Development of Reduced Chemistry Models for High Enthalpy and Plasma flows

A. Bellemans\textsuperscript{1+}, A. Munaf\textsuperscript{ö} and T.E. Magin\textsuperscript{3∗}

\textit{von Karman Institute for Fluid Dynamics, Rhode-Saint-Genèse, Belgium}

A. Parente\textsuperscript{4†} and G. Degrez\textsuperscript{5†}

\textit{Université Libre de Bruxelles, Bruxelles, Belgium}

This paper considers the development of reduced chemistry models for high enthalpy and plasma flows using Principal Component Analysis (PCA) based methods. Starting from detailed chemistry models, such as multi-temperature and collisional-radiative formulations, a reduction of the variable set (species mass fractions and temperatures) is proposed by projecting the full set on a reduced basis made up of its principal components. Consequently, an important reduction of calculation time is obtained as the governing flow equations are solved for these principal components only. This approach originates from the combustion field, where manifold generated principal component analysis (MG-PCA) has been developed as a successful reduction technique. In this work MG-PCA has been implemented and verified on shock tube simulations for argon plasma based on an electronic specific detailed chemistry model. The method has been validated with experimental results from UTIAS (University of Toronto, Institute for Aerospace Studies).

Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASDF</td>
<td>Atomic State Distribution Function</td>
</tr>
<tr>
<td>CR</td>
<td>Collisional-Radiative</td>
</tr>
<tr>
<td>$E$</td>
<td>energy level for each species</td>
</tr>
<tr>
<td>$e^e$</td>
<td>electron specific energy</td>
</tr>
<tr>
<td>$eV$</td>
<td>electron Volt</td>
</tr>
<tr>
<td>$g$</td>
<td>degeneracy</td>
</tr>
<tr>
<td>$h$</td>
<td>enthalpy of the mixture</td>
</tr>
<tr>
<td>$i,j$</td>
<td>indices for energy levels with $j &gt; i$</td>
</tr>
<tr>
<td>MT</td>
<td>Multi-Temperature</td>
</tr>
<tr>
<td>$p$</td>
<td>pressure</td>
</tr>
<tr>
<td>$p_e$</td>
<td>partial pressure of the electrons</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PV</td>
<td>Principal Variable</td>
</tr>
<tr>
<td>$u$</td>
<td>bulk velocity</td>
</tr>
<tr>
<td>$u_h$</td>
<td>heavy particle velocity</td>
</tr>
<tr>
<td>$u_e$</td>
<td>electron bulk velocity</td>
</tr>
<tr>
<td>$\rho$</td>
<td>mass density</td>
</tr>
<tr>
<td>$\omega_i$</td>
<td>chemical production term for species $i$</td>
</tr>
<tr>
<td>$y_e$</td>
<td>mass fraction of electrons</td>
</tr>
<tr>
<td>$y_i$</td>
<td>mass fraction of species $i$</td>
</tr>
<tr>
<td>$\Omega_e$</td>
<td>electronic energy production</td>
</tr>
</tbody>
</table>

\textsuperscript{1} PhD Candidate, AR department, Chaussée de Waterloo 72, 1640 Rhode-Saint-Genèse, Belgium.
\textsuperscript{2} PhD Candidate, AR department, Chaussée de Waterloo 72, 1640 Rhode-Saint-Genèse, Belgium.
\textsuperscript{3} Associate professor, AR department, Chaussée de Waterloo 72, 1640 Rhode-Saint-Genèse, Belgium.
\textsuperscript{4} Assistant professor, Service d'Aéro-Thermo-Mécanique, Avenue F.D. Roosevelt 50, 1050 Bruxelles, Belgium.
\textsuperscript{5} Professor, Service d'Aéro-Thermo-Mécanique, Avenue F.D. Roosevelt 50, 1050 Bruxelles, Belgium.

