Using genetic algorithm for automated optimization of reduced chemical schemes

N. Jaouen[*1,2], P. Domingo[1], L. Vervisch[1]

[1]CORIA CNRS, Normandie Université, INSA de Rouen, Saint-Etienne-du-Rouvray, France
[2] Safran, Snecma Site de Villaroche, Rond-Point René Ravaud-Réau, 77550 Moissy Cramayel, France

Abstract A strategy to optimize reduced chemical schemes is discussed. A set of one-dimensional premixed laminar flames is first simulated with detailed chemistry and given transport properties. A subset of species preserving the enthalpy budget is defined. Reference chemical source profiles are constructed, so that they properly close mass and energy budgets for this limited number of species. Then, a two-stage optimization procedure is applied: (1) With the reference chemical sources as targets, a preliminary guess of the Arrhenius parameters is obtained with a Newton method, assuming the reduced number of reactions to act separately. From this first approximation, a range of potential values is defined for the coefficients of every reaction. (2) A binary-type genetic algorithm is subsequently coupled with flame simulations to find the Arrhenius parameters that minimize the error over temperature, species and velocity profiles, for all the points of the reference laminar flames. The genes are defined as the Arrhenius coefficients to be optimized. Standard crossover and mutation rates are combined with a ranking selection of the chromosomes. After convergence, species profiles obtained with the reduced scheme are tested against those of the detailed one.

Introduction

Combustion detailed chemistry is essential to understand ignition, flame propagation and pollutants emission. However, because of the large number of degrees of freedom they introduce, advanced chemical schemes are still today impractical to simulate real combustion systems. Over the last decades various methods have been developed to downsize chemistry prior to its introduction in flow solvers. Tabulation methods are very attractive when the fresh and burnt gases conditions are fixed and well known. On the other hand, reduced chemical schemes may be more efficient to handle dilution by burnt gases, heat loss and complex flame structure, as unsteady vitiated partially premixed flames, which can hardly be tabulated with accuracy prior to the three-dimensional simulations.

Reduced chemical schemes based on a few global steps and involving major species, and sometimes also a few radical species, have been widely discussed in the literature. Some of these schemes are derived from skeletal mechanisms after formulating hypotheses of partial equilibrium and steady state for given species [1, 2]. Other mechanisms are obtained from more global physical considerations [3]. All rely on three ingredients: the choice of the species, their combination by the global steps and the parameters of the Arrhenius rates.

While a significant decrease of the CPU time and memory overload are achieved by chemistry reduction, the mechanisms necessarily lose accuracy when compared to the detailed reference ones. This departure is known to dramatically increase when varying the flow conditions (equivalence ratio, pressure and/or temperature) away from those at which the re-

duced scheme was developed.

To overcome some of these shortcomings, automatically generated reduced schemes for large ranges of operating conditions have been discussed in the literature using various optimization based approaches, either focussing on the reduction of the scheme itself [4], or reproducing the species profiles of detailed mechanism [5, 6], or again directly using the chemical sources distribution as targets for optimization [7]. A combination of the two formers is explored further in this paper with genetic algorithm coupled to flame simulations [8], to provide an automated approach that generates reduced chemical schemes parameters optimized for a given set of fresh gases conditions.

Target profiles and chemical sources

Freely propagating one-dimensional premixed flames are simulated at various equivalence ratios with detailed chemistry, to provide reference species mass fractions profiles, $Y_k^\star(x)$. The GRI mechanism [9] is used for CH_4/Air combustion along with the Hirschfelder and Curtiss approximation [10] to estimate the transport coefficients.

