, the computation could be carried out using only 3 scores. For an exact solution, at least 5 scores were needed. This is an encouraging result as the set of governing equations has been reduced with more than 90% which leads to an important speed-up of the calculation and a reduction of computational cost.

- When using a rotation method, such as the VARIMAX criterion in the present case, the source terms are smoother and allow the CFD code to converge more easily. It also offers a simpler way to interpret the score loadings with respect to the original data.

- The sensitivity study has shown that a given reduced model can be used in a large frame of test conditions while conserving good accuracy. This means only one training data set is needed for building a performing reduced model for free-stream conditions.

- A given reduced model can be extrapolated to other free-stream conditions. However, it has been shown a reduced model based on severe free-stream conditions provides satisfactory results.

- By combining two data sets one can build an interpolated model for intermediate conditions. It is an effective method for dealing with various conditions in a single simulation and is suitable for CFD applications.

ACKNOWLEDGMENTS

The research of A. Bellemans has been sponsored by a FRIA fellowship of the Belgian research fund F.R.S.-FNRS.

---

[1] D. M. Goebel and I. Katz, *Fundamentals of Electric Propulsion: Ion and Hall Thrusters* (Wiley, 2008).

[2] M. Thomas, D. Rafalskyi, T. Lafleur, and A. Aanesland, Plasma Sources Science and Technology **25**, 045018 (2016).

[3] S. V. Pancheshnyi, D. A. Lacoste, A. Bourdon, and C. O. Laux, IEEE Transactions on Plasma Science **34**, 2478 (2006).

[4] V. Adamovich and W. R. Lempert, Plasma Phys. Control. Fusion **57**, 014001 (2015).

[5] C. Park, *Nonequilibrium Hypersonic Aerothermodynamics* (Wiley, 1990).

[6] B. Helber, A. Turchi, J. B. Scoggins, A. Hubin, and T. E. Magin, International Journal of Heat and Mass Transfer **100**, 810 (2016).

[7] H. Klinkrad, *Space debris, models and risks analysis* (Springer, 2006).

[8] H. Goedbloed and S. Poedts, *Principles of Magnetohydrodynamics* (Cambridge, 2004).

[9] A. A. Laguna, A. Lani, H. Deconinck, N. Mansour, and S. Poedts, Journal of Computational Physics **318**, 252 (2016).

[10] A. Bultel, B. G. Chéron, A. Bourdon, O. Motapon, and I. F. Schneider, Phys. Plasmas **13**, 043502 (2006).

[11] M. G. Kapper and J.-L. Cambier, J. Appl. Phys. **109**, 113308 (2011).

[12] M. G. Kapper and J.-L. Cambier, J. Appl. Phys. **109**, 113309 (2011).

[13] A. Bogaerts, R. Gijbels, and J. Vlcek, Journal of Applied Physics **84**, 121 (1998).

[14] M. Panesi, T. E. Magin, A. Bourdon, A. Bultel, and O. Chazot, J. Thermophys. Heat Transfer **23**, 236 (2009).

[15] J. Annaloro and A. Bultel, Plasma Sources Sci. Technol. **22**, 025008 (2013).

[16] A. Guy, A. Bourdon, and M.-Y. Perrin, Chem. Phys. **420**, 15 (2013).

[17] M. Capitelli, G. Colonna, G. D'Ammando, K. Hassouni, A. Laricchiuta, and L. D. Pietanza, Plasma Processes and Polymers (2016).

[18] R. Macdonald, A. Munafò, C. Johnston, and M. Panesi, Physical Review Fluids **1**, 043401 (2016).

[19] T. Magin, M. Panesi, A. Bourdon, R.L., and D. W. Schwenke, Chem. Phys. **398**, 90 (2012).

[20] M. Panesi and A. Lani, Phys. Fluids **25**, 057101 (2013).

[21] A. Munafò and T. E. Magin, Phys. Fluids **26**, 097102 (2014).

[22] H. P. Le, A. R. Karagozian, and J. L. Cambier, Phys. Plasmas **20**, 123304 (2013).

[23] A. Munafò, M. Panesi, and T. E. Magin, Phys. Rev. E **89**, 023001 (2014).

[24] G. Colonna, G. DAmmandob, L. D. Pietanza, and M. Capitelli, in *Proc. of the 27th Int. Symposium on Rarefied Gas Dynamics* (AIP, 2010).

[25] J. Shlens, ArXiv e-prints **arXiv:1404.1100v1** (2014).

[26] A. Parente, J. C. Sutherland, L. Tognotti, and P. J. Smith, Proc. Combust. Inst. **32**, Pages 15791586 (2009).

[27] A. Coussement, O. Gicquel, and A. Parente, Combustion and Flame **159**, 2844 (2012).