American Institute of Aeronautics and Astronautics
I. Introduction

The modeling of high enthalpy and plasma flows is a complex multi-physical engineering problem because of the high degree of non-equilibrium in the composition of the various plasma components. Indeed, thermochemical non-equilibrium phenomena involve hundreds of reactions and thousands of species leading to expensive numerical simulations. As detailed multi-dimensional formulations are computationally expensive, one reduces either the dimensionality of the problem going from 3D to 1D or 0D expressions, or the physics by replacing detailed chemistry laws by simpler ones. Fortunately, although complete models involve hundreds of species and thousands of reactions, only a small fraction of these play a major role in aerospace applications. This is due to the fact that many chemical processes have much shorter time scales than the typical transport time scale in the system, just as in combustion processes. As a result, the chemical composition lies on a so-called low dimensional manifold. This observation opens the way for the development of reduced chemical models using the same kind of techniques developed for combustion processes.

This paper considers the development of reduced models for high enthalpy and plasma flows starting from detailed physics. The general objective is to apply existing chemical reduction techniques developed for turbulent combustion to plasma applications. The proposed research is an interesting alternative to physics based approaches as investigated by Munafo (Ref. 2-4).

The technique presented in this paper is based on Principal Component Analysis (PCA). PCA has been introduced recently as a mathematical method to retrieve the most active directions, or principal components, in a multivariate dataset by calculating the eigenvectors of the thermo-chemical state-space covariance matrix. Those eigenvectors form a low-dimensional basis on which the system is projected. Manifold Generated PCA (MG-PCA) uses PCA to reduce a physical model to a lower dimension. It allows for solving the transport equations for the principal variables only and reconstructs the remaining variables at each time step. Two different approaches of the method have been developed: a global one (MG-PCA) and a local one (MG-I-PCA). The local formulation divides the full dataset in clusters and applies MG-PCA on each of them. Different pre-processing techniques such as scaling and centering methods can be implemented and compared.

The current work is focused on the development and implementation of a MG-PCA based reduced model for plasma flows relying on combustion expertise. Optimal centering and scaling parameters will be investigated and compared to combustion cases. Preliminary results focus on the application of the reduction technique to data of 1D shock tube simulations for argon plasma. The model is validated with UTIAS shock tube experiments.

II. Physical modeling

A. Detailed chemistry modeling

Detailed physical models are needed to describe the plasma in thermo-chemical non-equilibrium. Two different types of models can be discerned:

- Multi-Temperature models (MT)
- Collisional-Radiative models (CR)

When considering MT models, the population of each internal energy state follows a Maxwell-Boltzmann distribution at a given temperature. To calculate the temperatures and the exchanged energy between all internal modes, conservation equations for thermochemical internal energy modes at non-equilibrium are added to the conventional conservation laws i.e. conservation of mass, momentum and total energy. When considering re-entry flows, the non-equilibrium phenomena and non-linear effects become important at the early stage of the trajectory phase, characterized by lower pressure conditions. The internal energy modes of all states will strongly be affected by these effects and be disturbed. A Maxwell-Boltzmann distribution can no longer be accounted for each state. Collisional-Radiative models do not rely on a Maxwell-Boltzmann distribution and give a more precise chemistry modeling than MT models.

Consequently, a collisional model has been chosen to model the ionizing shocks in argon plasma, as investigated during the thesis of Kapper (Ref. 5). A CR model describes the essence of chemistry, which are in this case the excitation and ionization processes in non-equilibrium conditions, allowing for non-Boltzmann distributions of bound electronic states. The main tasks of a CR model are to calculate the atomic state distribution function (ASDF), the effective conversion rates and the source terms for the energy equations.
The objective of his work is to investigate and model non-equilibrium phenomena behind the shock wave. The radiation part of the CR model is consequently not of a big importance. Radiation phenomena will be modeled by means of escape factors in a preliminary step.

The collisional model developed by Kapper (Ref. 5-7) for monatomic argon includes the 31 first excited levels of argon which contain most of the energy of the atom, and two ionized states of $Ar^+$. The influence of molecular $Ar^{++}$ has been neglected for this model. The number of species is thus 34 when counting in the electrons.

