

# Reduction of linear kinetic systems with multiple scales

A. Blouza<sup>a</sup>, F. Coquel<sup>b</sup>, F. Hamel<sup>b\*</sup>

<sup>a</sup> Laboratoire d'Analyse et Modélisation Stochastique, Université de Rouen

<sup>b</sup> Laboratoire d'Analyse Numérique, Université Paris VI

**Abstract.** – We present a simple and general reduction algorithm for stiff monomolecular kinetic systems. The reduction is based on algebraic techniques and consists in eliminating the fastest dynamics in the initial system without any change of basis. This process is systematic and is not based on chemical conventional assumptions or on singular perturbation techniques. Systems can be reduced even if they are not in the Tikhonov form. This reduction process is applied to kinetic systems with kinetic constants belonging to different scales. Error estimates for all species are given. Numerical tests are performed.

## 1 Introduction

Reacting flows occurring in a large variety of situations like air pollution, combustion processes or biochemistry can be modeled by mathematical equations where chemistry, transport and diffusion phenomena are considered together. For a faithful modelling, complex chemical networks have to be taken into account. The current air pollution models involve tens of chemical species and reactions [1], [3]. On the other hand, the chemical reactions usually have widely different time scales and the numerical resolution of such stiff chemical systems can reach 80%, or more, of the computational time for models with simple transport and diffusion phenomena. A first step is to isolate the chemistry. In the past decade, special effort has been carried out to build reduced systems approximating the detailed chemical mechanism with the best possible accuracy. Our goal in this paper is to define a systematic algorithm, different from the usual reduction methods, in order to reduce any monomolecular isothermal chemical kinetics.

At the level of modelling, two main conventional reduction methods based on chemical observations, the “quasi-steady state approximation” (q.s.s.a.) and the “partial equilibrium assumption” have been successfully used to reduce kinetic systems [2], [21], [22], [23], [27], [32]. The q.s.s.a. asserts that the consumption and production rates of some well-chosen “quasi-steady state species” (q.s.s. species) are high but very close to each other so that they cancel each other out. The q.s.s. species are given from algebraic relations which are obtained by zeroing the right-hand side of the first- or second-order time-derivatives of their concentrations. The other species obey a reduced differential system. To enhance the chemical intuition, the q.s.s. species can be found from numerical tests [7]. On the other hand, the partial-equilibrium

---

\*Part of this work was carried out during the year 1997-1998 at the Department of Mathematics at M.I.T.

approximation leads to algebraic relations by zeroing the first-order (or second-order) time-derivatives of the rates of the reversible reactions whose direct and opposite sides are high but very close to each other.

From a mathematical point of view, the stiffness of the chemical systems can usually be eliminated with two methods. The first approach consists in changing the basis of the species without reducing the number of unknowns. The reduced systems involve non-pure species (as those obtained from lumping techniques [11], [24]). The second approach consists in partitioning the species into slow and fast ones and using some constraints to express the fast species. That can be done by writing the original system under the standard Tikhonov form, with two “slow/fast” sub-systems involving two different scales:  $\dot{x}_1 = v_1(x_1, x_2, \varepsilon)$ ,  $\varepsilon \dot{x}_2 = v_2(x_1, x_2, \varepsilon)$ . Under some semi-stability conditions for the “fast part”  $v_2$  along the invariant manifold  $\{v_2(x_1, x_2, 0) = 0\}$ , singular perturbation techniques guarantee that the solution  $x(t) = (x_1(t), x_2(t))$  can be developed in the time scales  $t$  and  $t/\varepsilon$  and that the initial system can be approximated by the reduced system  $\dot{x}_1 = v_1(x_1, x_2, 0)$ ,  $v_2(x_1, x_2, 0) = 0$  [6], [9], [13], [19], [29]. This manifold can be pre-tabulated and its dimension can be given by partitioning the spectrum of the Jacobian matrix of the source term [17], [20]. However, the concrete choice of the fast and slow species to get a Tikhonov form is not obvious in general and may be based on the q.s.s.a. [9], [26], [28] or may also require a change of basis [31]. An alternative way of finding the invariant manifold has been developed by Lam and Goussis [14], [15]. It consists in expanding the chemical terms into a sum of terms which are ordered by increasing time scales and in zeroing them recursively, but without reducing the size of the differential system. We also refer to [25], [30] on other methods of expanding the source terms by keeping the size of the differential system unchanged.

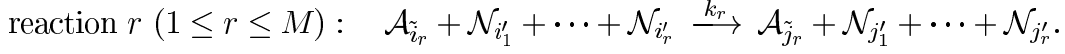
In this paper, our aim is to find a systematic and constructive reduction method for any kinetic system where the chemistry is monomolecular with respect to a set of limitant species whose concentrations have to be evaluated. We do not use either the quasi-steady state approximation or the partial-equilibrium assumption. As for the second approach described above, our goal is to construct, without any change of basis, a smaller and no longer stiff differential system involving “slow” species and to express the other “fast” species from the slow ones. As presented in section 2, the partitioning of the species is completely algorithmic and explicit. Although the starting point consists, like in [14], [15], in finding the fastest dynamics in the right-hand side of the original system, this is done in order to select some “slow” species and to eventually get a smaller differential system involving these species. Moreover, this reduction process can be done even if the system is not in the Tikhonov form. Unlike some optimization techniques used in combustion and pollution to obtain simplified models where only some species or group of species are evaluated [24], all species are here evaluated and time-global error estimates between the detailed and the reduced systems are given. This reduction mechanism is applied to kinetic systems with multiple scales, the scales being naturally given by the kinetic constants. Under some assumptions, we especially give a characterization of the size of the reduced system and of the fastest species. In section 5, we perform numerical tests by comparing the so-built reduced system with the detailed one and with the conventional quasi-steady state assumption. We study a theoretical example with three scales and a realistic example arising in biochemistry.

Some work has still to be carried out to deal with reactions which are at least bi-molecular in terms of the limitant species. For such systems, local instabilities may occur.

## 2 Definition of the reduction process and main results

### 2.1 Framework

Consider an isothermal mixture of  $N$  species  $\mathcal{A}_1, \dots, \mathcal{A}_N$  with relative concentrations  $y_i$ . There may also be some species  $\mathcal{N}_i$  in excess (like  $O_2$  or  $N_2$  in the air), whose concentrations are assumed to be constant. We assume that  $M$  chemical reactions take place simultaneously in the mixture and that each is mono-molecular with respect to the species  $\mathcal{A}_i$ , *i.e.*,



A reversible reaction  $\mathcal{A}_i \rightleftharpoons \mathcal{A}_j$  is represented by two reactions  $\mathcal{A}_i \rightarrow \mathcal{A}_j$  and  $\mathcal{A}_j \rightarrow \mathcal{A}_i$ . The rate of the reaction  $r$  is  $\omega_r = k_r y_{\tilde{i}_r}$  where, in this isothermal context,  $k_r$  is a given positive kinetic constant. We see that for each reaction  $r$  there is a unique pair  $(i, j)$  such that  $\mathcal{A}_i$  is a reactant and  $\mathcal{A}_j$  is a product. If the same pair  $(i, j)$  occurs in different reactions  $r_1, \dots, r_p$ , we combine these reactions together into a single reaction  $r$  by setting  $k_r = k_{r_1} + \dots + k_{r_p}$ . Then, a pair  $(\mathcal{A}_i, \mathcal{A}_j)$  can appear in at most one reaction as a pair reactant-product for such a reaction.

The evolution in time of each concentration  $y_i$  is given by the law of mass action:

$$\frac{dy_i}{dt} = - \sum_{r, \tilde{i}_r=i} \omega_r + \sum_{r, \tilde{j}_r=i} \omega_r, \quad i = 1, \dots, N.$$

In other words, the concentrations' vector  $y = (y_1, \dots, y_N)^T$  satisfies the following system  $\frac{dy}{dt} = S \omega$  where  $\omega = (\omega_1, \dots, \omega_M)^T$  is the reaction rates' vector and the rectangular stoichiometric matrix  $S$  of size  $N \times M$  is defined by:  $S_{i,r} = -1$  if  $i = \tilde{i}_r$ ,  $S_{i,r} = 1$  if  $i = \tilde{j}_r$  and  $S_{i,r} = 0$  otherwise. Since  $\omega_r = k_r y_{\tilde{i}_r}$ , the vector  $y(t)$  solves the differential system

$$(S) \quad \frac{dy}{dt} = Jy, \quad t \geq 0, \quad y(0) \text{ given} \quad (2.1)$$

where the real square matrix  $J$  of size  $N \times N$  is given by:

$$\begin{cases} J_{ii} = - \sum_{r, \tilde{i}_r=i} k_r, \\ J_{ij} = \begin{cases} 0 & \text{if } i \neq j \text{ and if } \forall r \ (\tilde{i}_r, \tilde{j}_r) \neq (j, i), \\ k_r & \text{if } i \neq j \text{ and if } \exists r \ (\tilde{i}_r, \tilde{j}_r) = (j, i). \end{cases} \end{cases} \quad (2.2)$$

Note that if  $i \neq j$  then there exists at most one reaction  $r$  such that  $(\tilde{i}_r, \tilde{j}_r) = (j, i)$ . In the sequel, one assumes that, up to normalization,  $\sum_{i=1}^N y_i(0) = 1$ .

The matrix  $J$  is said to be kinetic ([16], [28]): it satisfies

$$\begin{aligned} \text{(i)} \quad & \forall 1 \leq i \leq N, \quad J_{ii} \leq 0 \\ \text{(ii)} \quad & \forall 1 \leq i, j \leq N, \ i \neq j, \quad J_{ij} \geq 0 \\ \text{(iii)} \quad & \forall 1 \leq j \leq N, \ \sum_{i=1}^N J_{ij} = 0 \end{aligned} \quad (2.3)$$

and one has ([5], [28]): 1) the matrix  $J$  is “semi-stable”: all its nonzero eigenvalues have negative real parts and  $\text{Ker}(J) = \text{Ker}(J \cdot J)$ , 2) zero is an eigenvalue of  $J$  with multiplicity equal to the number of invariants in (2.1), and  $\mathbb{C}^N = \text{Im}(J) \oplus \text{Ker}(J)$ , 3) if  $\lambda$  is an eigenvalue of  $J$ , then  $\bar{\lambda}$  is also an eigenvalue and  $|\lambda| \leq 2 \max_{1 \leq i \leq N} |J_{ii}|$ , 4) lastly, any solution of the Cauchy problem (2.1) has a finite limit as  $t \rightarrow +\infty$ .