Formally, N species are involved in the detailed mechanism and $n < N$ in the reduced one, which represent 98% of the total mass. Target species source terms $\dot{\omega}_k^o$ are reconstructed for the reduced scheme from $Y_k^\star(x)$ for the n species, to ensure that mass budgets are properly closed for the reduced number of transported species:

$$\dot{\omega}_k^o = \nabla \cdot \left(\rho \left[\mathbf{u} + \mathbf{V}_k^\star - \sum_{\ell=1}^{n}(\mathbf{V}_\ell^\star Y_\ell^\star) \right] Y_k^\star \right), \quad (1)$$

with

$$\rho = \frac{P_o}{R\sum_{k=1}^{n}(Y_k^\star/W_k)T^\star} , \qquad (2)$$

$$\mathbf{u} = \frac{\rho^\star}{\rho}\mathbf{u}^\star, \qquad (3)$$

where W_k is the molar weight of the k-th species. The diffusion velocity reads $\mathbf{V}_k^\star = -D_k\nabla X_k^\star/X_k^\star$, with D_k the diffusion coefficient and X_k the molar fraction. As for the detailed chemistry simulation, a diffusive velocity correction is used in Eq. (1).

The Arrhenius reaction rate of the i-th reaction reads:

$$\dot{Q}_i = \mathcal{A}_i\, T^{\beta_i}\exp\left(-\frac{E_{ai}}{R\,T}\right)[A]^{n_{1i}}[B]^{n_{2i}} , \qquad (4)$$

with $\mathcal{A}_i$ the pre-exponential factor, E_{ai} the activation energy, the temperature exponent β_i, and, the exponent of concentrations n_{1i} and n_{2i}. $[A]$ and $[B]$ denote the concentration of the species involved in the chemical reaction step.

Given the high values associated to the pre-exponential factor, $a_i = \log\mathcal{A}_i$ is defined to conduct the optimization process.

The number of reactions steps is usually smaller than the number of species, therefore the problem is overdetermined in terms of reference reaction rates ($\dot{Q}_i^o$) to be determined from the chemical rates of the n species ($\dot{\omega}_k^o$), known from Eq. (1). In the following, a number of species equals to the number of reaction steps N_R to be optimized, are selected to compute the reaction rates Q_i^o from:

$$\dot{\omega}_k^o = W_k\sum_{i=1}^{N_R}\left(\mu_{k,i}^b - \mu_{k,i}^f\right)\dot{Q}_i^o , \qquad (5)$$

where $\mu_{k,i}^f$ (resp. $\mu_{k,i}^b$) is the forward (resp. backward) stoichiometric coefficient of the k-th species in the i-th reaction rate. This implies that we formulate at this stage the hypothesis that the reaction rates obtained while reproducing $N_R < n$ selected species, are not so different from the rates that would be obtained while trying to fairly distribute the error over all the species. An hypothesis relaxed in a second stage where optimization concerns all species, temperature and velocity profiles.

At the end of this preliminary step a set of species profiles $Y_k^\star(x)$ and of reaction rates $\dot{Q}_i^o(x)$ are available.

Fitness functions definition

The error between the reference rate Q^o verifying relation (5) and Q is calibrated as:

$$E_1(x_j;c^p) = |\ln(Q(x_j;c^p)) - \ln(Q^o(x_j))| . \qquad (6)$$

x_j denotes a point of the target profiles and c^p, a given set of the chemical parameters. (The i-th indice
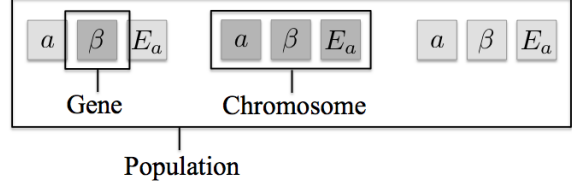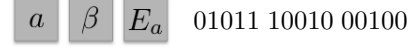

Figure 1: Structure of the population.

Figure 2: Chromosome

in Q_i, denoting the reaction number, is ommitted in this section for sake of clarity.)

A first fitness function is defined cumulating the error over the N_G points of the profiles:

$$
\begin{aligned}
f_1(c^p) &= \sum_{j=1}^{N_G} E_1(x_j;c^p) , \qquad (7)\\
&= \sum_{j=1}^{N_G}(-1)^{\sigma(x_j;c^p)}\Bigg[a(c^p)\ln(10)\\
&+ \beta(c^p)\ln(T^\star(x_j))\\
&- \frac{E_a(c^p)}{R\,T^\star(x_j)} + n_1(c^p)\ln([A]^\star(x_j))\\
&+ n_2(c^p)\ln([B]^\star(x_j)) - \ln(Q^0(x_j))\Bigg] .
\end{aligned}
$$

Where the temperature is computed from the local enthalpy and species concentration. $\sigma(x_i;c^p) = 0$ if $Q(x_i;c^p) > Q(x_i^o)$ and $\sigma(x_i;c^p) = 1$ otherwise.