Four processes are taken into account: electron impact excitation represented by Eq. (1), excitation by impact with the ground state represented by Eq. (2), electron impact ionization represented by Eq. (3) and ionization through impact with the ground state represented by Eq. (4).

$$Ar(i) + e^- \leftrightarrow Ar(j) + e^-$$  \hspace{1cm} (1)

$$Ar(i) + Ar(1) \leftrightarrow Ar(j) + Ar(1)$$  \hspace{1cm} (2)

$$Ar(i) + e^- \leftrightarrow Ar^+ + e^- + e^-$$  \hspace{1cm} (3)

$$Ar(i) + Ar(1) \leftrightarrow Ar^+ + e^- + e^-$$  \hspace{1cm} (4)

The indices $i$ and $j$ represent the excitation levels with $i < j$. The rate coefficients for electron impact excitation are represented by $C_{ij}$, for excitation through impact with the ground state by $K_{ij}$. The rate coefficients for electron impact ionization are represented by $S_i$ and for atom impact ionization by $V_i$.

### Table 1. Excited levels of argon. Considered excited levels of ionized argon: level ($i$), energy ($E$) and degeneracy ($g$).

<table>
<thead>
<tr>
<th>$Ar$</th>
<th>$g_i$</th>
<th>$E_i$ (eV)</th>
<th>$Ar$</th>
<th>$g_i$</th>
<th>$E_i$ (eV)</th>
<th>$Ar^+$</th>
<th>$g_i$</th>
<th>$E_i$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.00</td>
<td>16</td>
<td>3</td>
<td>13.86</td>
<td>1</td>
<td>4</td>
<td>15.76</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>11.55</td>
<td>17</td>
<td>5</td>
<td>13.90</td>
<td>2</td>
<td>2</td>
<td>15.94</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>11.72</td>
<td>19</td>
<td>7</td>
<td>14.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>11.83</td>
<td>20</td>
<td>5</td>
<td>14.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>12.91</td>
<td>21</td>
<td>5</td>
<td>14.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>13.08</td>
<td>22</td>
<td>3</td>
<td>14.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>13.09</td>
<td>23</td>
<td>7</td>
<td>14.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>13.15</td>
<td>24</td>
<td>3</td>
<td>14.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>13.17</td>
<td>25</td>
<td>5</td>
<td>14.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>13.27</td>
<td>26</td>
<td>5</td>
<td>14.23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>13.28</td>
<td>27</td>
<td>7</td>
<td>14.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>13.30</td>
<td>28</td>
<td>1</td>
<td>14.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>13.33</td>
<td>29</td>
<td>3</td>
<td>14.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>13.48</td>
<td>30</td>
<td>3</td>
<td>14.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>13.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Steady-State formulation

The Euler flow equations present equations for continuity, momentum and energy. Different approaches can be considered. The first approach is a two fluid description (2F) of the plasma. It consists in writing down all equations for the heavy particles and electrons. By doing so, one obtains a system of $2 \times 3$ equations (continuity, momentum, energy). On the contrary, a one fluid (1F) approach can be obtained by taking all equations together. This approach is valid when the heavy particle velocity and the electron bulk one are approximately the same: $u_e \approx u_h \approx u$. 

3

American Institute of Aeronautics and Astronautics
The considered plasma is ionized and contains free electrons which will thermalize rapidly with each other. For a sufficiently collisional electron gas, a Maxwellian distribution can be assumed for their translational mode and an electron temperature \( T_e \) can be associated to them. In such a case the electron thermal energy is convected independently from the total energy. This is why an extra equation is needed to describe this electron energy evolution. This kind of model is called a two-temperature (2T) model. The two considered temperatures are \( T \), the heavy particle temperature and \( T_e \), the temperature of the electrons. In this case the 2T model will be combined with a 1F approach. The Euler flow equations will be expressed for each species independently and completed by an extra energy equation for the electrons.\(^{5,11}\)

The numerical code has been adapted to reproduce the UTIAS shock tube experiments. Those experiments can be seen as snapshots and are consequently not evolving in time but only in space. One obtains a 1D problem in space depending on \( x \):

\[
\nabla \cdot \mathbf{F} = \dot{\Omega}
\]