## 2.2 Definition of the reduction process

We are now going to eliminate the fastest dynamics of (S) and the associated species. The eigenvalues of the matrix  $J$  (which may not be diagonalisable) can be ordered so that

$$\Re(\lambda_1) \leq \dots \leq \Re(\lambda_m) < \lambda_{m+1} = \dots = \lambda_N = 0. \quad (2.4)$$

Since the matrix  $J$  is real, its eigenvalues are conjugated and two conjugated eigenvalues have the same multiplicity. Hence, if  $\lambda \notin \mathbb{R}$  is an eigenvalue of  $J$  with multiplicity  $\mu$ , one can assume that  $\Im(\lambda) > 0$  and that  $\lambda_i = \lambda$ ,  $\lambda_{i+1} = \bar{\lambda}$ ,  $\lambda_{i+2} = \lambda, \dots, \lambda_{i+2\mu-1} = \bar{\lambda}$  for some  $i$ .

Call  $J^0 = J$  and choose a nonzero left eigenvector  $b^0 \in \mathbb{C}^N$  of  $J^0$  for the eigenvalue  $\lambda_1$ . Let  $i_1$  be the smallest integer such that  $|b_{i_1}^0| = \max |b_i^0|$  and define the matrix  $J^1 = (J_{ij}^1)_{i,j \neq i_1}$  by

$$\forall 1 \leq i, j \leq N, \quad i, j \neq i_1, \quad J_{ij}^1 = J_{ij}^0 - \frac{b_j^0}{b_{i_1}^0} J_{i_1 i}^0. \quad (2.5)$$

The choice of  $i_1$  is enlightened in Remark 4.3 in the case of a system with multiple scales. The complex valued matrix  $J^1$ , which only depends on  $J^0 = J$  and  $b^0$ , is called the ‘‘reduced matrix at the first step’’. Notice that if  $b^0$  is replaced with  $\alpha b^0$  ( $\alpha \in \mathbb{C}$ ), then  $J^1$  is unchanged.

Let us assume temporarily that the matrix  $J^1$  admits the same eigenvalues as  $J$  except  $\lambda_1$  (see Theorem 2.1 below). Now, two cases may occur:

– either  $\lambda_1 \in \mathbb{R}$ . Since  $J$  is real, we can then assume that the vector  $b^0$  is real. The matrix  $J^1$  is then real and one then considers a nonzero left eigenvector  $b^1$  of  $J^1$  for the eigenvalue  $\lambda_2$ .

– or  $\Im(\lambda_1) > 0$ . In this case, the matrix  $J^1$  may not be real. On the other hand, we know that  $\lambda_2 = \bar{\lambda}_1$ . Let us define the vector  $b^1 \in \mathbb{C}^{N-1}$  by

$$\forall i \neq i_1, \quad b_i^1 = \bar{b}_i^0 - \frac{b_i^0}{b_{i_1}^0} \bar{b}_{i_1}^0. \quad (2.6)$$

We claim that (see the proof in section 3.1)

$$b^1 \text{ is a nonzero left eigenvector of } J^1 \text{ for the eigenvalue } \lambda_2 = \bar{\lambda}_1. \quad (2.7)$$

By induction, at any step  $k < N$ , we can construct a nonzero left eigenvector  $b^{k-1}$  of the matrix  $J^{k-1}$  for the eigenvalue  $\lambda_k$ , by assuming temporarily that the eigenvalues of  $J^{k-1}$  are  $\lambda_k, \dots, \lambda_N$ . We then define the smallest integer  $i_k \in \{1, \dots, N\} \setminus \{i_1, \dots, i_{k-1}\}$  such that  $|b_{i_k}^{k-1}| = \max_{i \neq i_1, \dots, i_{k-1}} |b_i^{k-1}|$ . Lastly, we define the ‘‘reduced matrix’’  $J^k = (J_{ij}^k)_{i,j \neq i_1, \dots, i_k}$  by

$$\forall 1 \leq i, j \leq N, \quad i, j \neq i_1, \dots, i_k, \quad J_{ij}^k = J_{ij}^{k-1} - \frac{b_j^{k-1}}{b_{i_k}^{k-1}} J_{i_k i}^{k-1}. \quad (2.8)$$

We point out that if  $\lambda_k$  is real, then  $b^{k-1}$  can be assumed to be real ; if  $\Im(\lambda_k) > 0$ , then  $b^{k-1}$  is chosen arbitrarily but if  $\Im(\lambda_k) < 0$ , then  $b^{k-1}$  is chosen in terms of  $b^{k-2}$  and on  $\bar{b}^{k-2}$  as in (2.6):  $b_i^{k-1} = \bar{b}_i^{k-2} - (b_i^{k-2}/b_{i_{k-1}}^{k-2}) \bar{b}_{i_{k-1}}^{k-2}$  for each  $i \in \{1, \dots, N\} \setminus \{i_1, \dots, i_{k-1}\}$ .

The close relationship between the dynamics of  $J^0 = J$  and  $J^k$  is stated in the following

**Theorem 2.1** *For any  $k \geq 1$ , the matrix  $J^k \in \text{Mat}(\mathbb{C}^{N-k}, \mathbb{C}^{N-k})$  has the same eigenvalues as  $J$ , except  $\lambda_1, \dots, \lambda_k$ . The matrix  $J^k$  is also semi-stable. Furthermore, if  $k$  is such that  $\Im(\lambda_k) \leq 0$  (which also means that  $\Im(\lambda_{k+1}) \geq 0$ ), then  $J^k$  is a real matrix.*



**Remark 2.2** The matrices  $J^k$  are successive reduced forms of  $J$  but no longer involve the first  $k$  eigenvalues of  $J$  which have the most negative real parts. Moreover, the fundamental property of semi-stability is preserved and if the “eliminated” eigenvalues  $\lambda_1, \dots, \lambda_k$  are conjugated (*i.e.* if  $\Im(\lambda_k) \leq 0$ ), then  $J^k$  is still real. However, the reduced matrices may not be kinetic.

For any  $k$ , set  $\mathcal{K}_k = \{1, \dots, N\} \setminus \{i_1, \dots, i_k\}$ . Let us stop the reduction process at a step  $R$ ,  $1 \leq R \leq N - 1$ , such that  $\lambda_R \in \mathbb{R}$  or  $\Im(\lambda_R) < 0$ : the first  $R$  eigenvalues  $\lambda_1, \dots, \lambda_R$  of  $J$  are conjugated. Since the reduced matrix  $J^R$  only contains the  $N - R$  eigenvalues of  $J$  which have the lowest real parts (in absolute values), we choose to approximate the solutions  $y(t)$  of the Cauchy problem (2.1) by the following algebraic-differential system with unknowns  $y^R(t)$ :

$$(S^R) \quad \begin{cases} \frac{dy^R}{dt} = J^R y^R(t), \quad t \geq t_0, \quad (y_i^R(t_0))_{i \in \mathcal{K}_R} \text{ given,} & \text{for } (y_i^R)_{i \in \mathcal{K}_R} = (y_i^R)_{i \neq i_1, \dots, i_R} \\ y_{i_k}^R(t) = - \sum_{i \in \mathcal{K}_k} \frac{b_i^{k-1}}{b_{i_k}^{k-1}} y_i^R(t) & \text{if } \lambda_k \in \mathbb{R} \text{ or } \Im(\lambda_k) < 0, \quad k \leq R \\ y_{i_k}^R(t) = - \sum_{i \in \mathcal{K}_{k+1}} \left( \frac{b_i^{k-1}}{b_{i_k}^{k-1}} - \frac{b_{i_{k+1}}^{k-1}}{b_{i_k}^{k-1}} \frac{b_i^k}{b_{i_{k+1}}^k} \right) y_i^R(t) & \text{if } \Im(\lambda_k) > 0, \quad k \leq R. \end{cases}$$

System  $(S^R)$  is divided into two parts: first, a differential – and real – system for  $(y_i^R)_{i \in \mathcal{K}_R}$ , and, second,  $R$  algebraic relations giving  $(y_i^R)_{i=i_R, \dots, i_1}$ , in terms of the other species. More precisely,  $y_{i_R}^R$  is calculated in terms of  $(y_i^R)_{i \in \mathcal{K}_R}$ , next  $y_{i_{R-1}}^R$  is calculated in terms of  $(y_i^R)_{i \in \mathcal{K}_{R-1}}$ , and so on until  $y_{i_1}^R$ . If  $k+1 \leq R$  and  $\Im(\lambda_k) > 0$ , then  $y_{i_k}^R$  is actually given in terms of  $(y_i^R)_{i \in \mathcal{K}_{k+1}}$  (and not in terms of the  $(y_i^R)_{i \in \mathcal{K}_k}$  as in the other cases). The formula for  $y_{i_k}^R$  in the case  $\Im(\lambda_k) > 0$  is chosen so that it could involve real coefficients (*see* Theorem 2.3 below). However, due to the particular choice of the vector  $b^k$  (in terms of  $b^{k-1}$  and of  $\bar{b}^{k-1}$ ) in the case  $\Im(\lambda_{k+1}) < 0$ , it is easy to check that system  $(S^R)$  can be rewritten as

$$\begin{cases} \frac{dy^R}{dt} = J^R y^R(t), \quad t \geq t_0, \quad (y_i^R(t_0))_{i \in \mathcal{K}_R} \text{ given,} & \text{for } (y_i^R)_{i \in \mathcal{K}_R} \\ y_{i_k}^R(t) = - \sum_{i \in \mathcal{K}_k} \frac{b_i^{k-1}}{b_{i_k}^{k-1}} y_i^R(t) & \text{if } k \leq R. \end{cases} \quad (2.9)$$

Lastly,  $t_0$  is a new initial time for the reduced system  $(S^R)$ . It is specified in Theorem 2.3 below. If the eigenvalues  $\lambda_{R+1}, \dots, \lambda_N$  of  $J^R$  remain bounded and if  $\lambda_1, \dots, \lambda_R$  are very high, then the differential system for  $(y_i^R)_{i \in \mathcal{K}_R}$  is not stiff anymore. Under that assumption,  $t_0$  can be viewed as an exit time from a boundary layer for the so-called “fast” species  $y_i$ ,  $i = i_1, \dots, i_R$ .

The particular form met by this algebraic-differential system  $(S^R)$  is relevant in the sense that, given a number  $R$  of eliminated species, problem  $(S^R)$  provides a good approximation of the solutions  $(y_i)_{1 \leq i \leq N}$  of system  $(S)$  for times larger than  $t_0$ :

**Theorem 2.3** *i) If  $R$  is such that  $\lambda_R \in \mathbb{R}$  or  $\Im(\lambda_R) < 0$ , then  $J^R$  is a real matrix and all the coefficients of the algebraic relations governing the  $y_{i_k}^R$  ( $k = 1, \dots, R$ ) in  $(S^R)$  are real.*

*ii) Besides, for any  $R$ , there exists a constant  $C$  such that, for any time  $t_0 \geq \max(0, \max_{1 \leq k \leq R} (|\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|))$  and for any  $h \geq 0$ ,*

$$\begin{aligned} & \text{if } (\forall i \in \mathcal{K}_R, |y_i^R(t_0) - y_i(t_0)| \leq h), \\ & \text{then } (\forall 1 \leq i \leq N, \forall t \geq t_0, |y_i^R(t) - y_i(t)| \leq C (h + |\Re(\lambda_R)|^{-1})). \end{aligned} \quad (2.10)$$

In other words, the smaller the error between the exact solution and the reduced solution for the species  $(y_i)_{i \in \mathcal{K}_R}$  at a given and precisely estimated exit time  $t_0$  is and the larger the last eliminated dynamics  $\mathfrak{R}(\lambda_R)$  is, then the smaller the error between the exact solution and the reduced solution for all species at any further time is. In practice,  $y_i^R(t_0)$  can be obtained by solving numerically the initial system  $(S)$  on the interval  $[0, t_0]$  and  $h$  can be viewed as a numerical error with respect to the exact solution at the time  $t_0$ .