[28] K. Peerenboom, A. Parente, T. Kozak, A. Bogaerts, and G. Degrez, Plasma Sources Science and Technology **24**, 025004 (2014).

[29] A. Bellemans, A. Munafo, T. Magin, G. Degrez, and A. Parente, Physics of Plasmas **22**, 062108 (2015).

[30] I. I. Glass, *Over forty years of continuous research at UTIAS on nonstationary flows and shock waves* (Springer, 1991).

[31] D. L. Cauchon, NASA TM X-1402 **6** (1967).

[32] J. Vlcek, J. Phys. D **22**, 623 (1989).

[33] A. Bultel, B. van Ootegem, A. Bourdon, and P. Vervisch, Phys. Rev. E **65**, 046406 (2002).

[34] K. Radhakrishnan and A. C. Hindmarsh, *Description and use of* LSODE*, the Livermore solver for ordinary differential equations*, NASA Report 1327 (NASA, 1993).

[35] T. E. Magin, L. Caillault, A. Bourdon, and C. O. Laux, Journal of Geophysical Research **111**, E07S12 (2006).

[36] A. Munafò, *Multi-Scale Models and Computational Methods for Aerothermodynamics*, Ph.D. thesis, Ecole Centrale Paris, Châtenay-Malabry, France (2014).

[37] J. Sutherland and A. Parente, Proc. Combust. Inst. **32**, 15631570 (2009).

[38] A. Coussement, A. Parente, and O. Gicquel, Proceedings of the Combustion Institute **34**, 1117 (2013).

[39] B. Isaac, *Reduced-order modelling for reacting flows based on principal component analysis*, Ph.D. thesis, University of Utah (2014s).

[40] A. Parente and J. Sutherland, Combustion and Flame **160**, 340 (2013).

[41] A. Coussement, B. J. Isaac, O. Gicquel, and A. Parente, Combustion and Flame **168**, 83 (2016).

[42] H. F. Kaiser, Psychometrika **23**, 187 (1958).

[43] M. Richman, Journal of Climatology **6**, 293 (1986).

[44] I. D. Boyd, Prog. Aerosp. Sci. **41**, 669 (2005).

[45] A. A. Laguna, A. Lani, Y. Maneva, N. Ozak, H. Deconinck, and S. Poedts, (2016).

[46] A. Bellemans, A. Coussement, T. Magin, and A. Parente, *MG-local-PCA method for the reduction of a collisional-radiative argon plasma mechanism*, AIAA Conference Paper 2017 (AIAA, 2015) aVIATION 2015 Conference, Dallas, USA.

[47] T. E. Magin, L. Caillault, A. Bourdon, and C. O. Laux, Journal of Geophysical Research **111** (2006).

[48] J. C. Keck, Prog. Energy Combust. Sci. **16**, 125 (1990).

[49] J. C. Keck, AIP Conf. Proc. **1033**, 329 (2008).

[50] Y. Yang, S. B. Pope, and J. H. Chen, Combustion and Flame **160**, 1967 (2013).

[51] A. Biglari and J. C. Sutherland, Combustion and Flame **159**, 1960 (2012).

[52] J. A. Cumming and D. A. Wooff, Computational Statistics and Data Analysis **52**, 550 (2007).

[53] H. Mirgolbabaei and T. Echekki, Combustion and Flame **161**, 118 (2014).

[54] H. Mirgolbabaei, T. Echekki, and N. Smaoui, International Journal of Hydrogen Energy **39**, 4622 (2014).

[55] A. Yanguas-Gil, J. Cotrino, and A. R. Gonzalez-Elipe, Physics of Plasmas **11**, 5497 (2004).

[56] K. Miki, M. Panesi, E. E. Prudencio, A. Maurente, S. H. Cheung, J. J. Jagodzinski, D. Goldstein, S. Prudhomme, K. Schulz, C. Simmons, J. Strand, and P. Varghese, in *Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition* (AIAA, Orlando, Florida, 2010).

[57] G. P. M. Cabe, Technometrics **26**, 137 (1984).

[58] Z. Ren, S. B. Pope, A. Vladimirsky, and J. M. Guckenheimer, The Journal of Chemical Physics **124**, 114111 (2006).

[59] C. Park, J. Thermophys. Heat Transfer **7**, 385 (1993).

[60] C. Park, J. T. Howe, R. L. Jaffe, and G. V. Candler, J. Thermophys. Heat Transfer **8**, 9 (1994).

[61] G. Colonna, M. Tuttafesta, M. Capitelli, and D. Giordano, J. Thermophys. Heat Transfer **13**, 372 (1999).

[62] G. Colonna and M. Capitelli, J. Thermophys. Heat Transfer **15**, 308 (2001).