(5)

When expressing this for all species (with the index \( i \)) and adding an equation for the electron energy one finds:

\[
\frac{\partial}{\partial x} [\rho u y_i] = \omega_i
\]

(6)

\[
\frac{\partial}{\partial x} [\rho u^2 + p] = 0
\]

(7)

\[
\frac{\partial}{\partial x} \left[ \rho u \left( h + \frac{1}{2} u^2 \right) \right] = 0
\]

(8)

\[
\frac{\partial}{\partial x} [\rho u y_e e^e] = \Omega_e - p_e \frac{\partial}{\partial x} u
\]

(9)

In these expressions \( \rho \) is the total mass density, \( u \) the speed, \( y_i \) the mass fraction per species \( i \), \( \omega_i \) the source term for each species, \( p \) the pressure and \( h \) the enthalpy of the mixture. The variables related to the electrons are: \( y_e \) the mass fraction of the electrons, \( e^e \) the electronic specific energy, \( \Omega_e \) the electronic energy production and \( p_e \) the partial pressure of the electrons.

III. Principal Component Analysis

A. Principal Component analysis: Mathematical formulation

PCA offers a method to detect the most important dimensions within a multivariate dataset. It is based on an orthogonal transformation projecting the data on its principal dimensions or principal components. These components are the directions with the largest variance within the dataset. The global idea is thus to reduce the system by eliminating the less important dimensions and to build a reduced coordinate system based on those principal components.

To find the principal components one must solve an eigenvalue problem applied on the covariance matrix of the original data. The covariance matrix is given by Eq. (10).

\[
S = \frac{1}{(n-1)} X^T X
\]

(10)

In this expression \( X \) is a \( n \times Q \) matrix with \( n \) the number of observations, and \( Q \) the number of variables. This covariance matrix is the subject of an eigenvalue problem and can thus be decomposed its eigenvectors \( A \) and eigenvalues \( L \) as shown in Eq. (11).
\[ S = ALA^T \]  

(11)

The eigenvalues with the highest value indicate the directions with the most important variance. A new matrix \( A_q \) can thus be defined: it is the truncation of matrix \( A \), containing only the directions with the highest variance. The principal components can be retrieved by projecting this matrix \( A_q \) on the original dataset, see Eq. (12).

\[ Z_q = XA_q \]  

(12)

The relation in Eq. (12) can thus be inverted to have an approximated reconstruction of the original sample as shown in Eq. (13).

\[ \hat{X}_q = Z_qA_q^T \]  

(13)

Unfortunately, PCA has an important drawback. The error propagation during the reconstruction of the variables is important and will strongly affect the calculation of the source terms. This comes from the fact PCA distributes the error evenly over all state variables. This means very small and unimportant variables are affected by errors of the same magnitude as those accounted for the more important variables. As a result, MG-PCA has been developed to better control the error propagation during the reconstruction.

**B. Manifold Generated Principal Component Analysis**

MG-PCA or Manifold Generated PCA applies principal component analysis through the use of principal variables (PV's). Those principle variables are a subset of the space variables containing most of the variance of the system. As mentioned in the previous part, PCA introduces an important error propagation during the reconstruction of all variables. To better control this error propagation MG-PCA has been developed. The main idea is to reconstruct the non-transported variables, by using an optimal square and inverted subset of the eigenvector matrix \( A_q \), matrix \( B \):

\[ B = (A_q^T)^{-1} = X_q^{-1}Z_q \]  

(14)

The non-principal reconstructed variables, given by the vector \( \hat{X}(Q - q) \), are reconstructed by multiplying the vector containing the principal components, \( X(q) \), with this matrix \( B \) and the square subset of matrix \( A(Q - q) \):

\[ \hat{X}(Q - q) = X(q)BA(Q - q)_q^T \]  

(15)

This method has proven to be much better with respect to error propagation during the reconstruction as has been shown in the publication of Coussement (Ref. 12).