### 2.3 Applying the reduction to kinetic systems with multiple scales

In many situations, the kinetic constants  $k_r$  of the chemical reactions range in very different scales. The largest (resp. smallest) kinetic constants  $k_r$  correspond to the fastest (resp. slowest) reactions. In the sequel, we assume that there exist  $\varepsilon \in (0, 1)$  (considered to be small), two positive constants  $0 < C_1 < C_2$  (independent from  $\varepsilon$ ),  $S$  real numbers  $p_1 > \dots > p_S$  (not necessarily of the same sign) and a partition of the set  $\{1, \dots, M\}$  into  $S$  subsets  $\mathcal{R}_1, \dots, \mathcal{R}_S$ , such that, for any  $s \in \{1, \dots, S\}$  and for any  $r \in \mathcal{R}_s$ ,  $k_r \varepsilon^{p_s}$  does not depend on  $\varepsilon$  and

$$\forall 1 \leq s \leq S, \forall r \in \mathcal{R}_s, \quad C_1 \varepsilon^{-p_s} \leq k_r \leq C_2 \varepsilon^{-p_s}.$$

The reactions in the group  $\mathcal{R}_1$  have the highest kinetic constants; the kinetic constants of the reactions of the group  $\mathcal{R}_2$  belong to the second highest scale, etc. Define  $\mathcal{I}_1 = \{1 \leq i \leq N, \exists r \in \mathcal{R}_1, \tilde{i}_r = i\}$  the set of reactants  $\mathcal{A}_i$  in at least one reaction of the group  $\mathcal{R}_1$  and, by induction,

$$\mathcal{I}_s = \{1 \leq i \leq N, i \notin \bigcup_{1 \leq t \leq s-1} \mathcal{I}_t, \exists r \in \mathcal{R}_s, \tilde{i}_r = i\}, \quad 2 \leq s \leq S$$

and  $\mathcal{I}_{S+1} = \{1, \dots, N\} \setminus \bigcup_{1 \leq s \leq S} \mathcal{I}_s$ . The sets  $\mathcal{I}_1, \dots, \mathcal{I}_{S+1}$  partition the set  $\{1, \dots, N\}$ . Lastly, call  $R_s = \#\mathcal{R}_s$  ( $1 \leq s \leq S$ ) and  $I_i = \#\mathcal{I}_i$  ( $1 \leq i \leq S+1$ ), where  $\#E$  means the cardinal of  $E$ . Even if it means permuting the subscripts of the species, we can always assume that  $\mathcal{I}_1 = \{1, \dots, I_1\}, \dots, \mathcal{I}_S = \{I_1 + \dots + I_{S-1} + 1, \dots, I_1 + \dots + I_S\}, \mathcal{I}_{S+1} = \{I_1 + \dots + I_S + 1, \dots, N\}$ .

Let us now define  $S$  square matrices  $A^s$  ( $s = 1, \dots, S$ ) of size  $N \times N$ , as follows:

- $\forall 1 \leq i \neq j \leq N$ :  $A_{ij}^s = \varepsilon^{p_s} J_{ij} = \varepsilon^{p_s} k_r$  if  $\exists r \in \mathcal{R}_s, (\tilde{i}_r, \tilde{j}_r) = (j, i)$ ;  $A_{ij}^s = 0$  otherwise,
- $\forall 1 \leq i \leq N$ ,  $A_{ii}^s = -\varepsilon^{p_s} \sum_{r \in \mathcal{R}_s, \tilde{i}_r = i} k_r$ .

For each  $1 \leq s \leq S$ , the matrix  $\varepsilon^{-p_s} A^s$  only takes into account the kinetic constants of the reactions of the group  $\mathcal{R}_s$ . Each matrix  $A^s$  is kinetic and satisfies (2.3). Furthermore, the columns  $C_i^s$  of  $A^s$  vanish for  $i \geq I_1 + \dots + I_s + 1$ .

The initial system  $\dot{y} = Jy$  can then be written as follows:

$$\dot{y} = Jy = (\varepsilon^{-p_1} A^1 + \dots + \varepsilon^{-p_S} A^S) y. \quad (2.11)$$

The scales  $\varepsilon^{-p_1}, \dots, \varepsilon^{-p_S}$  appear in the columns of  $J$ . This is very different from the case where the system is directly divided into a slow and a fast subsystem of the kind:  $\dot{x}_1 = v_1(x_1, x_2)$  and  $\dot{x}_2 = \varepsilon^{-1} v_2(x_1, x_2)$  – the Tikhonov form – (see *e.g.* [6], [9], [29], or [8], [12] in the case of a finite number of such subsystems). Notice that if, for instance,  $\lambda_1$  is in the scale  $\varepsilon^{-1}$  and if the other eigenvalues are bounded as  $\varepsilon \rightarrow 0$ , the initial system (2.11) could be put in the Tikhonov form after a change of basis. But our goal is to find a reduced system involving pure species.

As in section 2.2, we can build a sequence of reduced matrices  $J^k$  by eliminating the fastest dynamics and some associated species. We now have to determine at which step  $R$  it seems

reasonable to stop the reduction process. Given a power  $p_q$  ( $1 \leq q \leq S$ ) and a real  $\alpha$  such that  $p_q > \alpha > p_{q+1}$  if  $q \neq S$  or  $p_S > \alpha$  if  $q = S$ , let  $R$  be the first step such that  $\Im(\lambda_R) \leq 0$  and

$$|\Re(\text{Tr}(J^R))| \leq \varepsilon^{-\alpha}. \quad (2.12)$$

Since  $\Re(\text{Tr}(J^R)) = \Re(\lambda_{R+1}) + \dots + \Re(\lambda_N)$  and  $\Re(\lambda_i) \leq 0$ , it follows that  $|\Re(\lambda_i)| \leq \varepsilon^{-\alpha}$  if  $i \geq R+1$ . Similarly, since the sequence  $(\Re(\lambda_i))$  is nondecreasing,  $|\Re(\lambda_i)| \geq \frac{1}{N-R-1}\varepsilon^{-\alpha}$  if  $i \leq R$ .

The following Theorem gives especially an estimation of the number of reduction steps and the nature of the eliminated species.

**Theorem 2.4** *i) Let  $I'_1$  be the rank of  $A^1$ . There exists  $\varepsilon_0 > 0$  such that if  $0 < \varepsilon < \varepsilon_0$ , then*

$$R \geq I'_1 \quad \text{and} \quad \forall 1 \leq k \leq I'_1, \quad 1 \leq i_k \leq I_1.$$

*ii) If  $\alpha$  is chosen in (2.12) such that  $p_1 > \alpha > p_2 + \beta(p_1 - p_2)$  with  $\beta = (I_1 - I'_1)/(I_1 - I'_1 + 1) \in [0, 1)$ , then there exist two positive constants  $C$  and  $\varepsilon_0$  such that if  $0 < \varepsilon < \varepsilon_0$ , then  $R = I'_1$ ,*

$$\begin{aligned} & \forall i \geq I'_1 + 1, \quad |\lambda_i| \leq C\varepsilon^{-p_2 - \beta(p_1 - p_2)} \\ & \begin{cases} \forall 2 \leq s \leq S, \forall j \in \mathcal{I}_s, \forall i \neq i_1, \dots, i_R, & |J_{ij}^R| \leq C\varepsilon^{-p_s} \\ \forall j \in \mathcal{I}_{S+1}, \forall i \neq i_1, \dots, i_R, & J_{ij}^R = 0 \end{cases} \end{aligned} \quad (2.13)$$

and  $J^R = o(\varepsilon^{-p_1})$  as  $\varepsilon \rightarrow 0$  in  $\text{Mat}(\mathbb{R}^{N-R}, \mathbb{R}^{N-R})$ .

There are at least  $I'_1$  steps to eliminate the fastest dynamics and the size of the differential system in  $(S^R)$  is at most of  $N - I'_1$ . Moreover, the “fastest” –eliminated algorithmically–  $I'_1$  species  $\mathcal{A}_{i_1}, \dots, \mathcal{A}_{i_{I'_1}}$  are reactants of at least one of the fastest reactions (group  $\mathcal{R}_1$ ).

If  $\alpha$  is such that  $p_1 > \alpha > p_2 + \beta(p_1 - p_2)$  and  $\varepsilon > 0$  is small enough, one then has an exact estimation of the number of reduction steps as well as very sharp estimates for the coefficients of  $J^R$ . In particular, the final reduced matrix  $J^R$  does not contain any term in the highest scale  $\varepsilon^{-p_1}$  and its eigenvalues belong to scales  $\mathcal{O}(\varepsilon^{-p_2 - \beta(p_1 - p_2)})$ . Therefore, the reduced system  $(S^R)$  is not stiff with respect to the initial system  $(S)$ . The scale  $\varepsilon^{-p_2 - \beta(p_1 - p_2)}$  can be viewed as a scale separating the highest dynamics  $\varepsilon^{-p_1}$  and the other ones.

**Remark 2.5** In the case  $I'_1 = I_1$ , then  $\alpha$  can be any number in  $(p_2, p_1)$  and, after reduction, all dynamics belong to the scale  $\mathcal{O}(\varepsilon^{-p_2})$ . This assumption  $I'_1 = I_1$  corresponds to the invertibility assumption in Tikhonov’s Theorem. However, systems of the type (2.11) can be reduced *even* if  $I'_1 < I_1$ , as for the following system

$$\frac{d}{dt} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -k_1\varepsilon^{-1} - k_3 & k_2\varepsilon^{-1} & 0 \\ k_1\varepsilon^{-1} & -k_2\varepsilon^{-1} & 0 \\ k_3 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad \text{associated to} \quad \left\{ \begin{array}{l} \mathcal{A}_1 \xrightleftharpoons{k_1\varepsilon^{-1}} \mathcal{A}_2 \\ \mathcal{A}_1 \xrightarrow{k_3} \mathcal{A}_3 \end{array} \right\},$$

where  $I'_1 = 1 < I_1 = 2$ . The reactions of the group  $\mathcal{R}_1$  only involve the first  $I_1$  species; hence, the sum of the first  $I_1$  lines of  $A^1$  is zero. This condition is sufficient for  $I'_1 < I_1$  to hold.