Similarly, a second fitness function is defined that combines errors measured on the temperature, species and velocity profiles across the premixed flame:

$$
\begin{aligned}
f_2(c^p) &= -\sum_{j=1}^{N_G}\Bigg[\left|\frac{T^\star(x_j) - T(x_j;c^p)}{T^\star(x_j)}\right| \qquad (8)\\
&+ \sum_{k=1}^{n}\left|\frac{Y_k^\star(x_j) - Y_k(x_j;c^p)}{Y_k^\star(x_j)}\right|\\
&+ \left|\frac{u^\star(x_j) - u(x_j;c^p)}{u^\star(x_j)}\right|\Bigg] .
\end{aligned}
$$

This second fitness function controls an iterative process, in which the profiles $T(x_j;c^p)$, $Y_k(x_j;c^p)$ and $u(x_j;c^p)$, are obtained with the reduced scheme for the c^p sets of coefficients solving the flame with Cantera [11].

First guess with uncorrelated reactions

A preliminary estimation of all the parameters of the reaction scheme is obtained by considering every reaction step independently. The source term fitness function given by (8) behaves linearly versus a, β,
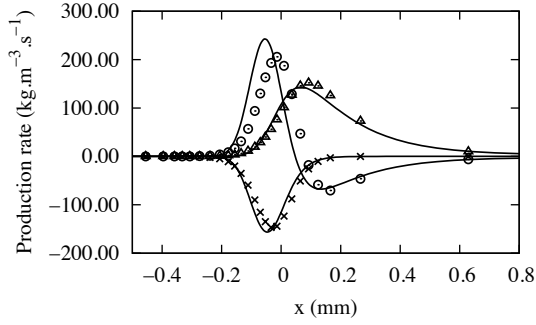
Figure 3: Major species burning rates. Stoichiometric flame. Lines: GRI-mech [9]. Symbols: C1 and C2 of Table 3 after first guess estimation. ×: CH_4. ⊙: CO. △: CO_2.
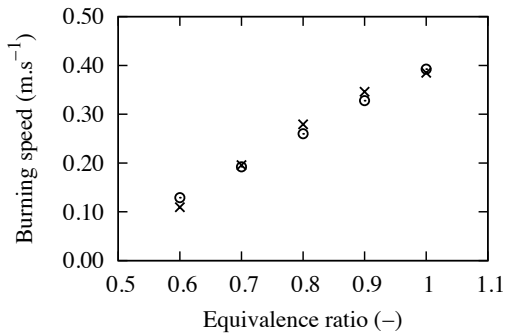

Figure 4: Flame speed response versus equivalence ratio. ×: GRI-mech [9]. ⊙: Present scheme.

E_a, n_1 and n_2, with a constant slope that changes its sign according to $(-1)^\sigma$.

Within this context, a direct approach based on a Newton algorithm is suitable to obtain a very first set of parameters for the reduced chemical rates. These values, a^1, β^1, E_a^1, n_1^1 and n_2^1 are taken as initial conditions for the next step, in which all reactions are now coupled.

Genetic optimization

A genetic algorithm applies the principles of natural selection and of evolution, so as to look for the optimal solution of a problem initialized by a population of potential solutions [8]. Although it is associated to a random search, it uses the information present in the group of potential solutions to converge toward the best compromise. These algorithms are attractive because they do not require the studied functions to be continuous neither monotonous.

The terms 'gene', 'chromosome' and 'population' are introduced for the optimization of the Arrhenius parameters. For simplicity in the notation, only a, β and E_a are considered in the following, see Fig. 1. In practice, the procedure is applied to all parameters of all reactions at once.