**C. Pre-processing techniques**

Before applying MG-PCA, prior actions can be undertaken to prepare the data, as presented by Parente (Ref. 13):

- Detect and remove the outliers
- Centering and scaling of the data

The outlier detection and removal will reduce the dataset and only subject the essential observation points to the PCA method. It has a positive effect on PCA as it reduces the number of directions to consider. To detect and discard the outliers, the so called Mahalanobis distance is used:

\[ D_M = (X - \bar{X})^T S^{-1} (X - \bar{X}) \]  

(16)
In this expression, $S^{-1}$ is the inverted covariance matrix, and $\bar{X}$ a matrix containing the average values of the original dataset. For a too large value of $D_M$, the corresponding observation is discarded.

Before applying PCA, it can be useful to apply some centering and scaling methods. These operations will optimize the PCA when it will be carried out. The centering operation will highlight the variables the most susceptible to variation. To do so, a scaling coefficient $\bar{x}_j$ is used:

$$\bar{x}_{\text{centered}} = x_j - \bar{x}_j$$

The scaling operation is a very important one. The data set contains variables with different units, temperatures and mass fractions for example, that can be scaled differently. The goal of the scaling operation is to set all variables to the same scale by using a scaling coefficient $d_j$. This way it will be easier for the PCA method to find correlations between the variables.

$$\bar{x}_{\text{scaled}} = \bar{x}_{\text{centered}}/d_j$$

There exist different scaling methods:

- **Auto scaling**: The standard deviation $s_j$ is taken as a scaling factor $d_j$. After scaling all elements of the dataset have a standard deviation equal to one.

- **Level scaling**: The scaling factors are defined as the mean value of the variables: $d_j = \bar{x}_j$. It is important to remove the outliers when using this type of scaling method to avoid the mean values are affected by wrong values.

- **Range scaling**: The difference between the maximum and minimum value is taken as scaling factor: $d_j = \max(\bar{x}_j) - \min(\bar{x}_j)$. Also in this case it is important to remove the outliers because they could wrongly influence the scaling coefficient.

- **Max scaling**: All values are bounded by zero and one after being normalized. This means they are all divided by their maximum value: $d_j = \max(\bar{x}_j)$.

- **PARETO scaling**: Every value is divided by the square root of its standard deviation: $d_j = s_j/\bar{x}_j$. The variance of every variable thus equals its standard deviation.

- **VAST (VAriable STability) scaling**: The standard deviation is also used for this scaling method: $d_j = s_j^2/\bar{x}_j$. Values with a small deviation are better taken into account.

### D. Principal variable extraction

Different methods exist to retain the principal variables within the dataset. Those methods are based on PCA as explained before. The principal variables are chosen in a way to have the best reconstruction. This paragraph will briefly describe these different PV-extraction methods as has been done in the publication of Parente (Ref. 14).

**B4 Forward method**: PCA is performed on the matrix containing the original dataset. The criterion to retain the principal variables depends on $l' = 0.7 \cdot \bar{l}$. In this expression $\bar{l}$ is the average eigenvalue size. If $Q_1$ components have an eigenvalue less than $l'$, the remaining $Q - Q_1$ components are analysed starting with the first component.

**B2 Backward method**: This method is similar to the B2 forward method. When analysing the remaining $Q - Q_1$ eigenvalues, the criterion $l'$ is applied on the last component.

**M2 Backward method**: PCA is performed on the original dataset, containing $Q$ variables, to evaluate the eigenvalues of the covariance matrix. A criterion is chosen to retain $q$ PCs scores in a matrix $Z_q$. The idea is to find a matrix $Z$ containing $m$ variables (with $q \leq Q$) which will reproduce $Z_q$ in the best way. To do so the M2 Backward method finds the subset of variables minimizing the following relation:

$$M_2 = \text{tr}(Z_q^T Z_q + Z^T Z - 2\Sigma)$$

6

American Institute of Aeronautics and Astronautics
In this expression $\Sigma$ is the matrix of singular values from the singular value decomposition of $\tilde{Z}^T Z_q$.