**Corollary 2.6** *Let  $\alpha \in (p_2 + \beta(p_1 - p_2), p_1)$  and assume  $p_1 > 0$ . Let  $\mu_{I'_1}$  be a nonzero eigenvalue of the matrix  $A^1$  whose real part is the closest to 0. There exists a constant  $C$  such that, for any  $\delta > |\Re(\mu_{I'_1})|^{-1}$ , there exists  $\varepsilon_0$  such that for any  $0 < \varepsilon < \varepsilon_0$  and for any  $t_\varepsilon \geq \delta\varepsilon^{p_1} |\ln \varepsilon^{p_1}|$ ,*

$$\left( \forall i \in \mathcal{K}_R, |y_i^R(t_\varepsilon) - y_i(t_\varepsilon)| \leq \varepsilon^{p_1} \right) \implies \left( \forall 1 \leq i \leq N, \forall t \geq t_\varepsilon, |y_i^R(t) - y_i(t)| \leq C\varepsilon^{p_1} \right).$$

This result—obtained thanks to Theorems 2.3 and 2.4—gives a precise estimation of the boundary layer exit time  $t_\varepsilon$  from the highest dynamics. Provided that the (numerical) error for the “slow” species  $(y_i)_{i \in \mathcal{K}_R}$  at this exit time is in the same scale  $\varepsilon^{p_1}$  as the time scale of the fastest dynamics, then the error with the exact solution stays in the same scale for all species at any further time  $t$ . Furthermore, the stiffer the initial system is, the smaller that error is. The values of  $y_i^R(t_\varepsilon)$  can be obtained by solving numerically the initial system  $(S)$  on the small boundary layer  $[0, t_\varepsilon]$ . Notice that on such an interval, the notion of fast and slow species is not suited.

**Remark 2.7** Under the assumptions of Corollary 2.6, the invariants of system  $(S)$  and the positivity of the concentrations are preserved in the reduced system  $(S^R)$  up to an error  $\mathcal{O}(\varepsilon^{p_1})$ .

### 3 General properties of the reduced systems

#### 3.1 Proof of Theorem 2.1

**Proof of the claim (2.7).** Let  $b^0$  be a nonzero left eigenvector of the matrix  $J$  for the eigenvalue  $\lambda_1$  and assume that  $\Im(\lambda_1) > 0$ . Let  $J^1$  be the matrix given by (2.5) and  $b^1$  be the vector in  $\mathbb{C}^{N-1}$  defined by (2.6). If the vector  $b^0$  is replaced with  $\gamma b^0$  ( $\gamma \in \mathbb{C}^*$ ), then  $J^1$  is not changed and  $b^1$  is replaced with  $\bar{\gamma} b^1$ . Hence, in order to prove that  $b^1$  is a left eigenvector of  $J^1$  for the eigenvalue  $\lambda_2$ , it is sufficient to consider the case where  $b_{i_1}^0 = 1$ . This means that  $b_j^1 = \bar{b}_j^0 - b_j^0$  for any  $j \neq i_1$  and  $J_{ij}^1 = J_{ij} - b_j^0 J_{ii_1}$  for any  $i, j \neq i_1$ .

First of all,  $b^1$  is not identically zero: otherwise  $b_j^0 \in \mathbb{R}$  for all  $j \neq i_1$ , then  $b^0$  would be a real and nonzero eigenvector of the real matrix  $J$  for the nonreal eigenvalue  $\lambda_1$ . This is impossible.

Let us now check that  $b^1 J^1 = \lambda_2 b^1 = \bar{\lambda}_1 b^1$ . For each  $j \neq i_1$ , we have

$$\sum_{k \neq i_1} b_k^1 J_{kj}^1 = \sum_{k \neq i_1} (\bar{b}_k^0 - b_k^0) (J_{kj} - b_j^0 J_{ki_1}) = \sum_{k \neq i_1} \bar{b}_k^0 J_{kj} - b_j^0 \sum_{k \neq i_1} \bar{b}_k^0 J_{ki_1} - \sum_{k \neq i_1} b_k^0 J_{kj} + b_j^0 \sum_{k \neq i_1} b_k^0 J_{ki_1}.$$

Since  $b^0 J = \lambda_1 J$  and since  $b_{i_1}^0 = 1$ , we deduce that

$$\sum_{k \neq i_1} b_k^1 J_{kj}^1 = (\bar{\lambda}_1 \bar{b}_j^0 - J_{i_1 j}) - b_j^0 (\bar{\lambda}_1 - J_{i_1 i_1}) - (\lambda_1 b_j^0 - J_{i_1 j}) + b_j^0 (\lambda_1 - J_{i_1 i_1}) = \bar{\lambda}_1 (\bar{b}_j^0 - b_j^0) = \bar{\lambda}_1 b_j^1.$$

**The matrix  $J^2$  is real if  $\Im(\lambda_1) > 0$ .** Under the above notation, let us now prove that the matrix  $J^2 = (J_{ij}^2)_{i, j \neq i_1, i_2}$  defined by (2.8) is real. Let  $i_2$  be the smallest integer in  $\{1, \dots, N\} \setminus \{i_1\}$  such that  $|b_{i_2}^1| = \max_{j \neq i_1} |b_j^1|$ . For any couple  $(i, j)$  such that  $i, j \neq i_1, i_2$ , one has

$$J_{ij}^2 = J_{ij} - b_j^0 J_{ii_1} - \frac{b_j^1}{b_{i_2}^1} (J_{ii_2} - b_{i_2}^0 J_{ii_1}) = J_{ij} + \left( \frac{b_j^1}{b_{i_2}^1} b_{i_2}^0 - b_j^0 \right) J_{ii_1} - \frac{b_j^1}{b_{i_2}^1} J_{ii_2}.$$

Remember that all the coefficients of  $J$  are real. Besides, for any  $j \neq i_1$ ,  $b_j^1 = \bar{b}_j^0 - b_j^0 \in i\mathbb{R}$  whence  $(b_j^1/b_{i_2}^1) \in \mathbb{R}$ . On the other hand, for any  $j \neq i_1$ ,

$$\frac{b_j^1}{b_{i_2}^1} b_{i_2}^0 - b_j^0 = \frac{(\bar{b}_j^0 - b_j^0) b_{i_2}^0 - b_j^0 (\bar{b}_{i_2}^0 - b_{i_2}^0)}{\bar{b}_{i_2}^0 - b_{i_2}^0} = \frac{\bar{b}_j^0 b_{i_2}^0 - b_j^0 \bar{b}_{i_2}^0}{\bar{b}_{i_2}^0 - b_{i_2}^0}.$$

In the right hand side, both numerator and denominator are purely imaginary. Hence,  $(b_j^1/b_{i_2}^1)b_{i_2}^0 - b_j^0 \in \mathbb{R}$  and eventually the matrix  $J^2$  is real.  $\square$

Furthermore, if  $\Im(\lambda_1) > 0$  and owing to the definition of  $(S^R)$ , the eliminated variables  $y_{i_1}^r$  and  $y_{i_2}^r$  depend algebraically on the  $y_j^r$ 's,  $j \neq i_1, i_2$ . The coefficients involved in these algebraic relations are:  $-b_j^1/b_{i_2}^1$  for  $y_{i_2}$  and  $-(b_j^0 - (b_j^1/b_{i_2}^1)b_{i_2}^0)$  for  $y_{i_1}$  (assuming that  $b^0$  is normalized so that  $b_{i_1}^0 = 1$ ). These coefficients turn out to be real from the above calculations.

Now assume temporarily that, at any step  $k$  of the reduction process, the eigenvalues of the matrix  $J^k$  are  $\lambda_{k+1}, \dots, \lambda_N$  (this is proved below in Theorem 2.1). If  $J^k$  is real and if  $\lambda_{k+1} \in \mathbb{R}$ , then the eigenvector  $b^k$  of  $J^k$  for the eigenvalue  $\lambda_{k+1}$  can be chosen in  $\mathbb{R}^{N-k}$  and  $J^{k+1}$  is then real. On the other hand, if  $J^k$  is real and if  $\Im(\lambda_{k+1}) > 0$ , then the above calculations yield that the matrix  $J^{k+2}$  is real. By induction,  $J^k$  is real at any step  $k$  such that  $\Im(\lambda_k) \leq 0$ .

Similarly, if  $\Im(\lambda_k) \leq 0$ , all the coefficients of the algebraic part of the reduced system  $(S^k)$  are real. This actually proves assertion i) of Theorem 2.3.

**Proof of Theorem 2.1.** Remember first that the eigenvalues of  $J$  can be ordered such that

$$\Re(\lambda_1) \leq \dots \leq \Re(\lambda_m) < \lambda_{m+1} = \dots = \lambda_N = 0$$

where  $m$  is the rank of  $J$ , and that  $\mathbb{C}^N = \text{Im}(J) \oplus \text{Ker}(J)$ . Since  $b^0 J = \lambda_1 J$ ,  $b^0$  is in  $\text{Im}(J)$ . Set  $b^{0,1} := b^0$ . From Schur reduction, there exists then a basis  $(b^{0,1}, \dots, b^{0,N})$  of  $\mathbb{C}^N$  and a matrix

$$T^0 = \begin{pmatrix} T' & 0 \\ 0 & 0 \end{pmatrix}$$

such that  $T'$  is lower triangular, its diagonal is  $(\lambda_1, \dots, \lambda_m)$ , and  $J = J^0 = P^0 T^0 (P^0)^{-1}$  with

$$P^0 = \begin{pmatrix} | & \dots & | \\ a^{0,1} & \dots & a^{0,N} \\ | & \dots & | \end{pmatrix} \quad \text{and} \quad (P^0)^{-1} = \begin{pmatrix} - & b^{0,1} & - \\ & \vdots & \\ - & b^{0,N} & - \end{pmatrix}.$$

Since  $J_{ij}^0 = \sum_{p,q \geq 1} a_i^{0,p} T_{pq}^0 b_j^{0,q}$  for each  $(i, j)$ , we get from the definition (2.5) of  $J^1$  that

$$\forall i, j \neq i_1, \quad J_{ij}^1 = \sum_{p,q \geq 1} a_i^{0,p} T_{pq}^0 (b_j^{0,q} - \frac{b_j^{0,1}}{b_{i_1}^{0,1}} b_{i_1}^{0,q}). \quad (3.1)$$

In the right hand side, the sum for  $q = 1$  is clearly equal to 0. Next, for  $p = 1$  and  $q \geq 2$ , we have  $T_{1q}^0 = 0$  since  $T^0$  is lower triangular. Hence, the sum (3.1) actually starts for  $p, q \geq 2$ . In other words,  $J^1 = P^1 T^1 Q^1$  where the matrix  $T^1$  is the submatrix of  $T^0$  obtained by dropping the first row and the first column, and where

$$P^1 = \begin{pmatrix} | & \dots & | \\ a^{1,2} & \dots & a^{1,N} \\ | & \dots & | \end{pmatrix}, \quad a_i^{1,p} = a_i^{0,p} \quad \forall i \neq i_1, p \geq 2, \quad (3.2)$$

$$Q^1 = \begin{pmatrix} - & b^{1,2} & - \\ & \vdots & \\ - & b^{1,N} & - \end{pmatrix}, \quad b_j^{1,q} = b_j^{0,q} - \frac{b_j^{0,1}}{b_{i_1}^{0,1}} b_{i_1}^{0,q} \quad \forall q \geq 2, j \neq i_1. \quad (3.3)$$

