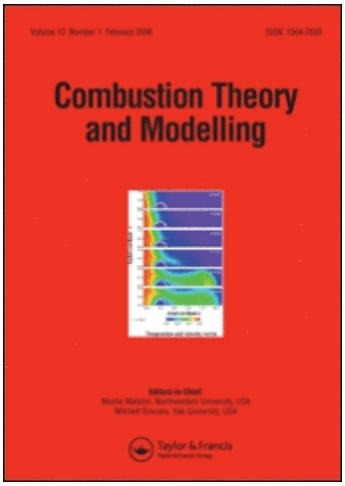

Combustion Theory and Modelling
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713665226

A greedy algorithm for species selection in dimension reduction of combustion chemistry

Varun Hiremath[a]; Zhuyin Ren[b]; Stephen B. Pope[a]
[a] Mechanical & Aerospace Engineering, Cornell University, Ithaca, NY, USA [b] ANSYS, Inc., Lebanon, NH, USA

Online publication date: 04 October 2010

To cite this Article Hiremath, Varun , Ren, Zhuyin and Pope, Stephen B.(2010) 'A greedy algorithm for species selection in dimension reduction of combustion chemistry', Combustion Theory and Modelling, 14: 5, 619 — 652
To link to this Article: DOI: 10.1080/13647830.2010.499964
URL: http://dx.doi.org/10.1080/13647830.2010.499964

Combustion Theory and Modelling
Vol. 14, No. 5, 2010, 619–652

A greedy algorithm for species selection in dimension reduction of combustion chemistry

Varun Hiremath[a*], Zhuyin Ren[b] and Stephen B. Pope[a]

[a]*Cornell University, Mechanical & Aerospace Engineering, Ithaca, NY 14853, USA;*
[b]*ANSYS, Inc., 10 Cavendish Court, Lebanon, NH 03766, USA*

(*Received 29 December 2009; final version received 5 June 2010*)

Computational calculations of combustion problems involving large numbers of species and reactions with a detailed description of the chemistry can be very expensive. Numerous dimension reduction techniques have been developed in the past to reduce the computational cost. In this paper, we consider the rate controlled constrained-equilibrium (RCCE) dimension reduction method, in which a set of constrained species is specified. For a given number of constrained species, the 'optimal' set of constrained species is that which minimizes the dimension reduction error. The direct determination of the optimal set is computationally infeasible, and instead we present a greedy algorithm which aims at determining a 'good' set of constrained species; that is, one leading to near-minimal dimension reduction error. The partially-stirred reactor (PaSR) involving methane premixed combustion with chemistry described by the GRI-Mech 1.2 mechanism containing 31 species is used to test the algorithm. Results on dimension reduction errors for different sets of constrained species are presented to assess the effectiveness of the greedy algorithm. It is shown that the first four constrained species selected using the proposed greedy algorithm produce lower dimension reduction error than constraints on the major species: CH_4, O_2, CO_2 and H_2O. It is also shown that the first ten constrained species selected using the proposed greedy algorithm produce a non-increasing dimension reduction error with every additional constrained species; and produce the lowest dimension reduction error in many cases tested over a wide range of equivalence ratios, pressures and initial temperatures.

Keywords: RCCE; greedy algorithm; optimal species; PaSR; dimension reduction

Nomenclature

Roman

h	enthalpy
p	pressure
n_e	number of elements
n_s	number of species
n_{rs}	number of represented species
n_{us}	number of unrepresented species
n_r	reduced dimension ($n_r = n_{rs} + n_e$)
$\mathbf{z}$	species specific moles
$\boldsymbol{r}$	reduced representation

*Corresponding author. Email: vh63@cornell.edu

ISSN: 1364-7830 print / 1741-3559 online

DOI: 10.1080/13647830.2010.499964
http://www.informaworld.com

z^r	specific moles of represented species
$z^{u,e}$	specific moles of elements in the unrepresented species
$\boldsymbol{S}$	chemical production rates
$\boldsymbol{R}(z, t)$	reaction mapping after time t starting from z
$\boldsymbol{B}$	constraint matrix
$\boldsymbol{r}(0)$	reduced representation at time $t = 0$
$z^{CE}(\boldsymbol{r})$	constrained-equilibrium composition at $\boldsymbol{r}$
$z^{DR}(t)$	reaction mapping, $\boldsymbol{R}(z^{DR}(0), t)$
$\boldsymbol{r}^{DR}(t)$	reduced representation after time t
N	number of test compositions
N_P	number of particles in the PaSR
Δt	reaction time step
T	PaSR initial temperature

Greek

$\boldsymbol{\Phi}$	set of all species
$\boldsymbol{\Phi}^r$	set of represented species
$\boldsymbol{\Phi}^u$	set of unrepresented species
$\boldsymbol{\Phi}^{ud}$	set of unrepresented determined species
$\boldsymbol{\Phi}^{uu}$	set of unrepresented undetermined species
$\boldsymbol{\Phi}^{opt}$	optimal set of represented species
$\boldsymbol{\Phi}^g$	set of good represented species generated by the greedy algorithm
$\epsilon(t, \boldsymbol{\Phi}^r)$	dimension reduction error in the reaction mapping, $\boldsymbol{R}(z^{DR}(0), t)$
$\epsilon(\boldsymbol{\Phi}^r)$	error used for defining the optimal set of species
ϕ	equivalence ratio
τ_{res}	PaSR residence time
τ_{mix}	PaSR mixing time scale
τ_{pair}	PaSR pairing time scale

1. Introduction

Modern chemical mechanisms for real fuels typically involve hundreds of species and thousands of reactions [1, 2]. Computational calculations of reactive flows involving such fuels with detailed chemistry are prohibitive even on a distributed computing platform.

Numerous techniques have been developed in the past to reduce the computational cost of implementing combustion chemistry. These include:

(1) **Skeletal mechanisms**: A skeletal mechanism consists of a subset of the species and reactions from the detailed mechanism. Many methods have been developed to systematically generate skeletal mechanisms from detailed mechanisms, such as the Directed Relation Graph (DRG) [3], DRG with error propagation (DRGEP) [4] and Simulation Error Minimization Connectivity Method (SEM-CM) [5].

(2) **Reduced chemical mechanisms (based on QSSA)**: The quasi-steady-state approximation (QSSA) [6, 7] has been widely applied to develop reduced chemical mechanisms. The QSSA method involves the identification of QSS species in the system, whose net rate of production is assumed to be zero, thereby reducing the governing differential

equation for the QSS species to an algebraic relation. These algebraic relations are used to eliminate the QSS species from the system.

(3) **Dimension reduction methods**: Another class of dimension reduction techniques is based on the observation that chemical systems involve reactions with a wide range of time scales. As a consequence, reaction trajectories are attracted to lower dimensional attracting manifolds in the composition space. Computations can be performed in a reduced space by identifying such low-dimensional manifolds, thereby reducing the overall computational cost. Methods based on this idea include rate-controlled constrained equilibrium (RCCE) [8, 9], computational singular perturbation (CSP) [10], intrinsic low-dimensional manifolds (ILDM) [11], trajectory-generated low-dimensional manifolds (TGLDM) [12], invariant constrained equilibrium-edge pre-image curve (ICE-PIC) [13] and one-dimensional slow invariant manifold (1D SIM) [14].

(4) **Storage retrieval methods**: In these approaches, combustion chemistry computations are stored in a table, and are used to build inexpensive approximate solutions at a later stage of computation to reduce the overall cost. Methods based on this idea include the structured look-up tabulation [15], repro-modelling [16], artificial neural network (ANN) [17], *in situ* adaptive tabulation (ISAT) [18, 19] and piecewise reusable implementation of solution mapping (PRISM) [20].

In recent times, combined methodologies have also been developed, wherein reduced reaction mechanisms or dimension reduction methods are used in conjunction with storage/retrieval methods, such as ISAT-QSSA [21], ISAT-RCCE [22], and recently ICE-PIC with ISAT [23].

In reactive flow calculations, the species concentrations are governed by two processes: chemical reaction and transport. We consider the important class of solution methods in which a splitting scheme is used, where the chemical reaction and transport processes are accounted for in two separate steps. In the computational modelling of the turbulent combustion using PDF methods [24], the fluid composition within the solution domain is represented by a large number of particles. In a full-scale PDF calculation, more than two million particles may be used, and the solution can advance for more than 5000 time steps, leading to approximately 10^9 particle-reaction sub-steps. If such a calculation involves 100 species with the chemistry represented by a detailed mechanism, then at each reaction sub-step, 100 coupled ODEs need to be solved to determine the species concentrations, which can be very expensive and computationally prohibitive. Instead a dimension reduction method (such as RCCE or ICE-PIC) integrated with ISAT can be used to perform the reactive flow calculations in terms of say 10 represented species; where the combined reduction-storage methodology determines and tabulates (*in situ*) the reduced space in terms of the 10 represented species based on the detailed mechanism.

In a reactive system, the reaction trajectories rapidly approach a hierarchy of attracting manifolds of decreasing dimensions and the reactive system's slow dynamics is well approximated by these low-dimensional attracting slow invariant manifolds (SIMs) in the reactive space. Numerous dimension reduction methods have been developed which exploit this property to represent the chemistry using a reduced set of variables. Here we focus on the RCCE and ICE-PIC dimension reduction methods, which have been successfully implemented and used in many reactive flow computations [23, 25–27]. The slow invariant manifold is approximated in the RCCE method by the constrained-equilibrium manifold (CEM) constructed using thermodynamic concepts, and in the ICE-PIC method by the invariant constrained-equilibrium edge (ICE) manifold which is a collection of all the reaction trajectories emanating from the CEM edge. Recent studies [14] have shown that the

consideration of the topology of classical thermodynamic functions may not provide a good approximation to the 1D SIM. However, in RCCE and ICE-PIC we typically use more than one constrained species (i.e. work in higher dimensions), and the errors in the approximation generally decrease in higher dimensions. We also note that the method described by [14] has yet to be applied in many dimensions.

The RCCE dimension reduction method (originally proposed in [8]) is based on the assumption that in a reactive system, the reaction trajectories relax to the complete equilibrium with a rate determined solely by the slow reactions, while the fast reactions tend to locally equilibrate the system subject to the constraints imposed by the slow reactions. Thus, the system reaches complete equilibrium by evolving through a sequence of constrained-equilibrium states on the CEM. As a result, only the rate equations of the slowly changing constraints need to be solved, though a different approach is used in our implementation of RCCE as described in Section 2.3. The constrained-equilibrium state can be determined (locally) by computing the state corresponding to the maximum entropy subject to the given set of constraints.

The very first step involved in the application of the RCCE (and ICE-PIC) method is the specification of the constraints for dimension reduction. The conservation of elements form the most basic time-independent constraints. Among the time-dependent constraints, a literature review [26–28] shows that the most commonly used constraints in RCCE include general linear constraints on the total number of moles; moles of active valence (AV); moles of free oxygen (FO), etc.

In our implementations of RCCE and ICE-PIC, which are integrated with ISAT [13, 22, 23, 29], the concentration of a specified set of constrained species form the constraints. The specification of good constraints is crucial for the accuracy of dimension reduction, but there are no systematic methods available to select good constraints in an automated way.

Ideally one wants to find the smallest set of constrained species that yields the dimension reduction errors below a specified tolerance; or one wants to find the 'optimal' set of constrained species that minimizes the dimension reduction errors for a fixed number of constrained species. This is a very hard problem, so we aim to devise an algorithm to select a 'good' set of constrained species, i.e. a near-optimal set of species which produces low dimension reduction errors.

The proposed method works by considering a computationally inexpensive representative test problem (the partially-stirred reactor (PaSR)), and directly measuring the dimension reduction error. A 'good' set of constrained species is selected by employing a greedy algorithm; which selectively adds 'good' species to the set (initially empty) one at a time to minimize the dimension reduction error at each stage. Here we consider the application of this method for RCCE, but the methodology developed is also applicable to ICE-PIC. The SEM-CM [5] method used for developing skeletal mechanisms employs a similar idea for identifying the species to be retained in the skeletal mechanism. In the SEM-CM method, the mechanism building procedure is started from a set of specified important species, and then species are added (based on ranking) one at a time until the simulation error using the skeletal mechanism becomes smaller than a required threshold.

The outline of the remainder of the paper is as follows. In Section 2, we develop a mathematical representation for a gas-phase reacting system. We then describe the RCCE dimension reduction procedure and define the dimension reduction errors involved in this procedure. In Section 3, we propose a definition for the 'optimal' set of species and later present the greedy algorithm for selecting a 'good' (near-optimal) set of species. In Section 4 we present results obtained using the greedy algorithm for the test case of a

partially-stirred reactor (PaSR) with methane premixed combustion. Finally, in Section 7 we draw conclusions based on the presented results.

2. Representation of chemistry

We consider a reacting gas-phase mixture consisting of n_s chemical species, composed of n_e elements. The set of all species is denoted by $\mathbf{\Phi}$. The thermochemical state of the mixture (at a given position and time) is completely characterized by the pressure p, the mixture enthalpy h, and the n_s-vector z of specific moles of the species. To simplify the exposition, we consider an adiabatic and isobaric system with h and p taken to be given constants, and so the thermochemical state is given by z.

It is useful to consider the species composition z to be an n_s-vector or a point in the n_s-dimensional *full composition space*.

2.1. *Reaction trajectories*

Due to chemical reactions, the chemical composition z evolves in time according to the following set of ordinary differential equations (ODEs)

$$\frac{dz(t)}{dt} = \boldsymbol{S}(z(t)), \tag{1}$$

where $\boldsymbol{S}$ is the n_s-vector of chemical production rates determined by the chemical mechanism used to represent the chemistry.

The *reaction mapping*, $\boldsymbol{R}(z, t)$, is defined to be the solution to Equation (1) after time t starting from the initial composition z. In this work, the *reaction mapping* is computed by numerically integrating the set of ODEs (1) using DDASAC [30].

2.2. *Dimension reduction*

The dimension reduction methods that we are interested in are methods based on low-dimensional manifolds, and in particular the RCCE and ICE-PIC dimension reduction methods. In this section we briefly describe the notation used in these dimension reduction methods; detailed descriptions are provided in [13, 23].