In this section, a given triplet of parameters values (a,β,E_a) constitutes a chromosome, while a gene

is the value taken by one of these parameters (Fig. 1). The population is composed of a set of potential solutions, *i.e.* a set of chromosomes.

For every parameter, a range of values is defined allowing a variation of $\pm 2\%$ centered on the initial condition, a^1, β^1 and E_a^1, obtained from the first guess as described above. For a desired precision Δa in the binary description on the parameter a, the number of required binary digits (bits) reads:

$$N_{bits}(a) \geq \frac{ln\left(\dfrac{a_{\max} - a_{\min}}{\Delta a}\right)}{ln(2)}, \qquad (9)$$

$N_{bits}(\beta)$, $N_{bits}(E_a)$ are defined similarly. The gene corresponding to the smallest value $a_{\min}$ (resp. the upper bound $a_{\max}$) is constructed setting all the bits to '0' (resp. '1') (Table 1).

During the evolution algorithm, real values of the parameters must be obtained from binary genes, in order to test them in flame simulations. Every gene is transformed into, $(a)_{10}$, a decimal number and then back into its real value (Table 1):

$$a = \frac{(a)_{10}(a_{\max} - a_{\min})}{2^{N_{bits}} - 1} + a_{\min}. \qquad (10)$$

	$(a)_2$		$(a)_{10}$		a
min	00000	$\longrightarrow$	0	$\longrightarrow$	3.92
	01011	$\longrightarrow$	11	$\longrightarrow$	3.97
max	11111	$\longrightarrow$	31	$\longrightarrow$	4.08

Table 1: Typical translation of one parameter from imposed minimum and maximum values $a \in [3.92, 4.08]$.

With these tools, the following procedure is applied to find the best parameters of the reduced chemistry:

1. M chromosomes (*i.e.* a population corresponding to a set of parameters values) are constructed by randomly selecting the values '0' or '1' for the $N_{bits}(a) + N_{bits}(\beta) + N_{bits}(E_a)$ bits of every chromosome (Fig. 2).

2. M flames are computed with the sets of chemical parameters (chromosomes), leading to $Y_k(x_j; c^p)$, $T(x_j; c^p)$, $u(x_j; c^p)$ for $p = 1, \cdots, M$.

3. The fitness function $f_2(c^p)$ (Eq. 8) is computed for every chromosome c^p. One of the objectives is to enhance in the solution evolution the influence of the parameters value having the highest fitness, therefore providing flame solutions the closest to the reference one. To do so, the chromosomes are simply stored in M
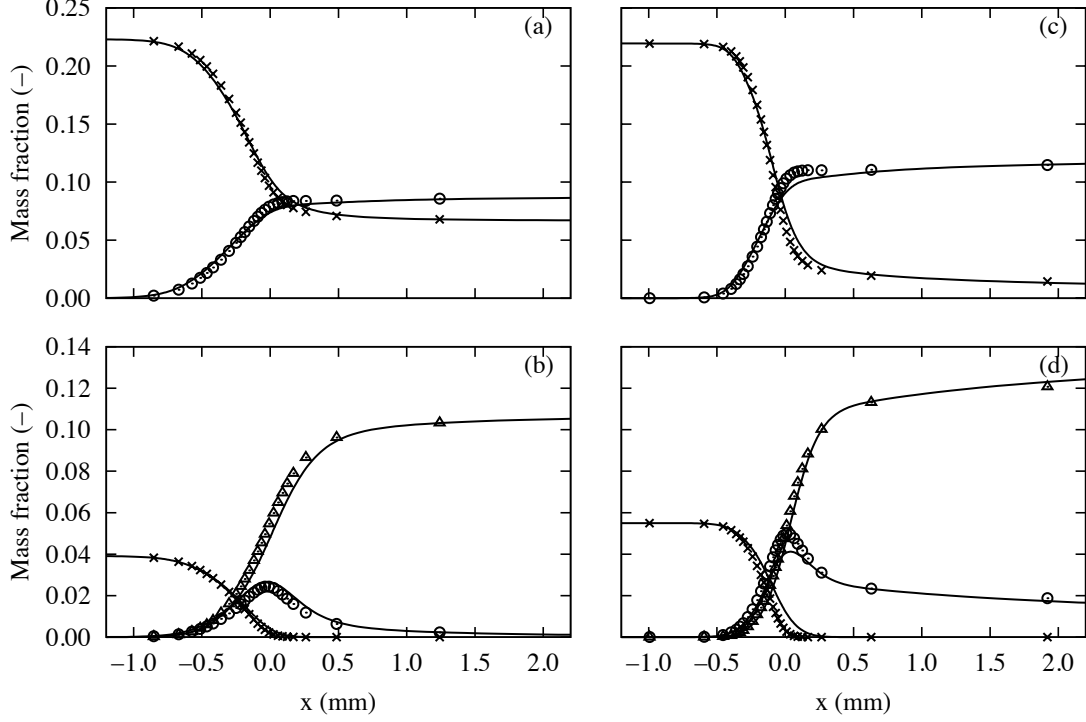
3