In particular, the matrix  $P^1$  is obtained from  $P^0$  by dropping the first column  $a^{0,1}$  and the  $i_1^{\text{th}}$  row. Let us now check that  $Q^1 = (P^1)^{-1}$ , *i.e.*  $Q^1 P^1 = I_{N-1}$ : for any  $p, q \geq 2$ , we have

$$b^{1,q} \cdot a^{1,p} = \sum_{i \neq i_1} (b_i^{0,q} - \frac{b_i^{0,1}}{b_{i_1}^{0,1}} b_{i_1}^{0,q}) a_i^{0,p} = (\delta_{p,q} - b_{i_1}^{0,q} a_{i_1}^{0,p}) - \frac{b_{i_1}^{0,q}}{b_{i_1}^{0,1}} (\delta_{p,1} - b_{i_1}^{0,1} a_{i_1}^{0,p}) = \delta_{p,q}.$$

In conclusion, the matrix  $J^1$  is similar to the matrix  $T^1 = \begin{pmatrix} T'^1 & 0 \\ 0 & 0 \end{pmatrix}$  where  $T'^1$  is lower triangular and its diagonal is  $(\lambda_2, \dots, \lambda_m)$ . This implies that the eigenvalues of  $J^1$  are those of  $J$  except  $\lambda_1$  and that  $b^{1,2}$  is a nonzero left eigenvector of  $J^1$  for the eigenvalue  $\lambda_2$ . This also yields that  $J^1$  is semi-stable. These properties hold good by induction at any step  $k$ .  $\square$

### 3.2 Proof of Theorem 2.3

First of all, the fact that the reduced system  $(S^R)$  is completely real if  $\Im(\lambda_R) \leq 0$  has been proved at the beginning of section 3.1 just before the proof of Theorem 2.1. The rest of this section is devoted to the proof of the error estimates (2.10) which is divided into several lemmas.

In the proof of Theorem 2.1, we considered a left nonzero eigenvector  $b^{0,1} = b^0$  of  $J$  for the eigenvalue  $\lambda_1$ . Next, we used a basis  $(b^{0,1}, \dots, b^{0,N})$  of  $\mathbb{C}^N$  and from this basis we built a basis of vectors  $(b^{1,2}, \dots, b^{1,N})$  of  $\mathbb{C}^{N-1}$  such that, for any  $q \geq 2$  and  $i \neq i_1$ ,  $b_i^{1,q} = b_i^{0,q} - (b_i^{0,1}/b_{i_1}^{0,1}) b_{i_1}^{0,q}$  and the vector  $b^{1,2}$  turned out to be a nonzero left eigenvector of  $J^1$  for the eigenvalue  $\lambda_2$ . In the reduction process defined in section 2.2, we only need a left eigenvector  $b^0 \in \mathbb{C}^N$  of  $J$  for the eigenvalue  $\lambda_1$ , next we need a left eigenvector  $b^1 \in \mathbb{C}^{N-1}$  of  $J^1$  for the eigenvalue  $\lambda_2$ , etc.

The following lemma states that the given vectors  $b^0, \dots, b^{k-1}$  could actually be given recursively from an orthogonal family  $(b^{0,1}, \dots, b^{0,k}) \in (\mathbb{C}^N)^k$ . In the sequel, we denote  $(a|b) = \sum_{1 \leq k \leq N} a_k \bar{b}_k$  and  $\|a\|^2 = (a|a)$  for any  $a, b \in \mathbb{C}^N$ .

**Lemma 3.1** *Let  $b^0 \in \mathbb{C}^N, \dots, b^{k-1} \in \mathbb{C}^{N-k+1}$  be nonzero left eigenvectors of the matrices  $J^0, \dots, J^{k-1}$  for the eigenvalues  $\lambda_1, \dots, \lambda_k$  (the matrices  $J^i$  are defined as in (2.8)).*

1) *There exists an orthogonal family of  $k$  nonzero vectors  $(b^{0,1}, \dots, b^{0,k}) \in (\mathbb{C}^N)^k$  such that*

$$\forall 0 \leq p \leq k-1, \quad b_{p+1}^p = b^p, \quad (3.4)$$

where the vectors  $b^{p,q} \in \mathbb{C}^{N-p}$ ,  $1 \leq p \leq k-1$ ,  $p+1 \leq q \leq k$  are defined recursively by

$$\forall 1 \leq p \leq k-1, \quad \forall p+1 \leq q \leq k, \quad \forall i \neq i_1, \dots, i_p, \quad b_i^{p,q} = b_i^{p-1,q} - (b_i^{p-1,p}/b_{i_p}^{p-1,p}) b_{i_p}^{p-1,q}. \quad (3.5)$$

2) *If a family  $(b^{0,1}, \dots, b^{0,k}) \in (\mathbb{C}^N)^k$  fulfills (3.4-3.5), then there exist  $\alpha_{q,r} \in \mathbb{C}$  such that*

$$\forall 1 \leq q \leq k, \quad b^{0,q} J = \lambda_q b^{0,q} + \alpha_{q,q-1} b^{0,q-1} + \dots + \alpha_{q,1} b^{0,1}.$$

**Remark 3.2** If a vector  $b^{0,q}$  is replaced with  $\gamma b^{0,q}$  ( $\gamma \in \mathbb{C}^*$ ), then the vector  $b^{q-1} = b^{q-1,q}$  is multiplied by  $\gamma$  whereas the vectors  $b^k$  for  $k = 0, \dots, q-2$  are unchanged. Since each matrix  $J^q$  is unchanged if  $b^{q-1}$  is multiplied by any complex number, we can finally assume that the basis  $(b^{0,1}, \dots, b^{0,N})$  is unitary. In the case  $k = N$ , this basis is then a Schur basis for the matrix  $J$ . Its explicit determination is not needed in the reduction process, only its existence is needed.

**Proof of Lemma 3.1.** The proof is done by induction on  $k$ . The case  $k = 1$  is obvious. For the sake of simplicity, we only do the induction from  $k = 1$  to  $k = 2$ . The general case is very similar but leads to very long calculations (see [4]).

Let us first prove part 1). Let  $b^0 \in \mathbb{C}^N$  and  $b^1 \in \mathbb{C}^{N-1}$  as in Lemma 3.1. Set  $b^{0,1} = b^0$  and  $b^{1,2} = b^1$ . From Remark 3.2, we can assume, up to normalization, that  $b_{i_1}^0 = 1 = \max |b_i^0|$  and  $b_{i_2}^1 = 1 = \max_{i \neq i_1} |b_i^1|$ . We look for a vector  $b^{0,2} \in \mathbb{C}^N$  such that (3.4-3.5) hold and  $(b^{0,1}|b^{0,2}) = 0$ . Formulæ (3.4-3.5) hold iff there exists  $\beta \in \mathbb{C}$  such that

$$\begin{cases} b_{i_1}^{0,2} &= \beta \\ b_i^{0,2} &= b_i^{1,2} + b_i^{0,1}\beta, \forall i \neq i_1. \end{cases} \quad (3.6)$$

Since  $b_{i_1}^{0,1} = 1$ , it follows that  $(b^{0,2}|b^{0,1}) = 0$  is true if and only if

$$\beta + \beta \sum_{i \neq i_1} b_i^{0,1} \overline{b_i^{0,1}} + \sum_{i \neq i_1} b_i^{1,2} \overline{b_i^{0,1}} = 0.$$

In other words, if  $\beta = -\|b^{0,1}\|^{-2} \sum_{i \neq i_1} b_i^{1,2} \overline{b_i^{0,1}}$  (all the quantities in the right hand side are known), then the family  $(b^{0,1}, b^{0,2})$  is orthogonal and satisfies (3.4-3.5). Notice that  $b^{0,2} \neq 0$ , otherwise  $b^{1,2}(=b^1)$  would also be 0 by (3.5).

To prove part 2), take a family  $(b^{0,1}, b^{0,2})$  satisfying (3.4-3.5). In particular,  $b^{0,1} = b^0$ , whence  $b^{0,1}J = \lambda_1 b^{0,1}$ . Furthermore, there exists  $\beta \in \mathbb{C}$  such that (3.6) is satisfied. We shall now prove that there exists  $\alpha_{2,1} \in \mathbb{C}$  such that  $b^{0,2}J = \lambda_2 b^{0,2} + \alpha_{2,1} b^{0,1}$ . Take first  $j \neq i_1$ . We have:

$$\begin{aligned} \sum_{1 \leq k \leq N} b_k^{0,2} J_{kj} &= \sum_{k \neq i_1} (b_k^{1,2} + \beta b_k^{0,1}) J_{kj} + \beta J_{i_1 j} \\ &= \sum_{k \neq i_1} b_k^{1,2} (J_{kj}^1 + b_j^{0,1} J_{ki_1}) + \beta \sum_{k \neq i_1} b_k^{0,1} J_{kj} + \beta J_{i_1 j} \quad (\text{from (2.5)}) \\ &= \sum_{k \neq i_1} b_k^{1,2} J_{kj}^1 + \left( \sum_{k \neq i_1} b_k^{1,2} J_{ki_1} \right) b_j^{0,1} + \beta \lambda_1 b_j^{0,1} \quad (\text{since } b^{0,1}J = \lambda_1 b^{0,1}) \\ &= \lambda_2 b_j^{1,2} + (\beta \lambda_1 + \sum_{k \neq i_1} b_k^{1,2} J_{ki_1}) b_j^{0,1} \quad (\text{since } b^{1,2}J^1 = \lambda_2 b^{1,2}) \\ &= \lambda_2 (b_j^{0,2} - \beta b_j^{0,1}) + (\beta \lambda_1 + \sum_{k \neq i_1} b_k^{1,2} J_{ki_1}) b_j^{0,1} \\ &= \lambda_2 b_j^{0,2} + [\beta(\lambda_1 - \lambda_2) + \sum_{k \neq i_1} b_k^{1,2} J_{ki_1}] b_j^{0,1}. \end{aligned}$$

On the other hand, for  $j = i_1$ , the right hand side of the last equality is equal to  $\beta \lambda_1 + \sum_{k \neq i_1} b_k^{1,2} J_{ki_1}$  since  $b_{i_1}^{0,1} = 1$ . The left hand side is equal to

$$\sum_{1 \leq k \leq N} b_k^{0,2} J_{ki_1} = \sum_{k \neq i_1} (b_k^{1,2} + \beta b_k^{0,1}) J_{ki_1} + \beta J_{i_1 i_1} = \beta \lambda_1 + \sum_{k \neq i_1} b_k^{1,2} J_{ki_1}$$

since  $b^{0,1}J = \lambda_1 J$  and  $b_{i_1}^{0,1} = 1$ . Hence,  $b^{0,2}J = \lambda_2 b^{0,2} + [\beta(\lambda_1 - \lambda_2) + \sum_{k \neq i_1} b_k^{1,2} J_{ki_1}] b^{0,1}$ .  $\square$

**Lemma 3.3** *There exists a constant  $C$ , which depends neither on  $t$  nor on  $J^k$ , such that*

$$\forall 0 \leq k \leq N, \forall t \geq 0, \quad \|e^{J^k t}\|_\infty = \sup_{x=(x_i)_{i \in \mathcal{K}_k} \in \mathbb{C}^{N-k}, \|x\|_\infty=1} \|e^{J^k t} x\|_\infty \leq C \quad (3.7)$$

where  $\|y\|_\infty = \max_{i \in \mathcal{K}_k} |y_i|$  for all  $y = (y_i)_{i \in \mathcal{K}_k} \in \mathbb{C}^{N-k}$ .