In RCCE and ICE-PIC, the set of species $\mathbf{\Phi}$ is decomposed as $\mathbf{\Phi} = \{\mathbf{\Phi}^r, \mathbf{\Phi}^u\}$, where $\mathbf{\Phi}^r$ is the set of *represented* species with cardinality n_{rs}, and $\mathbf{\Phi}^u$ is the set of *unrepresented* species with cardinality n_{us}, where $n_{rs} + n_{us} = n_s$ and $n_{rs} < n_s - n_e$.

The *reduced representation* of the species composition is denoted by $\boldsymbol{r} \equiv \{z^r, z^{u,e}\}$, where z^r is n_{rs}-vector of specific moles of represented species, $\mathbf{\Phi}^r$; and $z^{u,e}$ is an n_e-vector giving the specific moles of the elements in the unrepresented species, $\mathbf{\Phi}^u$ (for atom conservation). Thus, $\boldsymbol{r}$ is a vector of length $n_r = n_{rs} + n_e$ in the *reduced composition space*, and the dimension of the system is reduced from n_s to $n_r < n_s$. At any time t, the reduced representation, $\boldsymbol{r}(t)$, is related to the full composition, $z(t)$, as

$$\boldsymbol{r}(t) = \boldsymbol{B}^T z(t), \tag{2}$$

where $\boldsymbol{B}$ is a constant $n_s \times n_r$ matrix which can be determined for a specified set of represented species.

2.3. *Steps involved in dimension reduction*

In this section we briefly describe the four main steps involved in our implementation of RCCE. Since our implementation of RCCE is integrated with ISAT, some of the steps in our implementation of RCCE differ from other work found in the literature; those steps are highlighted and justified.

(1) The first important step in the application of the RCCE dimension reduction method is the *selection* of the set of represented (constrained) species, $\boldsymbol{\Phi}^r$. For a given set of represented species, $\boldsymbol{\Phi}^r$, the reduced representation is given as $\boldsymbol{r} \equiv \{z^r, z^{u,e}\}$.
Alternatively, in many of the RCCE implementations [26, 27] general linear constraints on species are specified. In our implementation of RCCE, to simplify the user interface and specification of constraints, we use the species specific moles of the represented species as the constraints.

(2) The next step is the *species reconstruction*, i.e. given a reduced representation $\boldsymbol{r}(0)$ at time $t = 0$, reconstruct an estimate of the full composition denoted by $z^{DR}(0)$.
In RCCE the *species reconstruction* is performed by computing the constrained-equilibrium composition for the given constraints. In our implementation of RCCE, the constrained-equilibrium composition is computed using the CEQ [31] code, with the constraints given by the reduced representation $\boldsymbol{r}$. The constrained-equilibrium composition at $\boldsymbol{r}$ is denoted by $z^{CE}(\boldsymbol{r})$. So the reconstructed composition in RCCE is given as

$$z^{DR}(0) = z^{CE}(\boldsymbol{r}(0)). \tag{3}$$

(3) The next step is to obtain the *reaction mapping*. Starting from the reconstructed composition, $z^{DR}(0)$, the set of ODEs (1) is integrated numerically in the full space using DDASAC to obtain the reaction mapping after time t, denoted by $z^{DR}(t)$ as shown in Figure 1.
An alternative approach for the RCCE method as suggested in [9] and also used in [26, 27] is to integrate the rate equations for the constraint potentials, which is more economical than integrating the ODEs (1) directly. In our implementation of RCCE, we chose the latter approach for the ease of integrating the RCCE dimension reduction method with ISAT, which is discussed in more detail in [22].

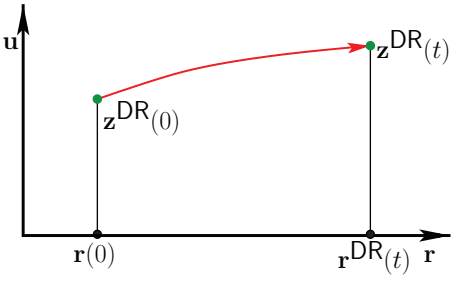

Figure 1. Sketch of the composition space (indicated by represented $\boldsymbol{r}$ and unrepresented $\boldsymbol{u}$ axes) showing the four steps involved in the dimension reduction procedure using RCCE and ICE-PIC: (i) the initial reduced composition is represented by $\boldsymbol{r}(0)$; (ii) the reconstructed composition at $\boldsymbol{r}(0)$ is represented by $z^{DR}(0)$; (iii) the reaction mapping starting from $z^{DR}(0)$ after time t is represented by $z^{DR}(t)$; (iv) the reduced composition after time t is represented by $\boldsymbol{r}^{DR}(t)$.

(4) The final step involved in the dimension reduction method is *reduction*, i.e. from the obtained *reaction mapping* after time t, $\boldsymbol{z}^{DR}(t)$, compute the *reduced representation* denoted by $\boldsymbol{r}^{DR}(t)$ (shown in Figure 1), given as:

$$\boldsymbol{r}^{DR}(t) = \boldsymbol{B}^T \boldsymbol{z}^{DR}(t). \tag{4}$$

To summarize, the key steps involved in the RCCE dimension reduction method are

(1) **Selection**: Identifying good constraints or the set of represented species, $\boldsymbol{\Phi}^r$, for dimension reduction.
(2) **Species reconstruction**: Given the constraints, $\boldsymbol{r}(0)$, reconstructing the full composition, $\boldsymbol{z}^{DR}(0)$.
(3) **Reaction mapping**: Starting from the *reconstructed composition* $\boldsymbol{z}^{DR}(0)$, computing the *reaction mapping* after time t in the full composition space $\boldsymbol{z}^{DR}(t)$.
(4) **Reduction**: From the *reaction mapping* $\boldsymbol{z}^{DR}(t)$, obtaining the *reduced representation* $\boldsymbol{r}^{DR}(t)$ after time t.

The ICE-PIC dimension method also involves the same four aforementioned steps, with the only difference being in the *species reconstruction* step, wherein the reconstructed composition is defined based on the invariant constrained-equilibrium (ICE) manifold [13, 23]. The remaining three steps: *selection*, *reaction mapping* and *reduction* are identical to the steps in RCCE.

Among these steps, the *selection* of the represented (constrained) species is an important step as the errors involved in the remaining three steps implicitly depend on the choice of the represented species, $\boldsymbol{\Phi}^r$. As also mentioned in [26, 27], identification of appropriate constraints is essential for the accuracy of the RCCE dimension reduction method. In the following sections, we develop an automated algorithm to select a 'good' set of represented species, $\boldsymbol{\Phi}^r$, for the accurate implementation of RCCE and ICE-PIC dimension reduction methods.

2.4. *Partially-stirred reactor*

In methods to develop QSSA based reduced mechanisms, it is useful to consider a range of test cases both to identify QSS species and to validate the resulting reduced mechanisms.

Here, we are interested in applying RCCE and ICE-PIC methods to LES/PDF calculations, for which the partially-stirred reactor (PaSR) is a (computationally cheaper) representative test case. We can vary the pressure p, temperature T, and the time step Δt to be representative of conditions in an LES/PDF calculation.

In this study, we consider the test case of a partially-stirred reactor (PaSR) involving premixed combustion of a methane–air mixture. A description of the PaSR is given in [18]; here we list only the important parameters involved.

There are two inflowing streams: a premixed stream of stoichiometric methane–air mixture at 600 K; and a pilot stream consisting of the adiabatic equilibrium products of a stoichiometric methane–air mixture at 2375 K (corresponding to an unburnt gas temperature of 600 K). The mass flow rates of these streams are in the ratio 0.95:0.05. Initially ($t = 0$), all particle compositions are set to be the pilot-stream composition. The pressure is atmospheric throughout.

Other important parameters involved are: number of particles, $N_P = 100$; residence time, $\tau_{res} = 10$ ms; mixing time scale, $\tau_{mix} = 1$ ms; pairing time scale, $\tau_{pair} = 1$ ms; time step = 0.1 ms (involving three fractional sub-steps of mixing, reaction and mixing); and reaction time step $\Delta t = 0.033$ ms. The PaSR is run for 3400 time steps, each involving three sub-steps over 100 particles, leading to more than 10^6 particle-sub-steps.

In this study, the GRI-Mech 1.2 mechanism involving 31 species is used to describe the methane combustion. The species involved are

$$\begin{aligned}\mathbf{\Phi} = \{&H_2, H, O_2, OH, H_2O, CH_3, CH_4, CO, CO_2, CH_2O, C_2H_4, O, HO_2,\\ &H_2O_2, C, CH, CH_2, CH_2(S), HCO, CH_2OH, CH_3O, CH_3OH, C_2H,\\ &C_2H_2, C_2H_3, C_2H_5, C_2H_6, HCCO, CH_2CO, HCCOH, N_2\}.\end{aligned} \tag{5}$$

To give some idea about the state of the PaSR, scatter plots of species specific moles of CH_4 and CO (retrieved from 10 selected particles from the PaSR) versus the temperature are shown in Figures 2 and 3, respectively. We see that the CH_4 concentration drops with temperature, as more and more CH_4 reacts to form products. The concentration of CO, on the other hand, increases with temperature and reaches a maximum at around 2000 K.

In our implementation of PaSR, computations can be performed using the full set of species, $\mathbf{\Phi}$, in the full composition space (without any dimension reduction) or using a smaller set of represented species, $\mathbf{\Phi}^r$, with one of the dimension reduction methods – RCCE or ICE-PIC. For a given test case, PaSR calculations are performed with and without dimension reduction, and the compositions obtained with the two approaches are compared to estimate the errors involved in dimension reduction.

In the next section we define the various errors involved in the dimension reduction steps. Subsequently, based on these definitions of error, we propose a definition of an 'optimal'

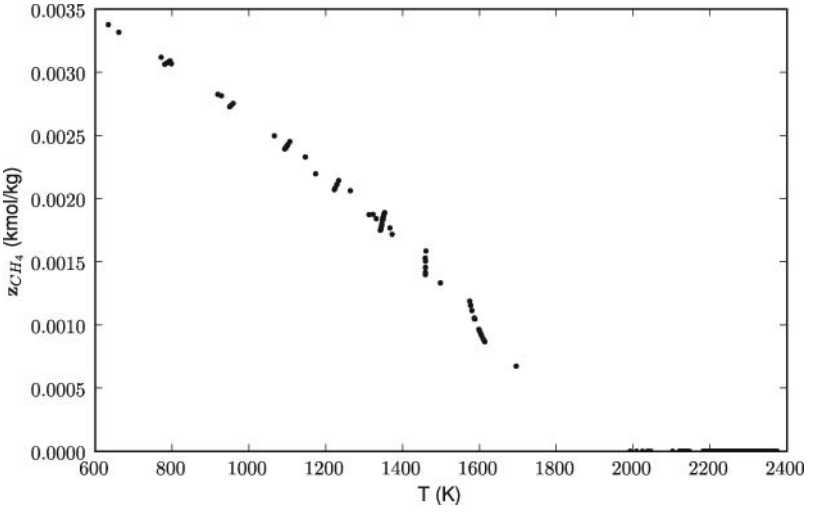

Figure 2. Scatter plot of the specific moles of CH_4, z_{CH_4} (retrieved from 10 selected particles from the PaSR) versus temperature, T, obtained from 3400 time steps of 0.1 ms each in the PaSR for a stoichiometric methane premixed combustion at atmospheric pressure and an initial temperature of 600 K.

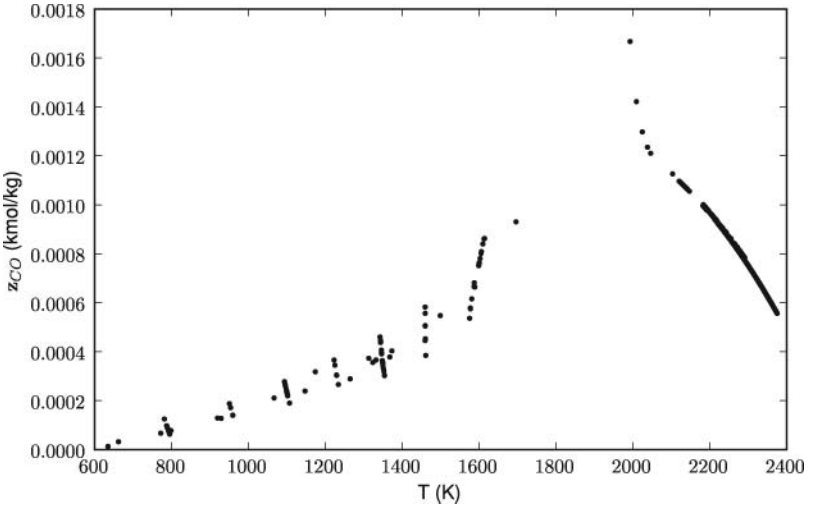

Figure 3. Scatter plot of the specific moles of CO, z_{CO} (retrieved from 10 selected particles from the PaSR) versus temperature, T, obtained from 3400 time steps of 0.1 ms each in the PaSR for a stoichiometric methane premixed combustion at atmospheric pressure and an initial temperature of 600 K.

set of represented species and present an algorithm to select a 'good', near-optimal set of represented species.

2.5. *Dimension reduction errors*

In this section we define the various errors involved in the dimension reduction process and describe the method employed to measure these errors using the PaSR.

Given a composition, $\boldsymbol{z}(0)$, in the full composition space, the *reaction mapping*, $\boldsymbol{R}(\boldsymbol{z}(0), t)$ (for $t \geq 0$) is more concisely denoted by $\boldsymbol{z}(t)$ (see Figure 4).

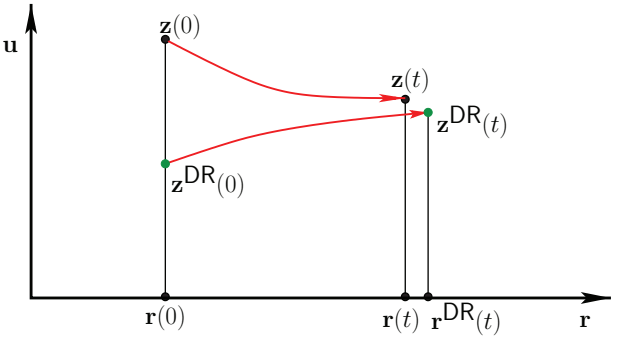

Figure 4. Sketch of the steps involved in the computation of dimension reduction errors. $\boldsymbol{z}(0)$ and its reaction mapping after time t, $z(t)$, are given compositions in the full space. For a specified set of represented species, $\boldsymbol{\Phi}^r$, the reduced representation at $\boldsymbol{z}(0)$ is denoted by $\boldsymbol{r}(0)$. The reconstructed composition at $\boldsymbol{r}(0)$ is denoted by $z^{DR}(0)$. The reaction mapping from $z^{DR}(0)$ after time t is denoted by $z^{DR}(t)$, and the reduced composition at $z^{DR}(t)$ is denoted by $\boldsymbol{r}^{DR}(t)$.