Figure 5: Species mass fractions. (a) & (b): $\phi = 0.7$. (c) & (d): $\phi = 1.0$. (a) & (c), $\times$: O_2. $\odot$: H_2O. (b) & (d), $\times$: CH_4. $\odot$: CO. $\triangle$: CO_2. Symbols: GRI-mech [9]. $-$: present scheme (Table 3).

boxes according to growing order of their fitness, the integer $I_b(c^p) \in [1, M]$ denotes the box position of the p-th chromosome. Then, a linear ranking procedure is applied to the chromosomes according to their fitness. The normalized rank $r(c^p)$ of the p-th chromosome reads:

$$r(c^p) = \frac{\sum_{\ell=1}^{p} \mathcal{F}(c^\ell)}{\sum_{m=1}^{M} \mathcal{F}(c^m)}, \qquad (11)$$

with

$$\mathcal{F}(c^p) = 2 - S_P + 2(S_P - 1)\frac{I_b(c^p) - 1}{M - 1}, \quad (12)$$

where a selective pressure $S_P = 2$ was chosen [12].

4. The population is advanced in time, in other words a new set of M chemical parameters (chromosomes) is built, according to three substeps:

(i) First, a crossover procedure is applied. $M/2$ pair of chromosomes are chosen. A random real number $\xi \in [0, 1]$ is generated and the p-th chromosome is selected if $\xi \in]r(c^\ell), r(c^p)]$, where the ℓ-th chromosome is ranked just below the p-th one (i.e. $I_b(c^p) - 1 = I_b(c^\ell)$). This procedure ensures that the chromosomes having a higher fitness are more likely to be selected. Once a given pair is selected, it has a probability of 30% to be left unchanged and a probability of 70% to undergo the crossover operation, in which two sets of parameters exchange a portion of their digits so as to create two new chromosomes (Table 2). The crossover point in the chromosome is also randomly selected.

01011 10010 00100 $\rightarrow$ 01011 10000 01001
11001 11000 01001 $\rightarrow$ 11001 11010 00100

Table 3: Crossover with one cutoff point.

(ii) Second, a random mutation is applied. Every digit of every chromosome may undergo a switch of it value ($0 \rightarrow 1$) or ($1 \rightarrow 0$), with a probability of 0.2%.

(iii) Third, to avoid damaging the best solution during these processes, an elitism operation is applied that consists of simply copying the best solution (e.g. $I_b(c^p) = M$ and also called 'elite') to the next generation.