**Remark 3.4** From Theorem 2.1, the reduced matrices  $J^k$  are semi-stable but may not have a kinetic structure. As a consequence, it can be easily shown that the constant  $C$  in (3.7) does not depend on  $t$ . More precisely, since all the eigenvalues of  $J^k$  have non-negative real parts and the null eigenvalue of  $J^k$  has a complete subspace of eigenvectors, then for a Cauchy problem  $V'(t) = J^k V + F(t)$  with a given  $V(0) = V_0$ , Duhamel's principle yields that  $|V(t)| \leq C(|V_0| + \int_0^t |F(\tau)| d\tau)$  for all  $t \geq 0$ , where  $C$  is a positive constant depending only on  $J^k$ . We shall go one step further by proving that  $C$  does not depend on the coefficients of  $J^k$  either.

**Proof of Lemma 3.3.** Let us first prove (3.7) for  $k = 0$ . Referring to the kinetic structure of  $J$  ((2.3)-iii), the solution  $y(t)$  of the Cauchy problem (2.1) obeys  $\sum_{i=1}^N y_i'(t) = 0$ , whence

$$\forall t \geq 0, \quad \sum_{i=1}^N y_i(t) = \sum_{i=1}^N y_i(0). \quad (3.8)$$

Next, since, for each  $i \in \{1, \dots, N\}$ ,  $y_i'(t) = \sum_{j \neq i} J_{ij} y_j + J_{ii} y_i$  with  $J_{ij} \geq 0$  and  $J_{ii} \leq 0$ , the domain  $\mathbb{R}_+^N = \{y_i \geq 0, i = 1, \dots, N\}$  is invariant. Assuming  $y_i(0) \geq 0$  for all  $i = 1, \dots, N$ , it follows that  $y_i(t) \geq 0$  and  $0 \leq y_i(t) \leq \sum_{i=1}^N y_i(0)$  for all  $t \geq 0$  and for all  $i \in \{1, \dots, N\}$ , whence

$$\forall t \geq 0, \quad \|y(t)\|_\infty = \|e^{Jt} y(0)\|_\infty \leq C_1 \|y(0)\|_\infty,$$

where  $C_1 (= N)$  depends neither on the time  $t$  nor on the coefficients of the matrix  $J$ . Next, by writing any  $x \in \mathbb{C}^{N-k}$  as  $x = \Re(x)^+ - \Re(x)^- + i\Im(x)^+ - i\Im(x)^-$  where  $\Re(x)^+, \Re(x)^-, \Im(x)^+, \Im(x)^- \in \mathbb{R}_+^N$ , it follows that  $\|e^{Jt} x\|_\infty \leq 4C_1 \|x\|_\infty$ . This gives the estimate (3.7) for  $k = 0$ .

The arguments above no longer work for the matrices  $J^k$  ( $k \geq 1$ ) since these matrices may not be kinetic. Nevertheless, up to multiplication of the vectors  $b^k$  by some complex numbers, Lemma 3.1 and Remark 3.2 yield the existence of a unitary basis  $(b^{0,1}, \dots, b^{0,N})$  of  $\mathbb{C}^N$  satisfying (3.4-3.5) and such that  $J$  can be written as  $J = P^0 T^0 (P^0)^{-1}$  where  $P^0$  and

$$(P^0)^{-1} = \begin{pmatrix} - & b^{0,1} & - \\ & \vdots & \\ - & b^{0,N} & - \end{pmatrix}$$

are unitary and where  $T^0$  is lower triangular. In particular,  $T^0 = (P_0)^{-1} J P^0$  and  $\|e^{T^0 t}\|_\infty \leq \|(P_0)^{-1}\|_\infty \|e^{Jt}\|_\infty \|P^0\|_\infty \leq C_2^2 \|e^{Jt}\|_\infty$  where  $C_2 = \sqrt{N}$ . Since  $\|e^{Jt}\|_\infty \leq 4C_1$ , we have  $\|e^{T^0 t}\|_\infty \leq C_3 = 4C_1 C_2^2$  for all  $t \geq 0$ .

On the other hand, since  $b^{0,1} = b^0$  is the first vector of the basis  $(b^{0,1}, \dots, b^{0,N})$  and since the formulæ (3.3) and (3.5) are identical, the calculations done in the course of the proof of Theorem 2.1 imply that  $J^1 = P^1 T^1 (P^1)^{-1}$  where

$$(P^1)^{-1} = \begin{pmatrix} - & b^{1,2} & - \\ & \vdots & \\ - & b^{1,N} & - \end{pmatrix}, \quad b^{1,2} = b^1,$$

where  $P^1$  is a submatrix of  $P^0$  and  $T^1$  is the submatrix of  $T^0$  obtained by dropping the first row and the first column. In particular,  $\|P^1\|_\infty \leq \|P^0\|_\infty \leq C_2$  and, for all  $t \geq 0$ ,

$$\|e^{T^1 t}\|_\infty = \sup_{x \in \mathbb{C}^{N-1}, \|x\|_\infty=1} \|e^{T^1 t} x\|_\infty = \sup_{y=(0, x_2, \dots, x_N) \in \mathbb{C}^N, \|y\|_\infty=1} \|e^{T^0 t} y\|_\infty \leq \|e^{T^0 t}\|_\infty \leq C_3.$$



Moreover, from the formulæ (3.3) and the definition of  $i_1$ , it is found that  $\|(P^1)^{-1}\|_\infty \leq 2\|(P^0)^{-1}\|_\infty \leq 2C_2$ . Eventually  $\|e^{J^1 t}\|_\infty \leq \|P^1\|_\infty \|e^{T^1 t}\|_\infty \|(P^1)^{-1}\|_\infty \leq 2C_2^2 C_3$  for all  $t \geq 0$ .

An immediate induction completes the proof of Lemma 3.3.  $\square$

We now investigate the accuracy with which the solution  $y^R(t)$  of the reduced system ( $S^R$ ) approximates the solution  $y(t)$  of the initial system ( $S$ ). Let us first deal with the case  $R = 1$ .

**Lemma 3.5** *Let  $y^1(t)$  be the solution of the algebraic-differential system ( $S^1$ ) with the notations of Theorem 2.3. There exists a constant  $C$  that does not depend either on  $t$  or on the matrix  $J$ , such that, for any  $t_0 \geq \max(0, |\Re(\lambda_1)|^{-1} \ln |\Re(\lambda_1)|)$  and for any  $h \geq 0$ ,*

$$\begin{aligned} (\forall i \in \mathcal{K}_1 \ (i \neq i_1), \ |y_i^1(t_0) - y_i(t_0)| \leq h) \\ \implies (\forall 1 \leq i \leq N, \ \forall t \geq t_0, \ |y_i^1(t) - y_i(t)| \leq C(h + |\Re(\lambda_1)|^{-1})). \end{aligned} \quad (3.9)$$

**Proof.** The proof is based on the existence of an algebraic-differential system which is equivalent to ( $S$ ) and which involves both the reduced matrix  $J^1$  and the algebraic part of ( $S^1$ ).

First of all, observe that the characteristic variable  $z^0(t) = b^0 \cdot y(t)$  solves  $(z^0)'(t) = \lambda_1 z^0$  since  $b^0 = b^{0,1}$  is a left eigenvector of  $J$  for the eigenvalue  $\lambda_1$ . As already underlined, one can assume, up to normalization, that  $b_{i_1}^{0,1} = b_{i_1}^{0,1} = 1$ . Then,  $z^0(t) = z^0(0)e^{\lambda_1 t}$  and  $y_{i_1}(t) = -\sum_{i \neq i_1} b_i^{0,1} y_i(t) + z^0(0)e^{\lambda_1 t}$ . Next, each  $y_i(t)$  for  $i \in \mathcal{K}_1 = \{1, \dots, N\} \setminus \{i_1\}$  satisfies

$$\left\{ \begin{aligned} \frac{d}{dt} y_i &= \sum_{j,k,l=1}^N a_i^{0,j} T_{jk}^0 b_l^{0,k} y_l \\ &= \sum_{j,k=1}^N \sum_{l \in \mathcal{K}_1} a_i^{0,j} T_{jk}^0 b_l^{0,k} y_l + \sum_{j,k=1}^N a_i^{0,j} T_{jk}^0 b_{i_1}^{0,k} y_{i_1} \\ &= \sum_{j,k=1}^N \sum_{l \in \mathcal{K}_1} a_i^{0,j} T_{jk}^0 (b_l^{0,k} - b_l^{0,1} b_{i_1}^{0,k}) y_l + z^0(0) e^{\lambda_1 t} J_{ii_1}^0 \\ &= \sum_{l \in \mathcal{K}_1} J_{il}^1 y_l + z^0(0) e^{\lambda_1 t} J_{ii_1}^0 \quad \text{from (3.1)}. \end{aligned} \right.$$

Therefore, system ( $S$ ) is equivalent to the following system ( $S_e$ ):

$$(S_e) \quad \left\{ \begin{aligned} \frac{d}{dt} y &= J^1 y + z^0(0) e^{\lambda_1 t} [J^0] \quad \text{for } (y_i)_{i \neq i_1}, \quad y_i(0) \text{ given for } i \neq i_1 \\ y_{i_1}(t) &= -\sum_{i \in \mathcal{K}_1} b_i^{0,1} y_i(t) + z^0(0) e^{\lambda_1 t} \end{aligned} \right.$$

where the vector  $[J^0] \in \mathbb{R}^{N-1}$  is given componentwise by  $[J^0]_i = J_{ii_1}^0$  for  $i \neq i_1$ .