For a given set of *represented* species, $\mathbf{\Phi}^r$, the *reduced* representation of the full composition, $z(0)$, is denoted by $\boldsymbol{r}(0)$ and is obtained by performing the *reduction* using (2) as

$$\boldsymbol{r}(0) = \boldsymbol{B}^T z(0). \tag{6}$$

At $\boldsymbol{r}(0)$, the reconstructed composition using a dimension reduction method is denoted by $z^{DR}(0)$. Starting from the reconstructed composition, the *reaction mapping*, $\boldsymbol{R}(z^{DR}(0), t)$, in the full composition space is more concisely denoted by $z^{DR}(t)$ (see Figure 4).

Now for a representative test problem, to estimate the errors incurred using a dimension reduction method, a number of test compositions are selected in the full space. Let the number of test compositions used be denoted by N. We perform a PaSR computation in the *full composition space* (without dimension reduction) and then pick N distinct test compositions in the full space denoted by $z^{(n)}(0)$, for $n = 1, \ldots, N$; and their corresponding *reaction mappings*, after a fixed constant time t, are denoted by $z^{(n)}(t)$.

At the N chosen compositions, $z^{(n)}(0)$, in the full space, and for a given set of *represented* species, $\mathbf{\Phi}^r$, we denote the corresponding reduced representations by $\boldsymbol{r}^{(n)}(0)$, the reconstructed compositions by $z^{DR(n)}(0)$, and the reaction mappings by $z^{DR(n)}(t)$.

Note that, given $z(0)$ and t, $z^{DR}(t)$ depends on the specification of the represented species, $\mathbf{\Phi}^r$. As needed, we show this dependence explicitly by the notation $z^{DR}(t, \mathbf{\Phi}^r)$.

At this stage, we define the error in the *reaction mapping* obtained after time t starting from the reconstructed composition to be

$$\epsilon(t, \mathbf{\Phi}^r) = \frac{[z^{DR(n)}(t, \mathbf{\Phi}^r) - z^{(n)}(t)]_{rms}}{[z^{(n)}(t)]_{rms}}, \tag{7}$$

where the operator $[\,]_{rms}$ is defined by, for example,

$$[z^{(n)}(t)]_{rms} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} ||z^{(n)}(t)||^2}, \tag{8}$$

where $||z||$ denotes the 2-norm.

In particular we have two important errors in the dimension reduction method corresponding to $t = 0$ and $t = \Delta t$:

(1) **Species reconstruction error**: This is the error in reconstructing the full composition given a reduced composition $\boldsymbol{r}(0)$ at $t = 0$ and is given by Equation (7) as $\epsilon(0, \mathbf{\Phi}^r)$.
(2) **Reaction mapping error**: This is the error in the reaction mapping obtained after time step Δt (reaction time step) starting from the reconstructed composition and is equal to $\epsilon(\Delta t, \mathbf{\Phi}^r)$.

Both the species reconstruction and reaction mapping errors depend on the choice of represented species, $\mathbf{\Phi}^r$, and the goal of this work is to identify a 'good' set of represented species which reduces these errors.

3. Selection of optimal species

3.1. *Optimal species*

In the previous section we looked at the various errors involved in the dimension reduction method, and with the goal of reducing these errors for accurate implementation of dimension reduction in combustion chemistry, here we propose a definition for optimal species based on the RCCE dimension reduction method.

The definition of the optimal set of species is based on the dimension reduction error, $\epsilon(t, \mathbf{\Phi}^r)$. We consider either the species reconstruction error, $\epsilon(0, \mathbf{\Phi}^r)$, or the reaction mapping error, $\epsilon(\Delta t, \mathbf{\Phi}^r)$; and the error used for defining the optimal set is concisely denoted by $\epsilon(\mathbf{\Phi}^r)$.

For a given definition of the error $\epsilon(\mathbf{\Phi}^r)$, and a given value of n_{rs}, the set of n_{rs} represented species, $\mathbf{\Phi}^r$, which minimizes the error, $\epsilon(\mathbf{\Phi}^r)$, is defined to be the *optimal* set of species and is denoted by $\mathbf{\Phi}^{opt}$.

The optimal set of species, $\mathbf{\Phi}^{opt}$, (by definition) produces the minimum error:

$$\epsilon(\mathbf{\Phi}^{opt}) = \min_{\mathbf{\Phi}^r} \epsilon(\mathbf{\Phi}^r). \tag{9}$$

3.2. *Objectives of species selection*

Ideally one wants to find the *optimal* set of represented species for implementing dimension reduction. For a given value of n_{rs}, a simple brute-force method for selecting the *optimal* set of species is to form all possible $\binom{n_s}{n_{rs}}$ number of sets of the represented species, and then compute the error, $\epsilon(\mathbf{\Phi}^r)$, for each of the sets. The set of species producing the minimum error is the required optimal set. Such a brute-force method involves $O\binom{n_s}{n_{rs}}$ number of computations, which can become very expensive even at small values of n_{rs} and n_s (for example, for $n_s = 30$ and $n_{rs} = 10$, over 30 million sets of represented species can be formed) and hence this brute-force method is in general impracticable to use.

An alternative approach is to use a *greedy algorithm*. A greedy algorithm proceeds in stages, making a locally optimal choice at each stage to find a near-optimal solution [32]. Greedy algorithms are shortsighted in their approach, making one greedy choice at a time without worrying about the consequences of such a choice in the future. In other words, a greedy algorithm never reconsiders its choices. Greedy algorithms are not guaranteed to give the optimal solutions, but provide good solutions for many mathematical problems.

A 'good' set (initially empty) of represented species denoted by $\mathbf{\Phi}^g$ is formed in stages using a greedy algorithm by selecting at each stage the species whose addition to the set produces the minimum dimension reduction error.

The idea is, for a given value of n_{rs}, to select a set of n_{rs} 'good' species, proceed in n_{rs} stages, from $1, 2, \ldots, n_{rs}$, selecting the best species at each stage, i.e. the species which minimizes the error. So, at stage 1, pick the first best species from $\mathbf{\Phi}$ corresponding to the minimum error. Next, at stage 2 pick the next best species from the remaining set of species which minimizes the error, and continue until n_{rs} species are selected.

At each stage S (for $S = 1, \ldots, n_{rs}$) of this algorithm, $(n_s - S + 1)$ number of computations are performed. Overall in n_{rs} stages only $O(n_s\, n_{rs})$ number of computations are performed and hence this method is economical.

Moreover, in the implementations of RCCE and ICE-PIC it is often desirable to start working with a given set of represented species, and if required to add more species to the existing set. For such a purpose, the greedy algorithm is ideal, as it selects the best *available* represented species from the remaining set of species.

In the next section we formally describe this automated greedy algorithm, and then present results.

3.3. *Greedy algorithm for species selection*

Notation

Here we define certain terms and quantities used to describe the algorithm:

- **Determined species**: In a given chemical system with a specified number of moles of elements, the species whose concentration can be determined by atom conservation alone are called the determined species. Obviously such species are not good choices for represented species as they are already determined.
- **Unrepresented determined species (Φ^{ud})**: For a given set of represented species, Φ^r, the set of other species whose concentration can be determined by element conservation alone are called the unrepresented determined species and are denoted by Φ^{ud}. (There may be no such species, in which case Φ^{ud} is the empty set.)
- **Unrepresented undetermined species (Φ^{uu})**: The set of unrepresented species which are not determined are called the unrepresented undetermined species and are denoted by Φ^{uu}. (If there are no such species, i.e. all the species are either represented or determined, then Φ^{uu} is the empty set.)

The greedy algorithm presented in the next section selects at each stage a good species (producing minimum dimension reduction error) from the set of unrepresented undetermined species, Φ^{uu}, to form a good set of represented species, Φ^g.

Greedy algorithm

The greedy algorithm is described below for finding the entire species ordering, i.e. until the set Φ^{uu} is empty, based on the defined error $\epsilon(\Phi^r)$.

(1) The n_s species in set Φ are assigned indices $1, 2, \ldots, n_s$ in an arbitrary order. We use the notation *species k* to denote the species with index k.
(2) The algorithm proceeds in S stages, numbered $1, 2, \ldots, S$ where S is at most $n_s - n_e$.
(3) At the end of the jth stage, there are j 'good' represented species selected by the algorithm, which form the represented set Φ^g_j.
(4) Initially, before the beginning of stage 1, the set Φ^g_0 is initialized to an empty set.
(5) At the beginning of the jth stage, based on the set of represented species from the previous stage, Φ^g_{j-1}, the set Φ^{uu}_j of unrepresented undetermined species is identified. If this set is empty, then the algorithm terminates. Let the set of indices of species in Φ^{uu}_j be denoted by $\boldsymbol{I}^{uu}_j$.
(6) In the jth stage, another *species* m_j for $m_j \in \boldsymbol{I}^{uu}_j$ is identified to be added to Φ^g_{j-1} to form Φ^g_j.

- For each *species k* ($k \in \boldsymbol{I}^{uu}_j$), $\Phi^g_{j,k}$ denotes the union of Φ^g_{j-1} and *species k* from Φ^{uu}_j.
- For each set $\Phi^g_{j,k}$ the defined error $\epsilon(\Phi^r)$ is computed as

$$\epsilon_{jk} = \epsilon(\Phi^g_{j,k}). \tag{10}$$

(7) The selected species $m_j \in \boldsymbol{I}_j^{uu}$ is that which minimizes the error, i.e.

$$\epsilon_{jm_j} \leq \epsilon_{jk} \quad \text{for all } k \in \boldsymbol{I}_j^{uu}. \tag{11}$$

(8) The species m_j is added to the set of good represented species, $\boldsymbol{\Phi}_{j-1}^{g}$, to yield

$$\boldsymbol{\Phi}_j^g = \boldsymbol{\Phi}_{j-1}^g \cup \{\text{species } m_j\}. \tag{12}$$

(9) The value of j is incremented, and the next stage is started at (5).

At the end of the algorithm, the ordered set $\boldsymbol{\Phi}^g$ presents a good choice of represented species for dimension reduction methods. For implementing dimension reduction at any given value of n_{rs}, the first n_{rs} number of species from the ordered set $\boldsymbol{\Phi}^g$ are used as the represented species.

Note that for any given reduced dimension n_{rs}, the above algorithm does not give the optimal set of species, $\boldsymbol{\Phi}^{opt}$, which minimizes the global error at that value of n_{rs}, but incrementally adds the best *available* species at each stage to the set of represented species computed from the previous stage.

4. Results

4.1. *Greedy algorithm results*

The greedy algorithm presented in the previous section is applied on the set of species (5) to obtain the species ordering based on the species reconstruction and reaction mapping errors. A total of $N = 2500$ test compositions are used in the full space to compute errors. The justification for choosing this value of N and the sensitivity of results to changes in N are discussed later in Section 4.3.

The first three stages of the algorithm for species selection based on the species reconstruction error are illustrated in Figure 5. At each stage j (for $1 \leq j \leq 3$), the error ϵ_{jk} (10) resulting from the addition of species k (for each $k \in \boldsymbol{I}_j^{uu}$) to the set of represented species from the previous stage, $\boldsymbol{\Phi}_{j-1}^{g}$, is plotted. The species producing the minimum error is selected at each stage. As N_2 is the only species in (5) containing nitrogen, it is a determined species and hence is not considered for selection. At each stage the determined and already-selected species, which are not part of the unrepresented undetermined species set, are marked with a dotted line. The x-axis labels show the entire species ordering obtained at the end of the algorithm. At stage 1, we see that the species CH_4 produces 25% less species reconstruction error than other species, when used as the represented species for dimension reduction, and hence CH_4 is selected by the greedy algorithm as the first 'best' represented species, $\boldsymbol{\Phi}_1^g = \{CH_4\}$. At stage 2, we pick the second species which when used along with the previously selected good species, CH_4, produces the minimum species reconstruction error, and as we see the species O_2 with CH_4 produces the minimum error, and so O_2 is the second species selected by the greedy algorithm, $\boldsymbol{\Phi}_2^g = \{CH_4, O_2\}$. At stage 3, we pick the third species which when used with the two previously selected species, i.e. $\{CH_4, O_2\}$, produces the minimum species reconstruction error, which is found to be the species C_2H_4, and hence is selected in the third stage, $\boldsymbol{\Phi}_3^g = \{CH_4, O_2, C_2H_4\}$. The algorithm continues in this fashion until the required number of represented species are selected.

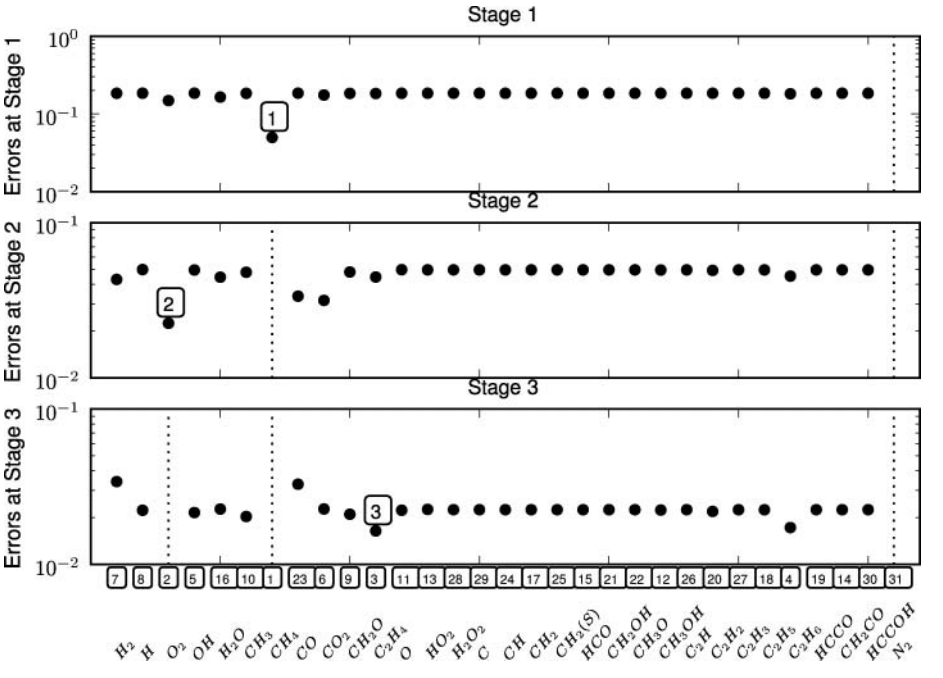

Figure 5. Illustration of the first three stages of species selection based on the species reconstruction error using the greedy algorithm. At each stage the species producing the minimum error is selected. At each stage, the determined and already selected species are not included in the selection, and are marked by a dashed line. The numbering on the x-axis shows the final species ordering obtained at the end of the algorithm.