[63] C. O. Laux, in *Physico-Chemical Models for High Enthalpy and Plasma Flows*, Lecture Series (von Karman Institute for Fluid Dynamics, 2006).

[64] L. Pierrot, L. Yu, R. J. Gessman, C. O. Laux, and C. Kruger, *Collisional-radiative modeling of nonequilibrium nitrogen plasmas*, AIAA Conference Paper 1999–3478 (30th Plasmadynamics and Lasers Conference, Norfolk, VA, 1999).

[65] R. J. Gessman, C. O. Laux, and C. Kruger, *Experimental study of kinetic mechanisms of recombining atmospheric pressure air plasmas*, AIAA Conference Paper 1997–2364 (28th AIAA Plasma-dynamics and Lasers Conference, Atlanta, GA, 1997).

[66] M. Panesi, T. E. Magin, A. Bourdon, A. Bultel, and O. Chazot, J. Thermophys. Heat Transfer **25**, 361 (2011).

[67] A. Munafò, A. Lani, A. Bultel, and M. Panesi, Phys. Plasmas **20**, 073501 (2013).

[68] G. Colonna, I. Armenise, D. Bruno, and M. Capitelli, J. Thermophys. Heat Transfer **20**, 477 (2006).

[69] M. Capitelli, C. M. Ferreira, B. F. Gordiets, and A. I. Osipov, *Plasma Kinetics in Atmospheric Gases* (Springer, 2000).

[70] I. V. Adamovich, S. O. Macheret, J. W. Rich, and C. E. Treanor, AIAA J. **33**, 1064 (1995).

[71] I. V. Adamovich, S. O. Macheret, J. W. Rich, and C. E. Treanor, AIAA J. **33**, 1070 (1995).

[72] A. Bellemans, A. Munafò, T. Magin, A. Parente, and G. Degrez, in *AIAA Aviation Conference* (Atlanta, USA, 2014).

[73] B. Helber, A. Turchi, O. Chazot, T. Magin, and A. Hubin, in *AIAA/ASME Joint Thermophysics and Heat Transfer Conference* (AIAA, Atlanta, USA, 2014).

[74] J. Kim and S. B. Pope, Combustion Theory and Modelling **18**, 388 (2014).

[75] S. B. Pope, FDA **03-02** (2003).

[76] W. Graef, *Zero-dimensional models for plasma chemistry*, Ph.D. thesis, Technische Universiteit Eindhoven (2012).

[77] B. Isaac, A. Coussement, O. Gicquel, P. J. Smith, and A. Parente, Combustion and Flame **161**, 2785 (2014).

[78] J. B. Scoggins and T. Magin, 11th AIAA/ASME Joint Thermphysics and Heat Transfer Conference (2014).

[79] Ph.D. thesis.

[80] A. Munafo, M. Panesi, R. Jaffe, G. Colonna, A. Bourdon, and T. Magin, Eur. Phys. Jour. D **66** (2012).

[81] M. Panesi, R. L. Jaffe, D. W. Schwenke, and T. E. Magin, J. Chem. Phys **138**, 044312 (2013).

[82] M. Panesi, A. Munafò, T. E. Magin, and R. L. Jaffe, Phys. Rev. E **90**, 013009 (2014).

[83] R. L. Jaffe, D. W. Schwenke, G. Chaban, and W. Huo, *Vibrational and rotational excitation and relaxation of nitrogen from accurate theoretical calculations*, AIAA Paper 2008–1208 (46th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, 2008).

[84] G. Chaban, R. L. Jaffe, D. W. Schwenke, and W. Huo, *Dissociation cross-sections and rate coefficients for nitrogen from acccurate theoretical calculations*, AIAA Paper 2008–1209 (46th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, 2008).

[85] R. L. Jaffe, D. W. Schwenke, and G. Chaban, *Theoretical analysis of* N_2 *collisional dissociation and rotation-vibration energy transfer*, AIAA Paper 2009–1569 (47th AIAA Aerospace Sciences Meeting and Exhibit, Orlando, FL, 2009).

[86] D. W. Schwenke, in *Non-Equilibrium Gas Dynamics - From Physical Models to Hypersonic Flights*, Lecture Series (von Karman Institute for Fluid Dynamics, 2008).

[87] Y. Liu, M. Vinokur, M. Panesi, and T. E. Magin, *A multi-group maximum entropy model for thermo-chemical nonequilibrium*, AIAA Paper 2010–4332 (10th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, Chicago, IL, 2010).

[88] Y. Liu, M. Panesi, M. Vinokur, and P. Clarke, *Microscopic simulation and macroscopic modeling of thermal and chemical non-equilibrium gases*, AIAA Paper 2013–3146 (44th AIAA Thermophysics Conference, San Diego, CA, 2013).

[89] M. Massot, Comptes Rendus Mathematique **335**, 93 (2002).