Let us now compare  $y^1(t)$  with  $y(t)$  and derive the error estimate (3.9) from a time  $t_0 \geq 0$  that shall be explicited later. For  $i = 1, \dots, N$ , denote  $\underline{e}_i^1(t) = y_i(t) - y_i^1(t)$ . Thanks to the equivalent system ( $S_e$ ), the vector  $e^1(t) := (\underline{e}_i^1(t))_{i \neq i_1}$  solves the Cauchy problem

$$\left\{ \begin{aligned} \frac{d}{dt} e^1 &= J^1 e^1 + z^0(0) e^{\lambda_1 t} [J^0] \\ \underline{e}_i^1(t_0) &= y_i(t_0) - y_i^1(t_0), \quad i \neq i_1. \end{aligned} \right. \quad (3.10)$$

Thus,  $e^1(t) = e^{J^1(t-t_0)} e^1(t_0) + \int_{t_0}^t z^0(0) e^{\lambda_1 \tau} e^{J^1(t-\tau)} [J^0] d\tau$  for all  $t \geq t_0$ . Assuming  $\|e^1(t_0)\|_\infty \leq h$ , it follows that

$$\begin{aligned} \forall t \geq t_0, \quad \|e^1(t)\|_\infty &\leq \|e^{J^1(t-t_0)}\|_\infty h + |z^0(0)| \int_{t_0}^t e^{\Re(\lambda_1)\tau} \|e^{J^1(t-\tau)}\|_\infty \|[J^0]\|_\infty d\tau \\ &\leq Ch + C \|[J^0]\|_\infty |z^0(0)| \int_{t_0}^t e^{\Re(\lambda_1)\tau} d\tau \end{aligned}$$

where the constant  $C$ , which is independent from  $t$ ,  $J$  and  $J^1$ , is given by Lemma 3.3. Remember now that the chemical species are assumed to be chosen at  $t = 0$  so that  $\sum_{i=1}^N y_i(0) = 1$  with  $y_i(0) \geq 0$  for all  $i$ . Hence,  $|z^0(0)| \leq \sum_{i=1}^N |b_i^{0,1}| y_i(0) \leq 1$  from the definition of  $i_1$ . On the other hand, it follows from the kinetic structure of the matrix  $J^0 = J$  (see (2.3)) that  $|J_{ij}^0| \leq |J_{jj}^0| \leq |Tr J^0| \leq N |\Re(\lambda_1)|$  for all  $i, j$ . Then, even if it means changing the constant  $C$ , one has

$$\forall t \geq t_0, \quad \|e^1(t)\|_\infty \leq Ch + C |\Re(\lambda_1)| \int_{t_0}^t e^{\Re(\lambda_1)\tau} d\tau = Ch + C e^{\Re(\lambda_1)t_0} |e^{\Re(\lambda_1)(t-t_0)} - 1|. \quad (3.11)$$

Taking any time  $t_0$  such that  $t_0 \geq \max(0, |\Re(\lambda_1)|^{-1} \ln |\Re(\lambda_1)|)$  eventually leads to

$$\forall t \geq t_0, \quad \|e^1(t)\|_\infty \leq C (h + |\Re(\lambda_1)|^{-1}). \quad (3.12)$$

To conclude the proof, we shall estimate  $\underline{e}_{i_1}^1(t)$  for  $t \geq t_0$ . We have  $\underline{e}_{i_1}^1(t) = y_{i_1}(t) - y_{i_1}^1(t) = -\sum_{i \neq i_1} b_i^{0,1} \underline{e}_i^1(t) + z^0(0) e^{\lambda_1 t}$ , whence

$$\forall t \geq t_0, \quad |\underline{e}_{i_1}^1(t)| \leq \sum_{i \neq i_1} |b_i^{0,1}| |\underline{e}_i^1(t)| + |z^0(0)| e^{\Re(\lambda_1)t} \leq \sum_{i \neq i_1} |\underline{e}_i^1(t)| + e^{\Re(\lambda_1)t}.$$

As a consequence, from (3.12) and from the choice of  $t_0$ , it is found that, for some constant  $C$ :

$$\forall t \geq t_0, \quad |\underline{e}_{i_1}^1(t)| \leq (N-1) \|e^1(t)\|_\infty + |\Re(\lambda_1)|^{-1} \leq C(h + |\Re(\lambda_1)|^{-1}). \quad \square$$

**Proof of Theorem 2.3, part ii).** Let us now turn to establish the similar error estimate analysis as above after  $R$  eliminations with  $R > 1$ . We prove (2.10) by induction on  $R$ .

Let us assume that (2.10) is true at the step  $R-1$ . In other words, there exists a constant  $C_1$  such that, for any  $t'_0 \geq \max(0, \max_{1 \leq k \leq R-1} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|)$  and for any  $h \geq 0$ , one has

$$\begin{aligned} & (\forall i \neq i_1, \dots, i_{R-1}, \quad |y_i^{R-1}(t'_0) - y_i(t'_0)| \leq h) \\ & \implies (\forall t \geq t'_0, \quad \forall i \in \{1, \dots, N\}, \quad |y_i^{R-1}(t) - y_i(t)| \leq C_1(h + |\Re(\lambda_{R-1})|^{-1})). \end{aligned} \quad (3.13)$$

where  $(y^{R-1}) = (y_i^{R-1})_{1 \leq i \leq N}$  solves the Cauchy problem  $(S^{R-1})$ , *i.e.* (2.9), from time  $t'_0$ .

Take any initial time  $t_0 \geq \max(0, \max_{1 \leq k \leq R} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|)$  and suppose that  $|y_i^R(t_0) - y_i(t_0)| \leq h$  for all  $i \neq i_1, \dots, i_R$ , where  $(y^R) = (y_i^R)_{1 \leq i \leq N}$  solves the Cauchy problem  $(S^R)$  from time  $t_0$ . We shall prove (2.10). As already underlined, one can assume, up to normalization, that  $b_{i_k}^{k-1} = b_{i_k}^{k-1,k} = 1$  for each  $k = 1, \dots, R$ . Observe first that the vector  $(y_i^R)_{i \neq i_1, \dots, i_{R-1}}$  solves

$$\begin{cases} \frac{d}{dt}(y_i^R) = J^R(y_i^R), & y_i^R(t_0) \text{ given,} & \text{for } i \neq i_1, \dots, i_R \\ y_{i_R}^R = -\sum_{i \neq i_1, \dots, i_R} b_i^{R-1,R} y_i^R \end{cases}$$

whereas the  $R-1$  other components are given by  $y_{i_k}^R = -\sum_{i \neq i_1, \dots, i_k} b_i^{k-1,k} y_i^R$  for  $k = 1, \dots, R-1$ . Consider now the solution  $(y_i^{R-1})_{1 \leq i \leq N}$  of the following Cauchy problem

$$\begin{cases} \frac{d}{dt}(y_i^{R-1}) = J^{R-1}(y_i^{R-1}), & y_i^{R-1}(t_0) = y_i(t_0), & i \neq i_1, \dots, i_{R-1} \\ y_{i_k}^{R-1} = -\sum_{i \neq i_1, \dots, i_k} b_i^{k-1,k} y_i^{R-1} & \text{for } k = 1, \dots, R-1 \end{cases} \quad (3.14)$$

The results of Lemma 3.5 can then be applied to the vectors  $(y_i^R)_{i \neq i_1, \dots, i_{R-1}}$  and  $(y_i^{R-1})_{i \neq i_1, \dots, i_{R-1}}$  of size  $N-R+1$ . Namely, for any  $i \neq i_1, \dots, i_R$ , we have  $|y_i^R(t_0) - y_i^{R-1}(t_0)| = |y_i^R(t_0) - y_i(t_0)| \leq h$  with  $t_0 \geq \max(0, \max_{1 \leq k \leq R} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|) \geq \max(0, |\Re(\lambda_R)|^{-1} \ln |\Re(\lambda_R)|)$ . Then

$$\forall t \geq t_0, \quad \forall i \neq i_1, \dots, i_{R-1}, \quad |y_i^R(t) - y_i^{R-1}(t)| \leq C_2(h + |\Re(\lambda_R)|^{-1}) \quad (3.15)$$

for some constant  $C_2$  which depends neither on  $t$ ,  $h$  nor on the coefficients of  $J^{R-1}$ .

Let us now find an upper bound for  $|y_i^R(t) - y_i^{R-1}(t)|$ , for  $i = i_1, \dots, i_{R-1}$ . We have

$$\begin{aligned} \forall t \geq t_0, \quad |y_{i_{R-1}}^R(t) - y_{i_{R-1}}^{R-1}(t)| &= |\sum_{i \neq i_1, \dots, i_{R-1}} b_i^{R-2, R-1} (y_i^R(t) - y_i^{R-1}(t))| \\ &\leq \sum_{i \neq i_1, \dots, i_{R-1}} |y_i^R(t) - y_i^{R-1}(t)| \\ &\leq (N - R + 1) C_2(h + |\Re(\lambda_R)|^{-1}) \quad \text{from (3.15).} \end{aligned}$$

We can repeat this calculation for  $|y_{i_{R-2}}^R(t) - y_{i_{R-2}}^{R-1}(t)|, \dots, |y_{i_1}^R(t) - y_{i_1}^{R-1}(t)|$ . An immediate induction eventually yields the existence of another constant  $C_3$  such that

$$\forall t \geq t_0, \quad \forall i \in \{1, \dots, N\}, \quad |y_i^R(t) - y_i^{R-1}(t)| \leq C_3(h + |\Re(\lambda_R)|^{-1}) \quad (3.16)$$

On the other hand, since the vector  $(y_i^{R-1})_{1 \leq i \leq N}$  solves (3.14) and since  $t_0 \geq \max(0, \max_{1 \leq k \leq R} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|) \geq \max(0, \max_{1 \leq k \leq R-1} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|)$ , (3.13) yields

$$\forall t \geq t_0, \quad \forall i \in \{1, \dots, N\}, \quad |y_i^{R-1}(t) - y_i(t)| \leq C_1 |\Re(\lambda_{R-1})|^{-1}. \quad (3.17)$$

As a conclusion, (3.16) and (3.17) imply that, for each  $t \geq t_0$  and each  $i$ ,

$$|y_i^R(t) - y_i(t)| \leq C_3 h + C_3 |\Re(\lambda_R)|^{-1} + C_1 |\Re(\lambda_{R-1})|^{-1} \leq (C_1 + C_3)(h + |\Re(\lambda_R)|^{-1})$$

since  $|\Re(\lambda_{R-1})| \geq |\Re(\lambda_R)|$ . That is the required result. The proof of Theorem 2.3 is complete.  $\square$

## 4 Reduction of a kinetic systems with multiple scales

This section is devoted to the application of the reduction process defined in section 2.2 to a kinetic system of the type (2.11). In (2.11), the matrix  $J$  is written as a finite sum of kinetic matrices  $\varepsilon^{-p_s} A^s$  with  $p_1 > \dots > p_S$ . It is then reasonable to think that the matrix  $J$  has some eigenvalues in the scale  $\varepsilon^{-p_1}$ , namely those of  $\varepsilon^{-p_1} A^1$  approximatively.