The full species ordering based on the species reconstruction error is shown in Figure 6. The figure illustrates the stage by stage selection of the best available species using the greedy algorithm. At each stage j, the ordering of the unrepresented undetermined species, $\mathbf{\Phi}_j^{uu}$, is shown, based on increasing species reconstruction error ϵ_{jk} (10) from bottom to top. The x-axis labels list the best species selected at the end of each stage, which corresponds to the bottom most species (which minimizes the error) appearing in the list at that stage. One important observation we make is that at each stage the unrepresented undetermined species are reordered significantly from the previous stage especially in the initial stages. For example, the species selected in stages 5–8 appear at the top of the ordering (produce high error) in the initial stages. The primary reason for this reordering is that the error at every stage depends on all the species selected in the previous stages, and as a consequence the species selected at each stage controls the error in the subsequent stages. As a result, the species selected in the first few stages of the algorithm have a significant effect on the rest of the species ordering.

Figure 7 shows the corresponding species reconstruction error values, ϵ_{jk} (10), at each stage j. The error values, ϵ_{jk}, are marked with a dot, and for clarity, only the species which produce the minimum and maximum error are numbered. We see that the range of errors at each stage is very narrow, except at stages 8, 9 and 13, and the minimum error decreases monotonically with every stage. At every stage, we also observe that certain species, if selected, result in an increase in the species reconstruction error. This behavior is analyzed further in the next section.

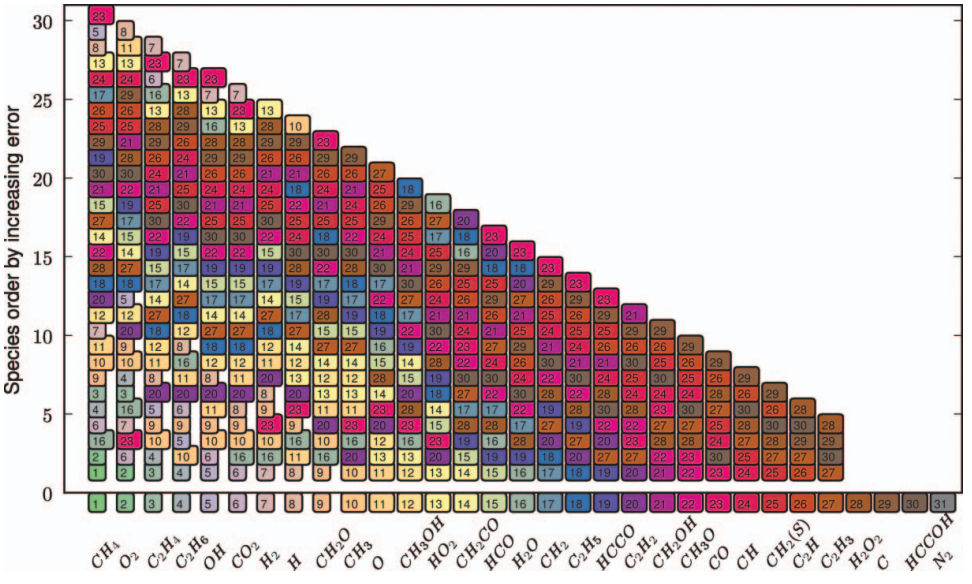

Figure 6. Illustration of the working of the greedy algorithm for species selection based on the species reconstruction error. The x-axis shows the stages $(1, \ldots, n_s = 31)$ for the 31 species present in the GRI-Mech 1.2 methane mechanism. At each stage, the unrepresented undetermined species are plotted with increasing error (from bottom to top) and the species producing the minimum error is selected, which is marked on the x-axis. The algorithm stops at stage 27, when $n_{rs} = n_s - n_e = 27$ because thereafter the species concentrations are determined using element conservation. The species ordering of the last four species is inconsequential.

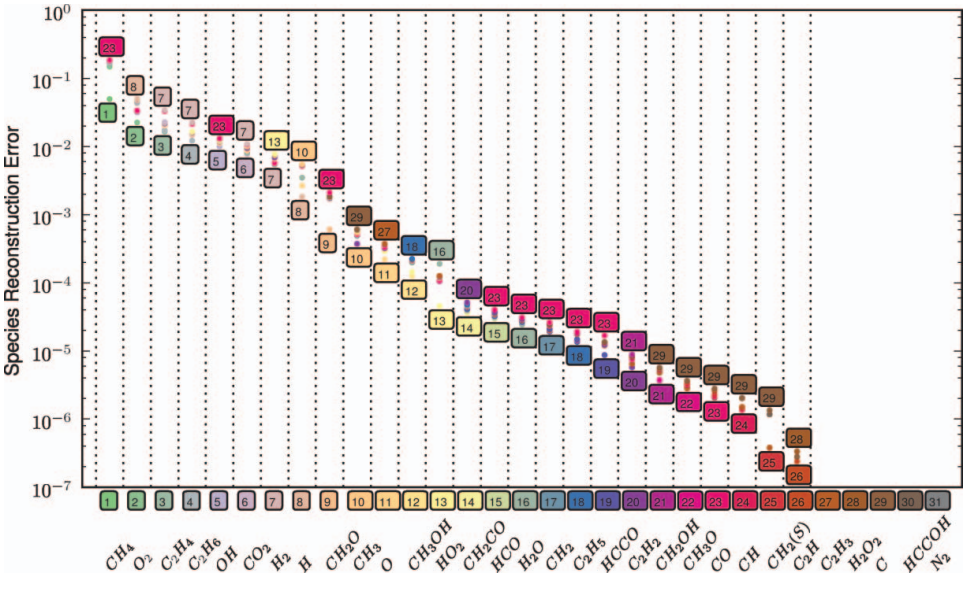

Figure 7. Plot showing the range of species reconstruction errors produced by the unrepresented undetermined species at each stage of the species selection using the greedy algorithm. The x-axis shows the stages $(1, \ldots, n_s = 31)$ for the 31 species present in the GRI-Mech 1.2 methane mechanism. At each stage, the species reconstruction error produced by adding each unrepresented undetermined species is marked with a dot, and the species that produce the minimum and maximum errors are numbered. The species which minimizes the error is selected at each stage.

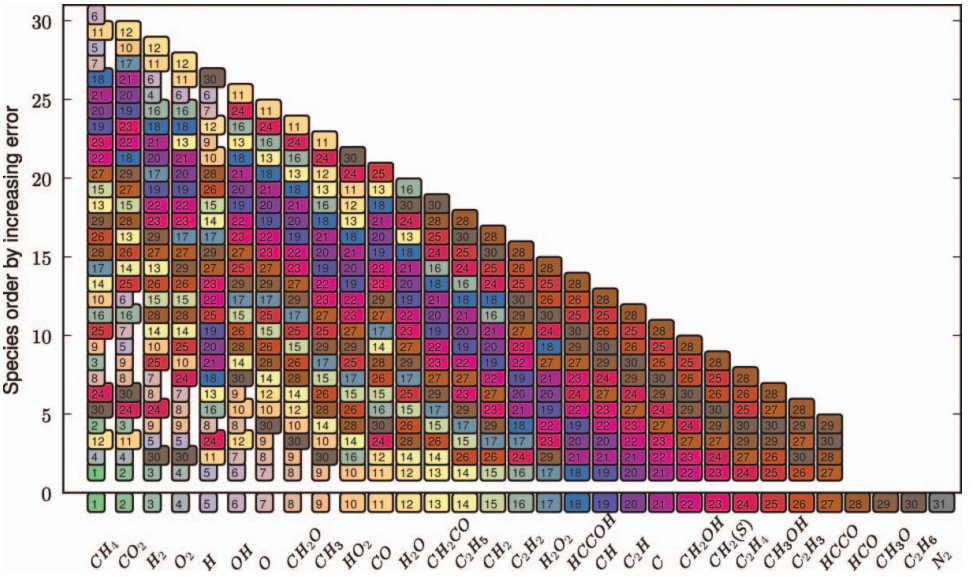

Figure 8. Illustration of the working of the greedy algorithm for species selection based on the reaction mapping error. The x-axis shows the stages $(1, \ldots, n_s = 31)$ for the 31 species present in the GRI-Mech 1.2 methane mechanism. At each stage, the unrepresented undetermined species are plotted with increasing error (from bottom to top) and the species producing the minimum error is selected, which is marked on the x-axis. The algorithm stops at stage 27, when $n_{rs} = n_s - n_e = 27$ because thereafter the species concentrations are determined using element conservation. The species ordering of the last four species is inconsequential.

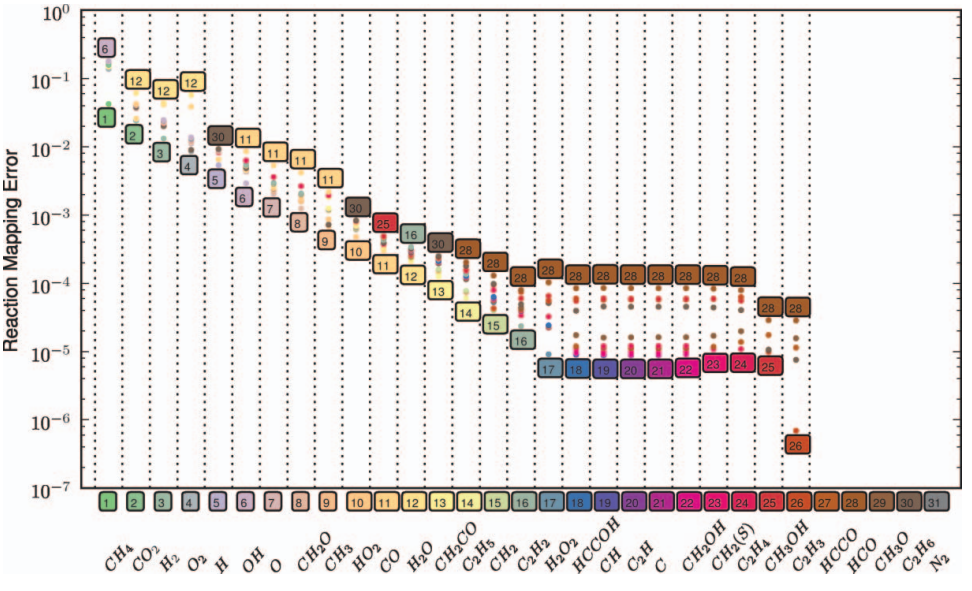

Figure 9. Plot showing the range of reaction mapping errors produced by the unrepresented undetermined species at each stage of the species selection using the greedy algorithm. The x-axis shows the stages $(1, \ldots, n_s = 31)$ for the 31 species present in the GRI-Mech 1.2 methane mechanism. At each stage, the species reconstruction error produced by adding each unrepresented undetermined species is marked with a dot, and the species that produce the minimum and maximum errors are numbered. The species which minimizes the error is selected at each stage.

Figures 8 and 9 show the same plots for species selection based on the reaction mapping error. While we make similar observations in Figure 8 as for the previous case, Figure 9 shows some interesting behavior after stage 16, where we see that the minimum reaction mapping error stays constant and then increases slightly at around stages 23 and 24 and then drops down again at stage 26. The species 28 also consistently shows high errors after stage 16. It appears as if the solution is 'trapped' in a local minimum, where addition of any more species does not result in any further decrease in error. This may be a shortcoming of using the greedy algorithm, which picks the locally optimal species at each stage without reconsidering the previous choices. As a consequence, the greedy algorithm can get trapped in a local minimum and may give a sub-optimal solution.

Nevertheless, from these results we can draw some important conclusions:

(1) The greedy algorithm has been successfully implemented, and is able to pick the best available species at each stage.
(2) At least up to $n_{rs} = 16$ (for the 31-species methane GRI-Mech 1.2 mechanism), the error decreases with each added represented species using the greedy algorithm.
(3) The error decreases by more than 30% in the first two stages, and on average, the error decreases by 8% with each added represented species in the first 16 stages.
(4) In order to achieve 1% and 0.1% levels of error, approximately 5 and 10 represented species are required, respectively.

4.2. *Worst case scenario*

In this section, to emphasize the importance of carefully selecting the represented species for dimension reduction, we perform a series of worst case analyses by computing errors incurred when using a bad set of species compared to using the species obtained from our algorithm.

To pick the 'worst' species, we again use the same greedy algorithm described in Section 3.3 with the exception that at every stage instead of picking the unrepresented species which minimizes the error, we pick the species which maximizes the error. From here on, we refer to this 'worst' species selection algorithm as the *greedy-worst* algorithm.

Figure 10 shows the worst species ordering obtained using the *greedy-worst* algorithm based on the reaction mapping error. It can be clearly seen that the error for this worst ordering of species remains almost constant with increase in the dimension, n_{rs}. Also the errors are orders of magnitude more than the errors obtained with the species ordering using our species selection algorithm.

We see that the major species – CH_4, O_2, CO_2 and H_2O – are the last species selected by the previous method. To check if manually including the major species first improves the errors, we performed another worst case test in which we first manually selected the major species in the system – CH_4, O_2, CO_2 and H_2O – and then picked the rest of the worst species using the *greedy-worst* algorithm. The species ordering for this case based on the reaction mapping error is shown in Figure 11. In this cases also, we see that including the major species in the represented set does not solve the problem fully; the errors still stay very high with the worst species ordering.

We can draw some important conclusions from these worst case scenarios:

(1) Increase in the number of represented species (or, equivalently, the dimension of the reduced space) does not necessarily result in a reduction of dimension reduction error.