**McCabe criterion:** McCabe proposes 3 criteria to select the principal variables. These criteria are based on the fact principal components satisfy some optimal criteria. A subset of the original variables that optimizes one of these criteria is selected as a set of McCabe principle variables.

If $X_1$ and $X_2$ are subsets of the set of variables $X$ and the covariance matrix $S$ can be decomposed as:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

One can then define the partial covariance matrix:

$$S_{22,1} = S_{22} - S_{21} S_{11}^{-1} S_{12}$$

Which appears in the 3 selection criteria together with its eigenvalues $\delta_k$:

$$MC1 = \max |S_{11}| = \min |S_{22,1}| = \min \prod_{k=1}^{m} \delta_k$$

$$MC2 = \min \text{tr}(S_{22,1}) = \min \prod_{k=1}^{m} \delta_k$$

$$MC3 = \min |S_{22,1}|^2 = \min \prod_{k=1}^{m} \delta_k^2$$

**IV. Preliminary results**

MG-PCA has been implemented in a one dimensional code reproducing argon shock tube experiments. The physical model is described by a two temperature model in a one fluid approach, combined with a collisional model reproducing detailed chemistry after the shock, as described previously. It is a code reproducing thermo-chemical effects after an ionizing shock in space. These kind of reduction techniques have already been tested for combustion problems involving turbulent reactive flows. The objective of this paper is to develop similar techniques for plasma and to compare the results with respect to combustion. In a first instance the correct implementation of the PCA code will be tested by running a full simulation. Afterwards, the model will piecewisely be reduced by discarding one variable at the time. Different MG-PCA pre-processing techniques will be tested and compared.

**A. Detailing the shock structure**

The heavy particle and electron temperature profiles, respectively denoted by $T_h$ and $T_e$, are visualized in function of the distance after the shock. These temperature evolutions will be detailed to explain the physics behind the CR processes. The shock structure can be divided into five regions as represented on Fig. 1. The first region (region I) is characterized by a peak in electron temperature close after the shock. This temperature peak corresponds to the initial production of priming electrons. The dominating chemical reaction is given by Eq. (3) which corresponds to ionization of a species by impact with the ground state of argon. The second region (region II) is also particular because of its characterizing dip in electron temperature. The chemistry that was dominated by heavy particle reactions is now dominated by electron impact reactions as represented in Eq. (4). The temperature of the electrons keeps

![Figure 1. Shock structure.](image)

*Figure 1. Shock structure. Heavy particle and electron temperature in function of the distance from the shock. The shock is located at the y-axis of the plot.*
increasing in the third region (region III). Consequently more energy is transferred from the heavy particles to the electrons. The species are now exited to higher levels due to electron and heavy particle impacts as shown in Eq. (1) and (2). The electron avalanche occurs in the fourth region (region IV). Enough electrons are produced to dominate the temperature of the plasma which results in a decrease of the heavy particle temperature. After the electron avalanche there is complete thermalization between the heavy particles and the electrons (region V). The plasma reaches an equilibrium and can be characterized by a single temperature.

Table 2 represents experimental argon shock tube results carried out by the Institute for Aerospace Studies, University of Toronto (UTIAS). Those experiments have been considered to validate the shock tube simulation code used during this research project. The seven considered experimental test cases are very alike. However, two cases can be retained because of their particular conditions: Case 4 represents the experiment with the lowest speed (Ma 13) and Case 5 represents the experiment with the highest speed (Ma 17.2) together with a low pre-shock pressure. Case 1 can be considered as an intermediate case, and will be used as a verification test case.

### B. MG-PCA implementation and verification

A first simulation with the full detailed chemistry model is needed to retrieve the principal components by solving an eigenvalue problem on the obtained dataset. The flow problem contains 36 variables: 34 species, the heavy particle temperature \( T \), and the electron bulk temperature \( T_e \). When applying a PCA based reduction method, one needs to carefully select the pre-processing techniques which include centering and scaling parameters, and chose a suitable PV-extraction method. In a first instance we will use the optimum parameters for combustion cases. The best scaling method for combustion is PARETO scaling and the optimum principal variable extraction method is the B2 Backward extraction method. In a later stadium of the research a complete comparison of pre-processing techniques and PV-extraction methods will be carried out between the plasma and combustion cases. For now the optimum combustion parameters will be considered.