A new set of chemical parameter is then available, which is used to repeat the process from point 2. The normalized variance of the fitness,

$$S_f = \frac{\sum_{p=1}^{M} f_2^{+2}(c^p) - \frac{1}{M}\left(\sum_{p=1}^{M} f_2^{+}(c^p)\right)^2}{4\sum_{p=1}^{M} f_2^{+}(c^p)\left(1 - \frac{1}{M}\sum_{p=1}^{M} f_2^{+}(c^p)\right)}, \quad (13)$$

where $f_2^{+} = (f_2 - f_{2_{\min}})/(f_{2_{\max}} - f_{2_{\min}})$ is monitored and the solution is considered converged
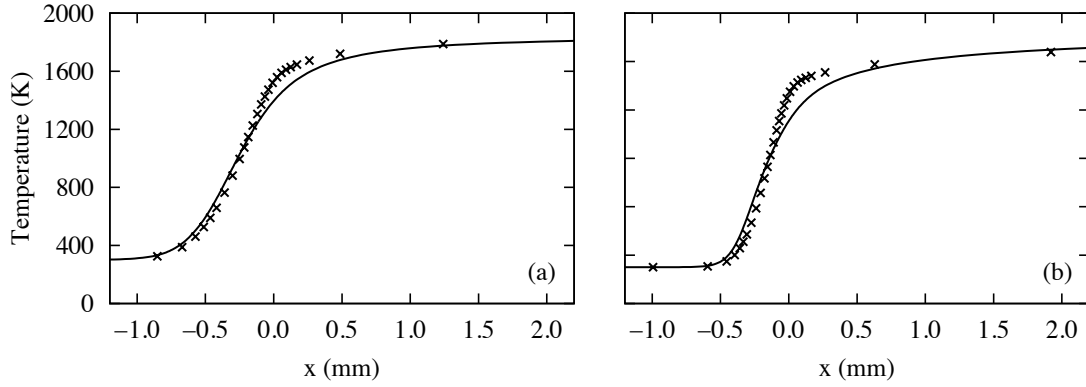
4

Figure 6: Temperature profile, (a): $\phi = 0.7$. (b): $\phi = 1.0$. $\times$: GRI-mech [9]. $-$: Present scheme (Table 3).

	Reaction		$\mathcal{A}^*$	β	E_a^*	n_1	n_2
C1	$CH_4 + 0.5\,O_2 \rightarrow CO + 2\,H_2$		2.592e15	-0.50	32099.41	0.799	0.732
H1	$H + O_2 \rightleftharpoons OH + O$		3.520e16	-0.70	17069.80		
H2	$H_2 + O \rightleftharpoons OH + H$		5.060e04	2.67	6290.63		
H3	$H_2 + OH \rightleftharpoons H_2O + H$		1.170e09	1.30	3635.28		
H4	$H + O_2 + M \rightarrow HO_2 + M$	k_0	5.750e16	-1.40	0.00		
		k_∞	4.650e12	0.44	0.00		
H5	$HO_2 + H \rightarrow 2\,OH$		7.080e13	0.00	294.93		
H6	$HO_2 + H \rightleftharpoons H_2 + O_2$		1.660e13	0.00	822.90		
H7	$HO_2 + OH + M \rightarrow H_2O + O_2 + M$	k_0	7.000e09	0.00	-1094.65		
		k_∞	4.500e14	0.00	10929.70		
H8	$H + OH + M \rightleftharpoons H_2O + M$		4.000e19	-2.00	0.00		
H9	$2H + M \rightleftharpoons H_2 + M$		1.300e15	-1.00	0.00		
C2	$CO + H_2O \rightleftharpoons CO_2 + H_2$		3.325e16	-0.60	32914.37	1.107	1.102

Table 2: Arrhenius coefficients for the reduced scheme. *Units are mol, s, cm^3, cal and K. n_1 and n_2 take stoichiometric values in the Hα reaction. Three body reactions see [2].

when its variation between two iterations is below a few percent.

Application to methane-air lean premixed combustion

The strategy is now applied to build a reduced mechanism for lean methane-air combustion. This mechanism is splitted in three parts: an oxydation reaction for CH_4, a set of reactions for hydro-oxygenated species (H_2/O_2 mechanism), and, a water-gas shift reaction. The parameters for both oxydation (C1 in Table 3) and the water-gas shift reactions (C2 in Table 3) are optimized. The constants of the H_2-O_2 mechanism (Hα in Table 3) are directly taken from [2]. With this approach, the hydrocarbon is involved in two reactions only, while the intermediate radical pool is controlled by nine elementary reactions. To focus on the optimization of the hydrocarbons reactions, no steady-state hypothesis is introduced in the H_2-O_2 mechanism, hypothesis that could however be used to reduce even more the mechanism size.