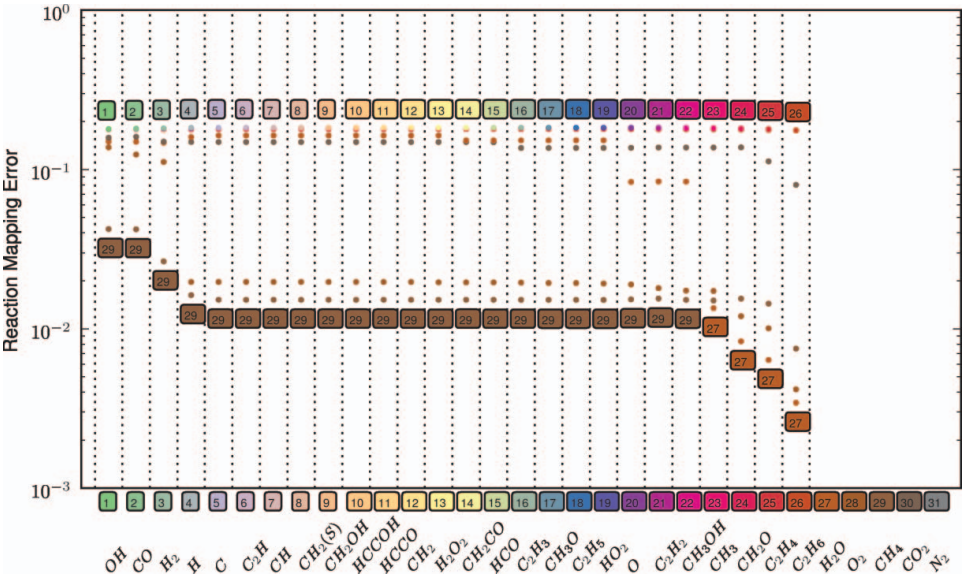

Figure 10. Plot showing the reaction mapping errors produced by selecting the worst species, i.e. species with maximum error at each stage of the greedy algorithm. At each stage, the reaction mapping errors in all the unrepresented undetermined species are marked with a dot, and the species which produce the minimum and maximum errors are numbered. The species which maximizes the error is selected at each stage, which is marked on the x-axis.

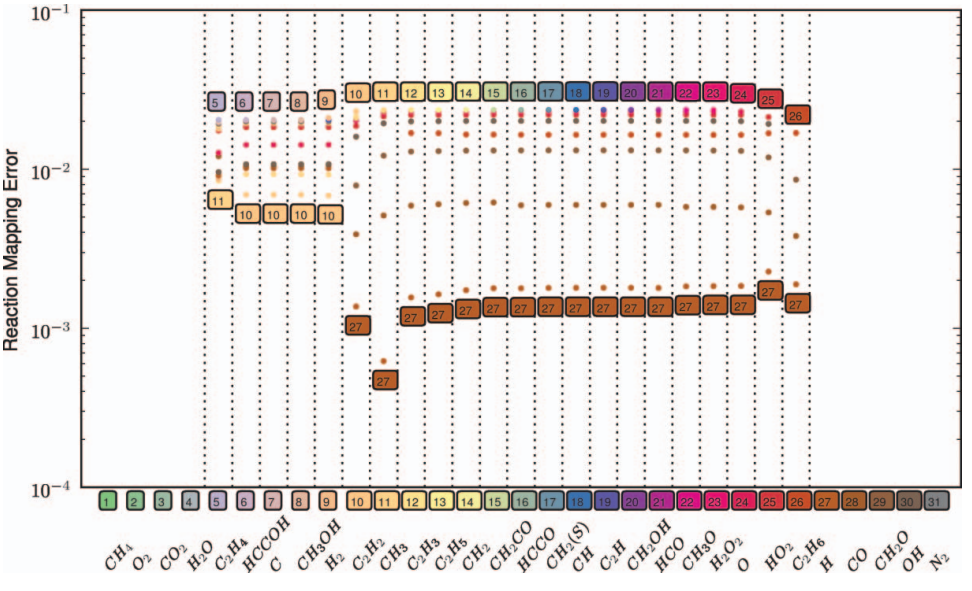

Figure 11. Plot shows the reaction mapping errors produced by selecting the worst species, i.e. species with maximum error at each stage of the greedy algorithm, with the major species: CH_4, O_2, CO_2 and H_2O manually fixed first in the ordering. At each stage, the reaction mapping errors in all the unrepresented undetermined species are marked with a dot, and the species which produce the minimum and maximum errors are numbered. The species which maximizes the error is selected at each stage, which is marked on the x-axis.

(2) A *bad* set of species can result in an error which is orders of magnitude greater than the error resulting from a well chosen set of species.
(3) Including major species does not always help; the errors can still remain very high if the rest of the species are not well selected.

4.3. *Sensitivity tests*

In this section we perform a series of tests to investigate the sensitivity of the species ordering to changes in the number of test compositions, N, used to compute errors; changes in various testing conditions like the pressure, initial temperature and the equivalence ratio; changes in the definition of the error, ϵ; and also to analyze how strongly the dimension reduction results depend on the choice of represented species.

Choice of test compositions

To compute accurate species orderings using the greedy algorithm, the errors involved need to be computed accurately. Since the errors are computed on the N chosen test compositions, the choice of test compositions is crucial.

At each stage of the greedy algorithm, for each candidate species in the unrepresented undetermined species set, the errors are computed at all the chosen N testing compositions to find the species producing the minimum error. Selecting a large number of testing compositions makes the algorithm expensive, and too small a value of N may not give accurate species orderings.

The species selected in the initial stages of the greedy algorithm are crucial as, in all the results presented so far, we see that the species reconstruction and the reaction mapping errors drop rapidly to below 10^{-2} in the first 8–10 species. Hence, N is chosen high enough such that the first 8–10 species in the species ordering remain unchanged with any further increase in N.

Species orderings obtained with increasing N at $\phi = 1$, $T = 600\,K$ and $p = 1$ atm are shown in Figure 12. From Figure 12(a) it is seen that the first 10 species are identical with $N = 2500$ and $N = 3000$; whereas with $N = 1000$ only the first species is the same. From Figure 12(b) we see that the tenth species differs between $N = 2500$ and $N = 3000$; whereas with $N = 1000$ there are three species which differ. At other testing conditions also, the species ordering results (not presented here for brevity) show that the first 8–10 species remain unchanged with $N \geq 2500$ test compositions. Hence all the results in this paper are presented with $N = 2500$ test compositions.

Sensitivity to changes in testing conditions

To investigate the species ordering sensitivity to changes in PaSR testing conditions, the species orderings are obtained at all of the 12 conditions listed in Table 1. The species orderings (first 10 species) obtained for these cases based on the reaction mapping error are listed in Table 2 and based on the species reconstruction error are shown in Table 3.

Table 1. The following set of testing conditions (overall 12 combinations) are considered for the PaSR tests.

Equivalence ratio (ϕ)	1	0.8	1.2
Pressure (atm)	1	10	
Initial temperature (K)	600	1200	

Table 2. Displayed are the first ten species selected by the greedy algorithm based on the reaction mapping error at all the 12 testing conditions listed in Table 1.

	$\phi = 1.0$ $T = 600$ $p = 1$	$\phi = 1.0$ $T = 1200$ $p = 1$	$\phi = 1.0$ $T = 600$ $p = 10$	$\phi = 1.0$ $T = 1200$ $p = 10$	$\phi = 0.8$ $T = 600$ $p = 1$	$\phi = 0.8$ $T = 1200$ $p = 1$	$\phi = 0.8$ $T = 600$ $p = 10$	$\phi = 0.8$ $T = 1200$ $p = 10$	$\phi = 1.2$ $T = 600$ $p = 1$	$\phi = 1.2$ $T = 1200$ $p = 1$	$\phi = 1.2$ $T = 600$ $p = 10$	$\phi = 1.2$ $T = 1200$ $p = 10$
1	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4
2	CO_2	CO_2	CO	CO_2	CO_2	CO_2	C_2H_6	C_2H_6	O_2	CO_2	O_2	CO_2
3	H_2	H_2	C_2H_6	CH_2O	H_2	H_2	OH	CO	CH_3	H_2	C_2H_6	H_2
4	O_2	H	H_2	C_2H_6	O_2	O_2	C_2H_2	CH_2O	C_2H_6	C_2H_6	CO_2	C_2H_6
5	H	C_2H_6	C_2H_4	H	C_2H_4	C_2H_4	CH_2CO	H_2	CO_2	O	CH_2O	CH_2O
6	OH	O	CH_2O	H_2	CH_2O	CH_2O	HCCOH	O	H_2O	CH_2O	H_2	OH
7	O	CH_2O	CH_3	O	CH_3OH	CO	CH	CH_3	H	CH_3OH	CH_3	O_2
8	CH_2O	OH	HO_2	CH_3	HO_2	O	C	CH_3OH	CH_2O	CH_2	C_2H_2	HO_2
9	CH_3	O_2	OH	C_2H_2	C_2H_2	OH	C_2H	C_2H_2	OH	C_2H_3	C_2H_4	CO
10	HO_2	HO_2	H	CH_2CO	CH_2CO	C_2H_2	HCCO	CH_2CO	C_2H_2	CH	C_2H_3	O

Table 3. Displayed are the first ten species selected by the greedy algorithm based on the species reconstruction error at all the 12 testing conditions listed in Table 1.

	$\phi = 1.0$ $T = 600$ $p = 1$	$\phi = 1.0$ $T = 1200$ $p = 1$	$\phi = 1.0$ $T = 600$ $p = 10$	$\phi = 1.0$ $T = 1200$ $p = 10$	$\phi = 0.8$ $T = 600$ $p = 1$	$\phi = 0.8$ $T = 1200$ $p = 1$	$\phi = 0.8$ $T = 600$ $p = 10$	$\phi = 0.8$ $T = 1200$ $p = 10$	$\phi = 1.2$ $T = 600$ $p = 1$	$\phi = 1.2$ $T = 1200$ $p = 1$	$\phi = 1.2$ $T = 600$ $p = 10$	$\phi = 1.2$ $T = 1200$ $p = 10$
1	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4	CH_4
2	O_2	O_2	O_2	O_2	O_2	O_2	O_2	O_2	O_2	O_2	O_2	O_2
3	C_2H_4	C_2H_4	C_2H_4	C_2H_6	C_2H_4	C_2H_4	C_2H_6	C_2H_6	C_2H_4	C_2H_4	H_2O	C_2H_6
4	C_2H_6	OH	C_2H_6	C_2H_4	C_2H_6	OH	C_2H_4	C_2H_4	CO_2	C_2H_6	CO	C_2H_4
5	OH	C_2H_6	CO_2	CH_2O	OH	C_2H_6	CO_2	CH_2O	H_2	CH_3	H_2	CO_2
6	CO_2	CH_3	H_2	H_2O	CO_2	O	H_2	OH	H	CO_2	CO_2	H_2
7	H_2	H	CH_2O	CO	H_2	H	CH_2O	CO_2	CH_2O	H_2O	OH	H
8	H	O	H	CH_3	H	CH_3	H	H_2	OH	H	C_2H_6	CH_2O
9	CH_2O	CO_2	OH	OH	CH_2O	CO_2	OH	O	C_2H_2	H_2	H	OH
10	CH_3	CH_2O	H_2O	H	O	CH_2O	HO_2	HO_2	CH_3	CH_3OH	C_2H_4	HO_2

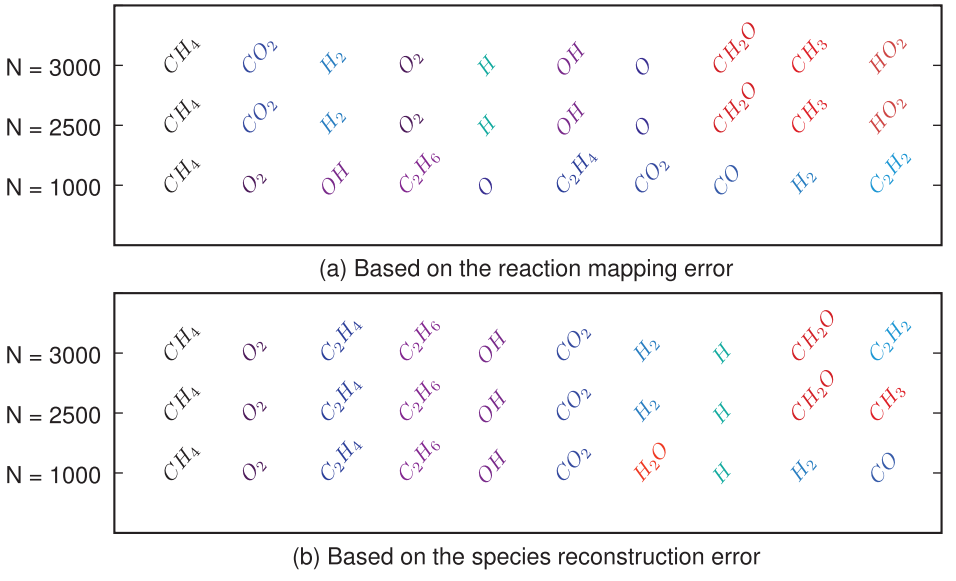

Figure 12. Displayed are the first ten species selected using the greedy algorithm based on (a) the reaction mapping error and (b) the species reconstruction error using the PaSR at $\phi = 1$, $T = 600$ K and $p = 1$ atm with increasing number of test compositions, N, used to compute the errors.

Next, using these species orderings we performed two sets of tests:

(1) Different PaSR tests at fixed species ordering:
We pick a species ordering obtained at a particular testing condition, and then using this species ordering we perform PaSR tests at all the conditions listed in Table 1 with dimension reduction and analyze the reaction mapping and species reconstruction errors at various values of n_{rs}. We are interested in determining whether the error in all these cases decreases monotonically or if it shows some irregular trends.
(2) Fixed PaSR test with different species orderings:
In this we perform a fixed PaSR test with dimension reduction using a selected number of species orderings, i.e, different sets of represented species (obtained at different testing conditions listed in Table 1) and analyze how the reaction mapping and species reconstruction errors vary with different choices of represented species.

Since we are more interested in the reaction mapping error which determines the error in the represented species concentration at the end of the reaction time step, Δt, we perform more tests based on the reaction mapping error than on the species reconstruction error.