In a first instance, the correct implementation of the MG-PCA method will be verified. The verification consists in retaining all 36 variables as flow variables. Consequently, no reduction will be obtained. To compare the reduced model to the full model one visualizes the temperature evolution of the heavy particles and the electron bulk temperature after the ionizing shock. The results of this verification can be seen on Fig. 2. There is a perfect correlation between the reduced and the full model.

After the verification test case the number of retained variables has been decreased to obtain a reduced model of the detailed chemistry. The formulation based on the 36 flow variables has been reduced to a model considering 26 principal variables. The results of this simulation can be visualized on Fig. 3a. To obtain a further reduction another test case has been considered. Shock tube experiment 4 has been retained because of its low pre-shock speed (Ma 13) and average pressure (5.01 torr). When applying MG-PCA on this test case the number of retained variables could be reduced to 23 which corresponds to reduction of 36%. The result is shown on Fig. 3b.

Unfortunately the code blows up when lowering the number of retained variables below 23 for Case 4, and 26 for Case 1. It couldn’t reach equilibrium (region V on Fig. 1). The region of the discontinuity corresponds to the electron avalanche process after

**Table 2. Shock tube experiment case studies.**

<table>
<thead>
<tr>
<th>Case</th>
<th>( \text{Ma} )</th>
<th>( P \ [\text{torr}] )</th>
<th>( T \ [\text{K}] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15.9</td>
<td>5.14</td>
<td>293.6</td>
</tr>
<tr>
<td>2</td>
<td>16.1</td>
<td>5.15</td>
<td>295.9</td>
</tr>
<tr>
<td>3</td>
<td>16.5</td>
<td>5.12</td>
<td>296.6</td>
</tr>
<tr>
<td>4</td>
<td>13.0</td>
<td>5.01</td>
<td>296.6</td>
</tr>
<tr>
<td>5</td>
<td>17.2</td>
<td>2.06</td>
<td>297.8</td>
</tr>
<tr>
<td>6</td>
<td>13.6</td>
<td>5.12</td>
<td>297.0</td>
</tr>
<tr>
<td>7</td>
<td>14.7</td>
<td>4.08</td>
<td>297.8</td>
</tr>
</tbody>
</table>

Figure 2. Verification of the MG-PCA code. Verification of the method on shock tube case 1. Extraction method: B2 Backward method. Scaling method: PARETO.
the shock, as explained previously (region IV on Fig. 1). This region is characterized by a steep decrease of the heavy particle temperature together with a decrease in electron temperature. The chemical reactions involved are heavy particle and electron impact excitation (Eq. (1) and (2)). It can be concluded the reduced model needs more principal components to reproduce the more complex chemistry occurring in this region after the shock.

By now the reduction method has been applied to the full dataset which contained the mass fractions of all species and the two temperatures. Another way of proceeding is to leave out the temperatures from the dataset and to apply MG-PCA on the mass fractions only. Unfortunately, no further reduction has been obtained with this approach.

V. Conclusion

This work demonstrates the use of chemical reduction techniques on high enthalpy and plasma flows. It has been proven the MG-PCA based methods developed for combustion problems can be applied to other domains like plasma. The method has successfully been tested on a one dimensional code reproducing argon shock tube experiments. A reduced model based on 23 principal components has been derived from a detailed collisional chemistry model through the use of a global MG-PCA method. This corresponds to a reduction of 36 percent.

Future work will focus on the comparison of different pre-processing techniques and the principal variable extraction methods of MG-PCA. Also the local formulation, MG-local-PCA will be implemented and verified.

References


Parente, A. and Sutherland, J. “Principal component analysis of turbulent combustion data: Data pre-processing and manifold sensitivity”, *Combustion and Flame*, 2012.