The initial condition for the chemical parameters is obtained with the first guess procedure reported above, with CH_4 and CO_2 used in Eq. (5). Figure 3 shows that a good match with the detailed mechanism is reached already for chemical sources of major species.

The number of chromosome is set to $M = 25$ (24 plus one elite). The precision is $\Delta a = \Delta \beta = \Delta n_1 = \Delta n_2 = 0.001$ and $\Delta E_a = 1$, leading to chromosomes of approximatively 100 bits. The method converges after about 15 iterations, which implies the simulations of about 375 flames. The required CPU

5

time is about 1hour on a laptop.

The resulting flame speed response versus equivalence ratio is shown in Fig. 4. The major species profiles are given in Fig. 5 and temperature in Fig. 6. The magnitude and the slopes of the profiles are captured. The present mechanism coefficients were obtained in a fully automated way, opening new perspectives to construct mechanism for specific operating conditions.

Conclusion

An automated approach based on a genetic-algorithm has been proposed to obtain chemical parameters from reference detailed chemistry sources and species profiles. The reduced mechanism is based on three parts: a global-step for fuel oxidation, a more detailed sub-mechanism for H_2/O_2 reactions and a global-step for water-shift. This method will be improved by coupling it with an additional automated approach to generate the reduced mechanism itself, prior to the optimization of its constants.

References

[1] N. Peters. Numerical and asymptotic analysis of systematically reduced reaction schemes for hydrocarbon flames. In Roland Glowinski, Bernard Larrouturou, and Roger Temam, editors, *Numerical Simulation of Combustion Phenomena*, volume 241 of *Lecture Notes in Physics*, pages 90–109. Springer Berlin Heidelberg, 1985.

[2] P. Boivin, C. Jiménez, A. L. Sánchez, and F. A. Williams. An explicit reduced mechanism for H_2 air combustion. *Proc. Combust. Inst.*, 33(1):517 – 523, 2011.

[3] W. P. Jones and R. P. Lindstedt. Global reaction schemes for hydrocarbon combustion. *Combust. Flame*, 73(3):233 – 249, 1988.

[4] N. Sikalo, O. Hasemann, C. Schulz, A. Kempf, and I. Wlokas. A genetic algorithm-based method for the automatic reduction of reaction mechanisms. *Int. J. Chem. Kinetics*, 46(1):41–59, 2013.

[5] A. Abou-Taouk, I. Sigfrid, W. Ronald, and L.-E. Eriksson. A four-step global reaction mechanism for CFD simulations of flexi-fuel burner for gas turbines. In *Turbulence, Heat and Mass Transfer 7*, volume 7, pages 785–788, 2012.

[6] B Farcy, A. Abou-Taouk, L. Vervisch, P. Domingo, and N. Perret. Two approaches of chemistry downsizing for simulating selective non catalytic reduction denox process. *Fuel*, 118:291–299, 2014.

[7] W. Polifke, W. Geng, and K. Döbbeling. Optimization of rate coefficients for simplified reaction mechanisms with genetic algorithms. *Combust. Flame*, 113(1–2):119 – 134, 1998.

[8] John H. Holland. *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA, USA, 1992.

[9] G. P. Smith, D. M. Golden, M. Frenklach, N. W. Moriarty, B. Eiteneer, M. Goldenberg, C. T. Bowman, R. K. Hanson, S. Song, W. C. Gardiner, V. V. Lissianski, and Z. Qin. Technical report, 1999. http://www.me.berkeley.edu/gri-mech/.

[10] C. F. Curtiss and J. O. Hischfelder. Transport properties of multicomponent gas mixtures. *Journal of Chemical Physics*, 17:550, 1949.

[11] D. Goodwin. Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes. http://code.google.com/p/cantera, 2009.

[12] N. M. Razali and J. Geraghty. Genetic algorithm performance with different selection straggles in solving tsp. In *Proceedings of the World Congress on Engineering*, volume II, London, 2011. WCE.