PaSR tests at fixed species ordering

We pick the species ordering based on the reaction mapping error obtained using the PaSR test at an equivalence ratio, $\phi = 1$, initial temperature, $T = 600$ K and pressure $p = 1$ atm. Using this species ordering we perform PaSR tests with dimension reduction for $1 \leq n_{rs} \leq 10$ at all the conditions listed in Table 1. For each value of n_{rs}, the first n_{rs} number of species from the chosen species ordering are used as represented species.

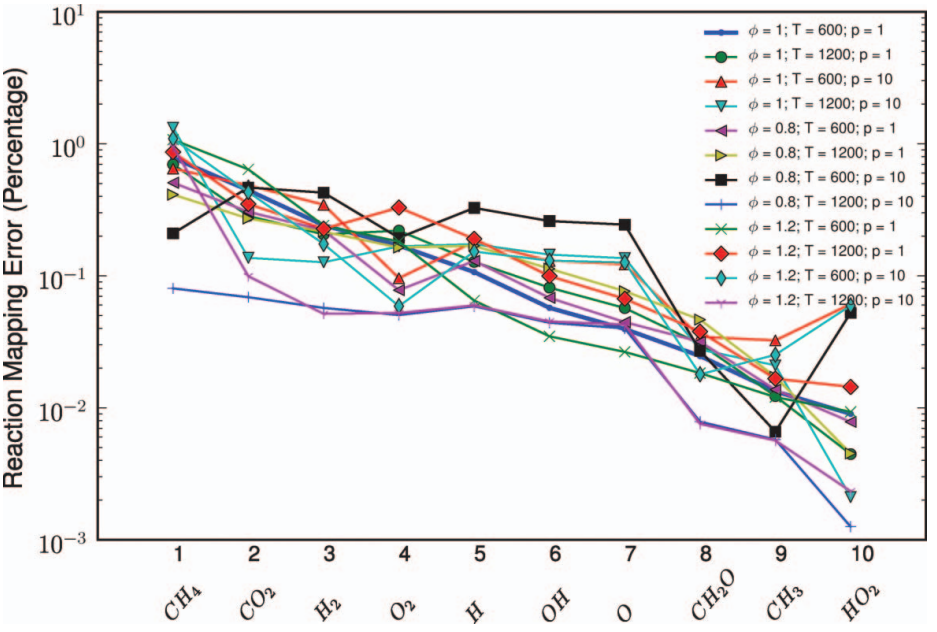

Figure 13. Plot of the reaction mapping error in PaSR tests performed at all the 12 conditions listed in Table 1 with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from the species ordering obtained based on the reaction mapping error at $\phi = 1$, $T = 600$ K and $p = 1$ atm.

The reaction mapping error obtained at each value of n_{rs} is plotted in Figure 13. We make the following observations:

(1) For a given value of n_{rs}, the error typically varies by a factor of 10 depending on the conditions.
(2) While the general behavior is for the error to decrease with increasing value of n_{rs}, there are many instances of the error increasing significantly: for example, in one case ($\phi = 0.8$, $T = 600$ K and $p = 10$ atm), as the value of n_{rs} increases from 9 to 10, the error increases by a factor of 10.

Fixed PaSR test with various species ordering

In the greedy algorithm presented, all the species are treated equally and there are no predefined 'major' or 'minor' species. But, in practice, it is often desirable to include the major species in the calculations. So, at a few selected testing conditions, we obtained species ordering using the greedy algorithm with the major species CH_4, O_2, CO_2 and H_2O fixed first in the ordering. The species orderings obtained (first 10 species) are listed in Table 4.

We performed PaSR tests with dimension reduction at $1 \leq n_{rs} \leq 10$ using various species orderings:

Table 4. Displayed are the species orderings (first 10 species) obtained using the greedy algorithm based on the reaction mapping error at five selected testing conditions (from Table 1) with the four major species: CH_4, O_2, CO_2 and H_2O manually fixed first in the species ordering.

	$\phi = 1.0$ $T = 600$ $p = 1$	$\phi = 1.0$ $T = 1200$ $p = 1$	$\phi = 1.0$ $T = 600$ $p = 10$	$\phi = 0.8$ $T = 1200$ $p = 1$	$\phi = 1.2$ $T = 1200$ $p = 1$
1	CH_4	CH_4	CH_4	CH_4	CH_4
2	O_2	O_2	O_2	O_2	O_2
3	CO_2	CO_2	CO_2	CO_2	CO_2
4	H_2O	H_2O	H_2O	H_2O	H_2O
5	CH_3	CH_3	H_2	CH_3	CH_3
6	H	H	CH_2O	OH	H
7	C_2H_6	OH	HO_2	H	OH
8	OH	H_2	CH_3	CH_2O	H_2
9	CH_2O	O	CH_3OH	C_2H_6	HO_2
10	C_2H_2	CO	H_2O_2	HO_2	CH_3OH

- Figure 14 shows the reaction mapping error obtained using the PaSR test performed at $\phi = 1$, $T = 600$ K and $p = 1$ atm with various species orderings based on the reaction mapping error. We see that the species ordering obtained at the same testing conditions ($\phi = 1$, $T = 600$ K and $p = 1$ atm) produces the lowest error at all values of n_{rs} except at $n_{rs} = 10$. The error with this species ordering decreases by 10% on average with every stage, while at other species orderings the error values are greater and at many instances stay constant with increase in n_{rs}. This case clearly demonstrates that

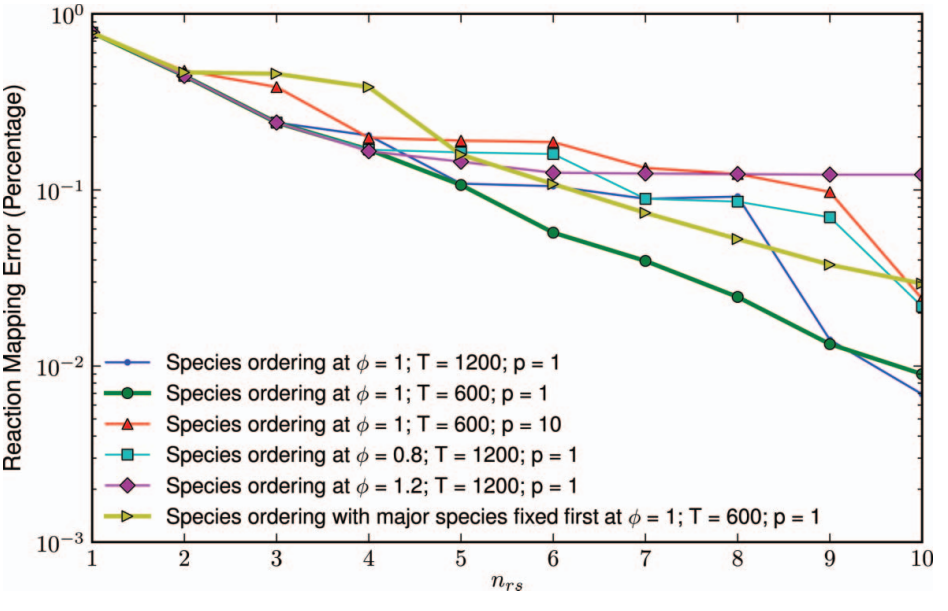

Figure 14. Plot of the reaction mapping error in the PaSR test performed at $\phi = 1$, $T = 600$ K and $p = 1$ atm with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from various species orderings based on the reaction mapping error, given in Tables 2 and 4.

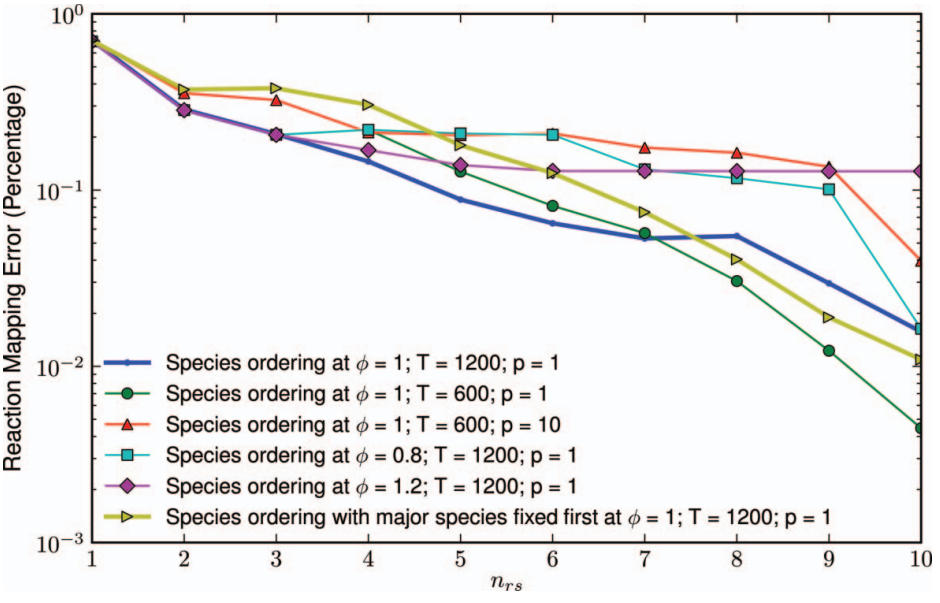

Figure 15. Plot of the reaction mapping error in the PaSR test performed at $\phi = 1$, $T = 1200$ K and $p = 1$ atm with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from various species orderings based on the reaction mapping error, given in Tables 2 and 4.

the greedy algorithm is successfully able to select 'good' represented species which produce low dimension reduction error.

- Figure 15 shows results for the PaSR test performed at a slightly higher temperature of $T = 1200$ K, $\phi = 1$ and $p = 1$ atm. Here also we see that the species ordering obtained at the same testing conditions ($T = 1200$ K, $\phi = 1$ and $p = 1$ atm) produces the lowest error for $n_{rs} \leq 7$; thereafter the error values are still low but slightly higher than species ordering obtained with major species fixed first and species ordering obtained at $\phi = 0.8$, $T = 1200$ K and $p = 1$ atm. This again shows that the greedy algorithm successfully captured the best species in the initial stages, but since the greedy algorithm does not reconsider its choices, it fails to capture the best possible set of represented species at higher values of n_{rs}.
- Figure 16 shows results for the PaSR test performed at higher pressure of $p = 10$, $\phi = 1$ and $T = 600$ K. Here we see that for $n_{rs} \leq 3$ the species ordering obtained at the same testing conditions ($p = 10$, $\phi = 1$ and $T = 600$ K) produces the lowest error, but is soon out-performed by some of the other species orderings. In the range $5 \leq n_{rs} \leq 9$ the species ordering with major species fixed first produces the lowest error. For this case, even though the greedy algorithm does not give the best possible set of species at many values of n_{rs}, the error with the species ordering obtained at the same testing conditions decreases monotonically and remains very close to the lowest error values achieved at all the values of n_{rs}. The other species orderings which produce lower errors show a highly irregular trend, with the error increasing/decreasing by more than 10% at many values of n_{rs}.
- Figure 17 shows the results for lean premixed combustion at $\phi = 0.8$, $T = 1200$ K and $p = 1$ atm. In this case also we see behavior similar to the previous case: the

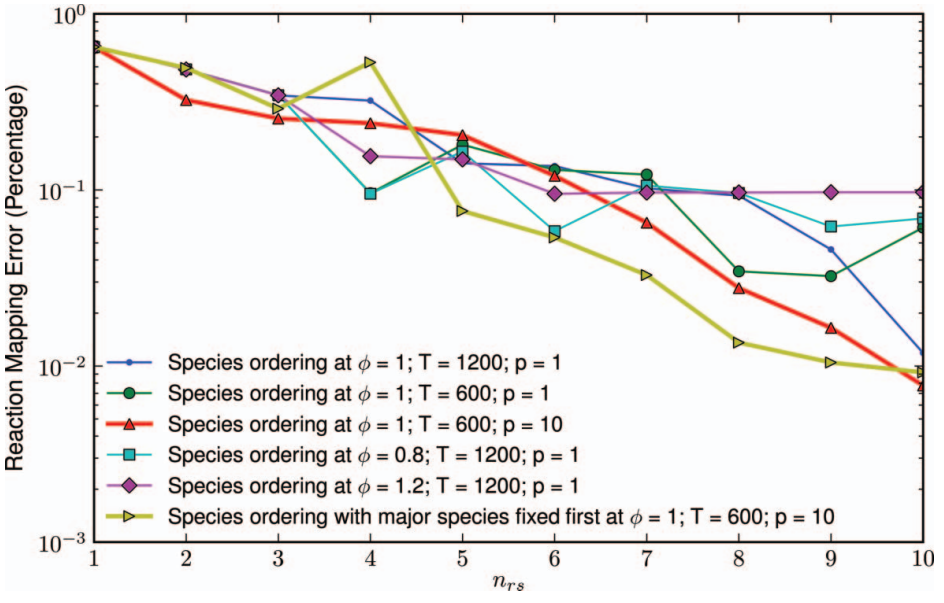

Figure 16. Plot of the reaction mapping error in the PaSR test performed at $\phi = 1$, $T = 600$ K and $p = 10$ atm with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from various species orderings based on the reaction mapping error, given in Tables 2 and 4.

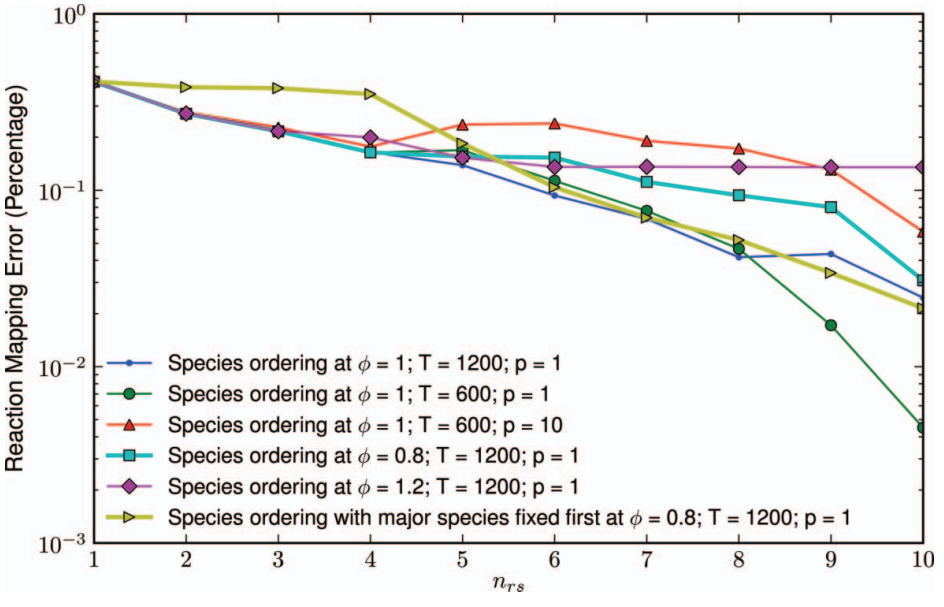

Figure 17. Plot of the reaction mapping error in the PaSR test performed at $\phi = 0.8$, $T = 1200$ K and $p = 1$ atm with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from various species orderings based on the reaction mapping error, given in Tables 2 and 4.

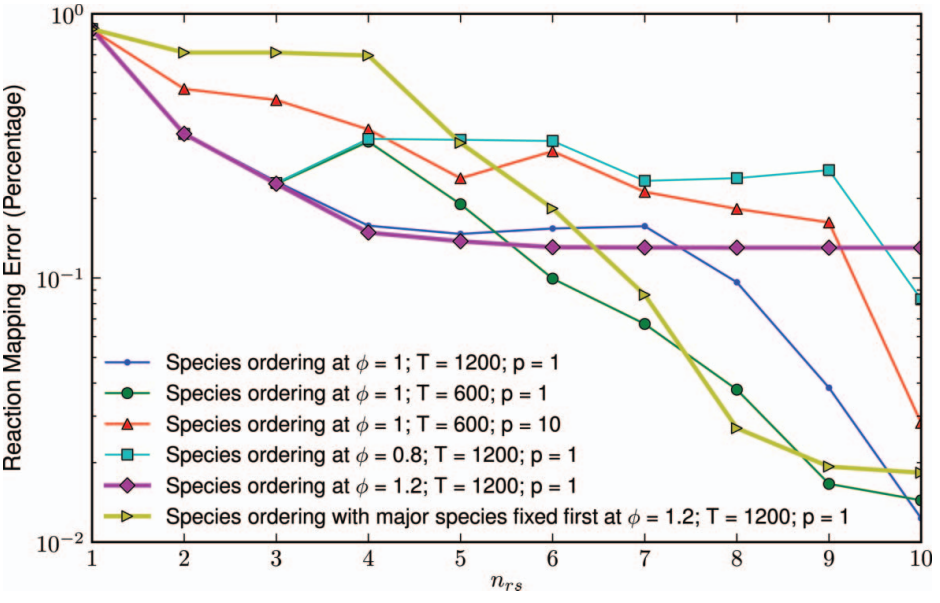

Figure 18. Plot of the reaction mapping error in the PaSR test performed at $\phi = 1.2$, $T = 1200$ K and $p = 1$ atm with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from various species orderings based on the reaction mapping error, given in Tables 2 and 4.

species ordering obtained at the same conditions ($\phi = 0.8$, $T = 1200$ K and $p = 1$ atm) produces the lowest error for $n_{rs} \leq 3$; it remains close to the lowest error achieved at higher values of n_{rs}, and decreases monotonically.

- Figure 18 shows the results for a rich premixed combustion at $\phi = 1.2$, $T = 1200$ K and $p = 1$ atm. In this case, interestingly we see that the species ordering obtained using the greedy algorithm at the same conditions ($\phi = 1.2$, $T = 1200$ K and $p = 1$ atm) produces the lowest error for $n_{rs} \leq 5$ but thereafter the error decreases at a very slow rate. This again could be a case where the greedy algorithm got 'trapped' in a local minimum and as a consequence the error value almost stays constant with any further addition of represented species. The other species orderings including the species ordering with the major species fixed, start with very high errors at low values of n_{rs}, but show some improvement at higher values of n_{rs}, with the error values decreasing but not very significantly.

Apart from the individual observations made in each of the cases, the results for all five cases also show the following:

(1) The species ordering obtained (without any constrained major species) at the same testing conditions at which the PaSR test is performed, always shows the lowest error at low values of n_{rs}, suggesting that the greedy algorithm successfully picks the best species in the initial stages (as it certainly does on the first stage).
(2) The species ordering obtained at the same testing conditions at which the PaSR test is performed is found to produce a non-increasing error with each addition of represented species for $n_{rs} \leq 10$ (even when no such restrictions are enforced by the algorithm itself), which is not always true with other species orderings.

(3) At low values of n_{rs}, the error generated by the species ordering with major species fixed is higher than that generated by other species orderings in all cases.
(4) At high values of n_{rs}, the greedy algorithm does not always give the best possible set of represented species producing the lowest error. But, this is an expected outcome of using a greedy algorithm as it picks the locally optimum species at every stage and is not guaranteed to give the global optimum set of species at all values of n_{rs}.

In short, the results demonstrate that the greedy algorithm (within its own limitations) is able to produce a 'good' set of represented species for a majority of cases tested over a wide range of testing conditions.

Sensitivity to changes in the definition of error

The species ordering obtained using the greedy algorithm depends on the given definition of error, ϵ, i.e. the dimension reduction error that we want to minimize.

The dimension reduction error in the reaction mapping after time t starting from the reconstructed composition, $z^{DR}(0)$, is given by $\epsilon(t, \mathbf{\Phi}^r)$.

We have so far presented results based on two errors:

(1) species reconstruction error, $\epsilon(0, \mathbf{\Phi}^r)$;
(2) reaction mapping error, $\epsilon(\Delta t, \mathbf{\Phi}^r)$.

We have already seen that the species ordering obtained with these two definitions of error are significantly different, as seen in Figures 6 and 8, and also in the various species orderings obtained at different testing conditions in Tables 2 and 3.

Among the two definitions of the error used, the species reconstruction error is cheaper and easier to compute because it does not involve any ODE integrations. So, to examine how much reaction mapping error is incurred if we use represented species selected based on the species reconstruction error instead of the reaction mapping error, we perform a PaSR test at $\phi = 1$, $T = 600$ K and $p = 1$ atm with dimension reduction using represented species from (i) species ordering based on the reaction mapping; and (ii) species ordering based on the species reconstruction error, and then compare the errors for the two cases. The results are shown in Figure 19, and we see that the species ordering based on the species reconstruction error results in higher reaction mapping errors than species ordering based on the reaction mapping error, except at $n_{rs} = 1$, 2 and 8 where they produce approximately the same error. Thus, the species ordering based on the species reconstruction error may not work well for problems involving the computation of reaction mappings, but it still provides a quick and systematic way of obtaining a decent set of represented species to start working with.

From the results presented so far, we see that there is no one good definition of the error which will work all the time to select good species using the greedy algorithm. A good definition of the error is more problem-specific, and depends on the specifications of the problem we are working on and what we want to achieve from using the dimension reduction method. If one is investigating a steady state problem involving only species reconstructions, then an error based at $t = 0$ will work best. In reactive flow problems, depending on the reaction time step one can choose a specific value of time, t, to define the error, or one may also define an error averaged over the time from $t = 0$ to some specific time t. However, note that the greedy algorithm presented in this paper is independent of the definition of the error, and works unchanged with any given definition of the error.

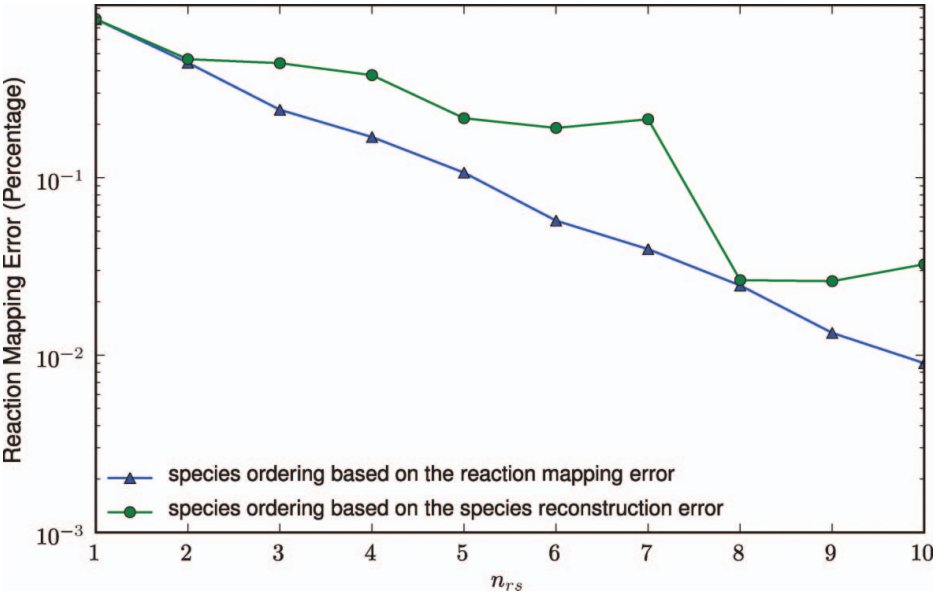

Figure 19. Plot of the reaction mapping error in the PaSR test performed at $\phi = 1$, $T = 600$ K and $p = 1$ atm with dimension reduction (for $1 \leq n_{rs} \leq 10$) using represented species from species orderings obtained at the same conditions ($\phi = 1$, $T = 600$ K and $p = 1$ atm) based on the reaction mapping and species reconstruction errors, given in Tables 2 and 3.

4.4. *Computational cost*

In order to compare the computational cost involved in the dimension reduction steps (species reconstruction and reaction mapping) and to assess the overall cost of the greedy algorithm, the CPU times were measured on a 2.2 GHz Quad-Core AMD Opteron® Processor and are reported in this section.

The average CPU time taken per species reconstruction to compute $\mathbf{z}^{DR}(0) = \mathbf{z}^{CE}(\mathbf{r}(0))$ using CEQ, and the time taken to compute the reaction mapping, $\mathbf{z}^{DR}(\Delta t)$, starting from $\mathbf{z}^{DR}(0)$ using DDASAC are shown in Figure 20. We see that solving the full set of ODEs

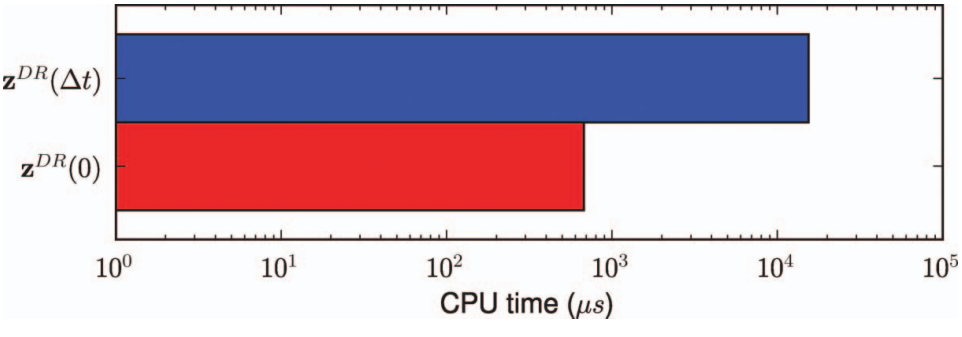

Figure 20. Plot showing the average CPU time taken (in microseconds) per species reconstruction, i.e. to compute $z^{DR}(0)$ given $\boldsymbol{r}(0)$ using CEQ; and to compute the reaction mapping, $z^{DR}(\Delta t)$, starting from $z^{DR}(0)$ using DDASAC. The CPU times are measured on a 2.2 GHz Quad-Core AMD Opteron® Processor.

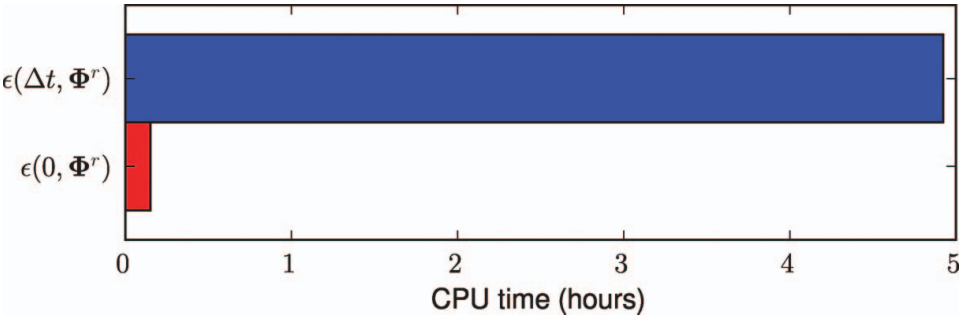

Figure 21. Plot showing the total CPU time taken (in hours) by the greedy algorithm to compute the full species ordering for the GRI-Mech 1.2 mechanism involving 31 species based on the species reconstruction error, $\epsilon(0, \mathbf{\Phi}^r)$, and based on the reaction mapping error, $\epsilon(\Delta t, \mathbf{\Phi}^r)$, using $N = 2500$ test compositions. The CPU time is measured by running a serial implementation of the greedy algorithm on a 2.2 GHz Quad-Core AMD Opteron® Processor.

(1) using DDASAC to compute the reaction mapping takes approximately 15 times more CPU time than that required for performing a species reconstruction using CEQ.

The computation of full species ordering for a chemical mechanism involving n_s species using the greedy algorithm on N chosen test compositions based on:

(1) the species reconstruction error involves $O(N\, n_s^2/2)$ species reconstructions using CEQ;
(2) the reaction mapping error involves computation of $O(N\, n_s^2/2)$ species reconstructions using CEQ and their reaction mappings using DDASAC.

The total CPU time taken to compute the full species ordering based on the species reconstruction error, $\epsilon(0, \Phi^r)$, and based on the reaction mapping error, $\epsilon(\Delta t, \Phi^r)$, is shown in Figure 21. Since the computation of the reaction mapping alone takes 15 times more CPU time than that taken for the species reconstruction, we see that computing the full species ordering based on the reaction mapping error takes approximately 25 times more CPU time than that required for species ordering based on the species reconstruction error.

In Figure 21, the CPU timings are reported for computing the full species ordering. In practice, however, we are mostly interested in only the n_{rs} most important species for a dimension reduction with n_{rs} represented species. So, we can stop once n_{rs} number of species are selected using the algorithm, which involves $O(N\, n_s\, n_{rs})$ number of computations. So, for selecting n_{rs} species, with a fixed number of test compositions, N, the algorithmic cost increases linearly with the number of species, n_s, in the system.

5. Comparison with time scale based methods

As mentioned earlier, the RCCE and ICE-PIC methods are based on the observation that chemical systems involve a wide range of time scales, and as a result the reaction trajectories are attracted to a hierarchy of low dimensional slow invariant manifolds (SIMs). The key issue in the RCCE and ICE-PIC methods is to identify the slow time scales or the slowly evolving species (known as represented species in our work) to best approximate the SIMs.

The greedy algorithm described in this paper provides one way of selecting the representing species. One could also select represented species based on a time scale analysis. Though we are unaware of any implementation of RCCE in which species or constraints are

selected based on a time scale analysis, there exist many methods for generating reduced reaction mechanisms based on QSS [33], CSP [34, 35] and ILDM [36], wherein a time scale analysis is used to identify the unimportant species and reactions in the detailed mechanism which are eliminated to obtain reduced mechanisms.

It may be insightful to compare the species ordering obtained by the greedy algorithm with the order in which species are removed in these reduced mechanism generating methods. In [34] an automatic reduction of chemistry with the CSP (ARC-CSP) method is developed to generate reduced mechanisms. In this method, CSP is used to analyze perfectly-stirred reactor (PSR) data (obtained using a detailed mechanism) to identify and eliminate QSS species from the detailed mechanism to obtain reduced mechanisms. The ARC-CSP method when applied to develop a 10-step reduced mechanism from the GRI Mech 1.2 mechanism (using PSR data from a wide range of operating conditions) identifies and eliminates 17 QSS species in the following order (see Figure 5 in [34]): $CH_2(S)$, C_2H, HCO, CH_3O, CH, C_2H_5, CH_2OH, CH_2, C_2H_3, H_2O_2, HO_2, C, C_2H_6, CH_3, CH_2CO, HCCO, CH_3OH. The 10-step reduced mechanism retains the following 13 species (in no particular order): CH_4, O_2, H_2O, CO_2, O, CO, H_2, OH, H, C_2H_2, C_2H_4, CH_2CO, HCCOH (the inert species N_2 is not considered).

It is interesting to note that only three QSS species – CH_3 (ninth), HO_2 (tenth) and CH_2CO (thirteenth) – appear in the first 13 species selected using the greedy algorithm based on the reaction mapping error (which involves reaction kinetics) as seen in Figure 9. Also, in the first 10 species selected by the greedy algorithm based on the reaction mapping error at various testing conditions listed in Table 2, at most four QSS species are seen in each case, and overall eight out of the 17 QSS species appear in Table 2. We note that five out of the eight QSS species appearing in Table 2 are at the edge of the 10-step cut-off limit (Figure 5 in [34]) and, as pointed out in [34], these five species: CH_3OH, HCCO, CH_2CO, CH_3 and C_2H_6 have approximately the same normalized time scales and are hard to distinguish, and hence are treated as a group of QSS species. However, in the greedy algorithm the species are treated individually and are selected based on an error criterion, and hence some of these species could possibly be selected in the early stages of the algorithm if they produce the lowest error.

Finally, we note that the species CH_4 appears tenth in Figure 5 [34], based on the time scale analysis, and any safety factor of $\alpha < 30$ will generate a reduced mechanism without the CH_4 species. However, as we have seen, the greedy algorithm selects CH_4 as the optimum species (producing the minimum error) at stage 1 in all the cases reported here, and so this clearly shows that a time scale analysis based method for selecting represented species may not work at low values of n_{rs}.

In summary, we do see some similarities between the species ordering obtained using the greedy algorithm with the ordering in which species are removed in time scale based reduced mechanism generating methods. However, we also note that a time scale analysis based method alone may not give 'good' represented species for the RCCE dimension reduction method, especially at low values of n_{rs}.

6. Limitations and variants of the greedy algorithm

Based on the results reported here, we see that the current implementation of the greedy algorithm does not always give a near-optimal set of species, especially at high values of n_{rs}. This may be due to one or more of the following limitations:

(1) The greedy algorithm is shortsighted in nature and never reconsiders its choices.

(2) The species in the current implementation are treated individually without considering any connections (imposed by the reaction mechanism) between species and treating (strongly connected) species in groups as is done in many reduced mechanism building methods [3, 5].

To address some of these limitations, implementation of the following variants of the greedy algorithm are being considered for future work:

(1) Selection-rejection strategy:
The inherent flaw in the greedy algorithm is that choices made in the previous stages are never reconsidered. To compensate for this drawback, we can consider a selection–rejection algorithm, in which at every third or fourth stage of the algorithm, we reject the worst species from the current set of selected species and resume the algorithm. The cost of rejection goes linearly with number of selected species, $O(n_{rs})$, and so is fairly inexpensive and may help in improving the results.

(2) Selecting species in groups:
The greedy algorithm selects the best species at every stage; instead a variant of this algorithm would involve selecting species in groups of two or three species. So considering the case in which we want to select two good species at every stage, then at stage 1 we perform tests with all possible $\binom{n_s}{2}$ number of sets of species and pick the set which produces the minimum error; next at stage 2 we test the remaining $\binom{n_s-2}{2}$ sets of species to pick the next two good species; and so on until the required number of species are selected. Obviously this is an expensive algorithm since the number of tests performed scales as $O(N\, n_s^3)$, but by using a fewer number of test points, N, the computational cost can be reduced. This approach is expected to produce a more nearly optimal set of species.

7. Conclusions

An automated algorithm for selecting a *good* set of species for the accurate implementation of dimension reduction methods has been presented. The following conclusions can be drawn from the results presented in this paper:

(1) The dimension reduction errors in RCCE are highly sensitive to the choice of constrained species; errors can differ by orders of magnitude at the same dimension with two different sets of represented species as seen in Figures 7 and 11.
(2) Constraints based on the major species concentrations are not always the best constraints for the RCCE method as is evident from the results presented in Figures 14–18
(3) At low values of $n_{rs} \leq 5$, the greedy algorithm is successfully able to select a near-optimal set of represented species.
(4) At high values of n_{rs}, the greedy algorithm is not guaranteed to produce near-optimal sets of represented species (see Figure 18) but nevertheless is found to produce a good set of species for the majority of cases tested (as demonstrated in Figures 14–17) over a wide range of testing conditions.
(5) In all the cases tested, the species ordering generated by the greedy algorithm for $n_{rs} \leq 10$ produces a *non-increasing* error with every addition of represented species.

Acknowledgements

This work is supported by Department of Energy under Grant DE-FG02-90ER and the National Science Foundation under Grant CBET-0426787. SBP has a financial interest in Ithaca Combustion Enterprise, LLC, which has licensed the software ISAT-CK and CEQ used in this work.

References

[1] W.J. Pitz, N.P. Cernansky, F.L. Dryer, F.N. Egolfopoulos, J.T. Farrell, D.G. Friend, and H. Pitsch, *Development of an experimental database and kinetic models for surrogate gasoline fuels*, Society of Automotive Engineers SAE-2007-01-0175 (2007).

[2] C.K. Westbrook, W.J. Pitz, O. Herbinet, H.J. Curran, and E.J. Silke, *A comprehensive detailed chemical kinetic reaction mechanism for combustion of n-alkane hydrocarbons from n-octane to n-hexadecane*, Combust. Flame 156 (2009), pp. 181–199.

[3] T. Lu and C.K. Law, *A directed relation graph method for mechanism reduction*, Proc. Combust. Inst. 30 (2005), pp. 1333–1341.

[4] P. Pepiot-Desjardins and H. Pitsch, *An efficient error-propagation next term-based reduction method for large chemical kinetic mechanisms*, Combust. Flame 154 (2008), pp. 67–81.

[5] T. Nagy and T. Turnyi, *Reduction of very large reaction mechanisms using methods based on simulation error minimization*, Combust. Flame 156 (2009), pp. 417–428.

[6] M. Bodenstein and S.C. Lind, *Geschwingigkeitder bildung des bromwasserstoffs aus seinen elementen*, Z. Phys. Chem. 57 (1906), p. 168.

[7] M.D. Smooke, *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, Springer-Verlag, Berlin, 1991.

[8] J.C. Keck and D. Gillespie, *Rate-controlled partial equilibrium method for treating reacting gas-mixtures*, Combust. Flame 17 (1971), pp. 237–241.

[9] J.C. Keck, *Rate-controlled constrained equilibrium theory of chemical reactions in complex systems*, Prog. Energy Combust. Sci. 16 (1990), pp. 125–154.

[10] S.H. Lam and D.A. Goussis, *The CSP method for simplifying kinetics*, Int. J. Chem. Kinetics 26 (1994), pp. 461–486.

[11] U. Maas and S.B. Pope, *Simplifying chemical-kinetics: intrinsic low-dimensional manifolds in composition space*, Combust. Flame 88 (1992), pp. 239–264.

[12] S.B. Pope and U. Mass, *Simplifying chemical kinetics: trajectory-generated low-dimensional manifolds*, Report FDA 93-11, Cornell University, Ithaca, 1993.

[13] Z. Ren, S.B. Pope, A. Vladimirsky, and J.M. Guckenheimer, *The invariant constrained equilibrium edge preimage curve method for the dimension reduction of chemical kinetics*, J. Chem. Phys. 124 (2006), p. Art. No. 114111.

[14] A.N. Al-Khateeb, J.M. Powers, S. Paolucci, A.J. Sommese, J.A. Diller, J.D. Hauenstein, and J.D. Mengers, *One-dimensional slow invariant manifolds for spatially homogenous reactive systems*, J. Chem. Phys. 131 (2009), pp. 024118 (1–19).

[15] J.Y. Chen, W. Kollmann, and R.W. Dibble, *PDF modeling of turbulent methane–air nonpremixed jet flames*, Combust. Sci. Technol. 64 (1989), pp. 315–346.

[16] T. Turanyi, *Parameterization of reaction mechanisms using orthonormal polynomials*, Comput. Chem. 18 (1994), pp. 45–54.

[17] F.C. Christo, A.R. Masri, E.M. Nebot, and S.B. Pope, *An integrated PDF/neural network approach for simulating turbulent reacting systems*, Proc. Combust. Inst. 26 (1996), pp. 43–48.

[18] S.B. Pope, *Computationally efficient implementation of combustion chemistry using in situ adaptive tabulation*, Combust. Theory Model. 1 (1997), pp. 41–63.

[19] L. Lu and S.B. Pope, *An improved algorithm for in situ adaptive tabulation*, J. Comput. Phys. 228(2) (2009), pp. 361–386.

[20] S.R. Tonse, N.W. Moriarity, N.J. Brown, and M. Frenklach, *PRISM: Piecewise Reusable Implementation of Solution Mapping. An economical strategy for chemical kinetics*, Israel J. Chem. 39 (1999), pp. 97–106.

[21] Q. Tang, J. Xu, and S.B. Pope, *PDF calculations of local extinction and NO production in piloted-jet turbulent methane/air flames*, Proc. Combust. Inst. 28 (2000), pp. 133–139.

[22] Q. Tang and S.B. Pope, *Implementation of combustion chemistry by in situ adaptive tabulation of rate-controlled constrained equilibrium manifolds*, Proc. Combust. Inst. 29 (2002), pp. 1411–1417.
[23] Z. Ren, V. Hiremath, and S.B. Pope, *Dimension reduction and tabulation of combustion chemistry using ICE-PIC and ISAT*, 6th US National Combustion Meeting, May 17–20, 2009 Ann Arbor, Michigan, 2009.
[24] S.B. Pope, *PDF methods for turbulent reactive flows*, Prog. Energy Combust. Sci. 11 (1985), pp. 119–192.
[25] S. Rigopoulos, *The Rate-Controlled Constrained Equilibrium (RCCE) method for reducing chemical kinetics in systems with time-scale separation*, Int. J. Multiscale Comput. Eng. 5 (2007), pp. 11–19.
[26] M. Janbozorgi and H. Metghalchi, *Rate-controlled constrained-equilibrium theory applied to the expansion of combustion products in the power stroke of an internal combustion engine*, Int. J. Thermodyn. 12 (2009), pp. 44–50.
[27] M. Janbozorgi, S. Ugarte, H. Metghalchi, and J.C. Keck, *Combustion modeling of mono-carbon fuels using the rate-controlled constrained-equilibrium method*, Combust. Flame 156 (2009), pp. 1871–1885.
[28] D. Hamiroune, P. Bishnu, M. Metghalchi, and J.C. Keck, *Rate-controlled constrained-equilibrium method using constraint potentials*, Combust. Theory Model. 2 (1998), pp. 81–94.
[29] Z. Ren, V. Hiremath, and S.B. Pope, *ICE-PIC: "A library of Fortran 90 routines for using the dimension reduction methods RCCE and ICE-PIC with ISAT*", 2009, http://eccentric.mae.cornell.edu/∼tcg/ice-pic.
[30] M. Caracotsios and W.E. Stewart, *Sensitivity analysis of initial value problems with mixed ODEs and algebraic equations*, Comput. Chem. Eng. 9 (1985), pp. 359–365.
[31] S.B. Pope, *The computation of constrained and unconstrained equilibrium compositions of ideal gas mixtures using Gibbs function continuation*, Cornell University report + FDA 03-02, 2003.
[32] T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press and McGraw-Hill, 2001.
[33] A.C. Zambon and H.K. Chelliah, *Explicit reduced reaction models for ignition, flame propagation, and extinction of $C_2H_4/CH_4/H_2$ and air systems*, Combust. Flame 150 (2007), pp. 71–91.
[34] T. Lu, Y. Ju, and C.K. Law, *Complex CSP for chemistry reduction and analysis*, Combust. Flame 126 (2001), pp. 1445–1455.
[35] M. Valorani, F. Creta, D.A. Goussis, J.C. Lee, and H.N. Najm, *An automatic procedure for the simplification of chemical kinetic mechanisms based on CSP*, Combust. Flame 146 (2006), pp. 29–51.
[36] V. Bykov and U. Maas, *Extension of the ILDM method to the domain of slow chemistry*, Proc. Combust. Inst. 31 (2006), pp. 465–472.