Since the columns  $C_i$  of  $A^1$  are equal to 0 for  $i > I_1$ ,  $A^1$  can be written as  $A^1 = \begin{pmatrix} B & 0 \\ C & 0 \end{pmatrix}$  where  $B$  is a square matrix of size  $I_1 \times I_1$  (remember that  $I_1$  is the number of species which are reactants of at least one of the fastest reactions). The eigenvalues  $\mu_i$  of the kinetic matrix  $A^1$  can be ordered so that  $\Re(\mu_1) \leq \dots \leq \Re(\mu_{I_1}) < \mu_{I_1+1} = \dots = \mu_N = 0$  where the rank  $I_1'$  of  $A^1$  satisfies  $I_1' \leq I_1$ . For each  $i \leq I_1'$ , let  $m_i$  be the multiplicity of the eigenvalue  $\varepsilon^{-p_1} \mu_i$  of  $\varepsilon^{-p_1} A^1$ .

**Lemma 4.1** *There exist two constants  $C_0$  and  $\varepsilon_0 > 0$  such that if  $0 < \varepsilon \leq \varepsilon_0$ , then the eigenvalues  $\lambda_1, \dots, \lambda_N$  of the matrix  $J$  can be ordered so that (2.4) holds and*

$$\left\{ \begin{array}{ll} \forall 1 \leq i \leq I_1', & |\lambda_i - \varepsilon^{-p_1} \mu_i| \leq C_0 \varepsilon^{-p_1 + (p_1 - p_2)/m_i} \\ \forall I_1' < i \leq I_1 + \dots + I_S, & |\lambda_i| \leq C_0 \varepsilon^{-p_2 - \beta(p_1 - p_2)} \\ \forall I_1 + \dots + I_S < i \leq N, & \lambda_i = 0, \end{array} \right. \quad (4.1)$$

where  $\beta = (I_1 - I_1')/(I_1 - I_1' + 1)$ . Moreover,  $|\lambda_i| \leq 2C_2 R \varepsilon^{-p_1}$  for each  $1 \leq i \leq N$ .

For  $\varepsilon$  small enough, the system  $\dot{y} = Jy$  contains then exactly  $I'_1(\leq I_1)$  dynamics in the scale  $\varepsilon^{-p_1}$ . Moreover, all the eigenvalues  $\lambda_i$  for  $i \geq I'_1 + 1$  belong to lower scales than the first  $I'_1$  ones and if  $I_1 = I'_1$ , then those eigenvalues range in scales  $\mathcal{O}(\varepsilon^{-p_2})$ . However, the determination of all the dynamics of  $J$  in the scale  $\varepsilon^{-p_2}$  or in smaller scales is not clear in general.

The proof of Lemma 4.1 can be done by comparing the characteristic polynomials of  $J$  and  $\varepsilon^{-p_1}A^1$  (see [4] for the technical details) and it especially uses the bounds:

$$\forall 1 \leq s \leq S, \forall i \in \mathcal{I}_s, \begin{cases} 0 \leq J_{ij} \leq C_2 \varepsilon^{-p_s} & \text{if } j \neq i \\ -(R_s + \dots + R_S)C_2 \varepsilon^{-p_s} \leq J_{ii} \leq -C_1 \varepsilon^{-p_s}. \end{cases} \quad (4.2)$$

**Remark 4.2** Lemma 4.1 can be seen as an application of Puiseux series expansions for a special type of a matrix  $J$  whose highest part  $\varepsilon^{-p_1}A^1$  has only a finite number of nonzero columns (see [10] for similar general results).

Let us now turn to the

**Proof of Theorem 2.4, part i).** First, remember that the eigenvalues  $\lambda_i$  of  $J$  are ordered as in (2.4) and that the final reduction step  $R$  is the first step such that  $\Im(\lambda_R) \leq 0$  and  $|\Re(\text{Tr}(J^R))| \leq \varepsilon^{-\alpha}$  (where  $p_q > \alpha > p_{q+1}$ ). From Lemma 4.1, there exists  $C > 0$  such that

$$\forall 1 \leq i \leq I'_1, \quad |\lambda_i| \geq |\Re(\lambda_i)| \geq C \varepsilon^{-p_1} \quad (\text{for } \varepsilon \text{ small enough}). \quad (4.3)$$

On the other hand, we saw in section 2.3 that  $|\Re(\lambda_i)| \leq \varepsilon^{-\alpha}$  for  $i \geq R + 1$ . Since  $p_1 > \alpha$ , one has  $R \geq I'_1$  for  $\varepsilon$  small enough.

From now on, we set  $p_{S+1} = -\infty$  and  $\varepsilon^{+\infty} = 0$ . Consider now the first step in the reduction process. The vector  $b^0 \in \mathbb{C}^N$  satisfies  $b^0 J = \lambda_1 b^0$  and  $i_1$  is the smallest integer such that  $|b_{i_1}^0| = \max_{1 \leq i \leq N} |b_i^0|$ . Let  $2 \leq s \leq S + 1$  and  $j \in \mathcal{I}_s$  and let us prove that  $|b_j^0| < |b_{i_1}^0|$  for  $\varepsilon$  small enough, which yields  $i_1 \leq I_1$ . Since  $\lambda_1 b^0 = b^0 J$ , one has  $b_j^0 = \lambda_1^{-1} \sum_{i=1}^N b_i^0 J_{ij}$ . Hence,  $|b_j^0| \leq |b_{i_1}^0| |\lambda_1|^{-1} \sum_{i=1}^N |J_{ij}|$ . By (2.3) and (4.2) applied to  $j \in \mathcal{I}_s$ , we have  $\sum_{i=1}^N |J_{ij}| = 2|J_{jj}| \leq 2C_2 M \varepsilon^{-p_s}$  ( $M$  is the number of chemical reactions). Combining that with (4.3) gives

$$|b_j^0| \leq 2C_2 M C^{-1} \varepsilon^{p_1 - p_s} |b_{i_1}^0|.$$

Since  $p_1 > p_2 \geq p_s$ , this yields  $|b_j^0| < |b_{i_1}^0|$  for  $\varepsilon$  small enough. Eventually,  $1 \leq i_1 \leq I_1$ .

Take now any  $i, j \neq i_1$ . From (2.3) and (4.2) and owing to the definition (2.5) of  $J^1$ , we get:

- if  $j \in \mathcal{I}_1$  and if  $i \neq j$ , then  $|J_{ij}^1| = |J_{ij} - (b_j^0)/(b_{i_1}^0)J_{ii_1}| \leq 2C_2 \varepsilon^{-p_1}$ ,
- if  $j \in \mathcal{I}_1$  and if  $i = j$ , then  $|J_{ii}^1| \leq (M + 1)C_2 \varepsilon^{-p_1}$ ,
- if  $j \in \mathcal{I}_s$  with  $2 \leq s \leq S + 1$  and if  $i \neq j$ , then  $|J_{ij}^1| \leq C_2 \varepsilon^{-p_s} + (2C_2 M C^{-1} \varepsilon^{p_1 - p_s})(C_2 \varepsilon^{-p_1}) \leq C_2(1 + 2C_2 M C^{-1}) \varepsilon^{-p_s}$ ,
- if  $j \in \mathcal{I}_s$  with  $2 \leq s \leq S + 1$  and if  $i = j$ , then  $|J_{ii}^1| \leq C_2 M(1 + 2C_2 C^{-1}) \varepsilon^{-p_s}$ .

In any case, there exists a constant  $C$  such that for any  $i, j \neq i_1$  and  $j \in \mathcal{I}_s$  with  $1 \leq s \leq S + 1$ , then  $|J_{ij}^1| \leq C \varepsilon^{-p_s}$ .

At the second reduction step, one has  $b^1 J^1 = \lambda_2 b^1$  and  $i_2$  is the smallest integer in  $\{1, \dots, N\} \setminus \{i_1\}$  is such that  $|b_{i_2}^1| = \max_{i \neq i_1} |b_i^1|$ . From (4.3), we have  $|\lambda_2| \geq C \varepsilon^{-p_1}$  for  $\varepsilon$  small enough (provided that  $I'_1 \geq 2$ ). Since the matrix  $J^1$  satisfies the same kind of bounds (4.2) as  $J$ , the same arguments as above yield that, for  $\varepsilon$  small enough,  $i_2 \leq I_1$  and that the new reduced matrix  $J^2$  satisfies  $|J_{ij}^2| \leq C \varepsilon^{-p_s}$  for any  $s$  and for any  $i, j \neq i_1, i_2$  with  $j \in \mathcal{I}_s$ .

This result holds by induction at any  $k^{\text{th}}$  step in the reduction process, provided that  $k \leq I'_1$ . **Proof of Theorem 2.4, part ii).** Assume now that  $p_1 > \alpha > p_2 + \beta(p_1 - p_2)$ . From (2.12), we observed in section 2.3 that  $|\Re(\lambda_i)| \geq \frac{1}{N-R-1}\varepsilon^{-\alpha}$  for each  $1 \leq i \leq R$ . On the other hand, Lemma 4.1 states that  $|\lambda_i| \leq C_0\varepsilon^{-p_2-\beta(p_1-p_2)}$  for each  $i > I'_1$ . This finally implies that  $R \leq I'_1$  for  $\varepsilon$  small enough, whence  $R = I'_1$  from the part i) of Theorem 2.4.

The formulæ (2.13) for the final reduced matrix  $J^R = J^{I'_1}$  are a particular case of the bounds given in the previous paragraphs for the coefficients of the reduced matrices  $J^k$  for  $k \leq I'_1$ .

Let us now prove that  $J^r = o(\varepsilon^{-p_1})$  as  $\varepsilon \rightarrow 0$ . Take a sequence  $\varepsilon \rightarrow 0$ . For  $\varepsilon$  small enough, the conclusion of Lemma 4.1 as well as the first part of Theorem 2.4 are valid. We have  $i_1(\varepsilon) \in \{1, \dots, I_1\}$ . Up to extraction of some subsequence, we can then assume that  $i_1(\varepsilon) = i_1$  does not depend on  $\varepsilon$ , with  $b_{i_1}^0(\varepsilon) = 1$  (up to normalization), and that  $(b^0(\varepsilon))_\varepsilon$  converges to some vector  $b^0$  such that  $b_{i_1}^0 = 1 = \max |b_i^0|$ . The vectors  $b^0(\varepsilon)$  satisfy

$$b^0(\varepsilon)J(\varepsilon) = \varepsilon^{-p_1}b^0(\varepsilon)A^1 + \dots + \varepsilon^{-p_S}b^0(\varepsilon)A^S = \lambda_1(\varepsilon)b^0(\varepsilon).$$

Multiplying this equation by  $\varepsilon^{p_1}$  and passing to the limit  $\varepsilon \rightarrow 0$  leads to  $b^0A^1 = \mu_1b^0$  since  $\varepsilon^{p_1}\lambda_1(\varepsilon) \rightarrow \mu_1$  as  $\varepsilon \rightarrow 0$  by Lemma 4.1.

As done for  $J^1$  in section 2.2, define the matrices  $A^{s,1}(\varepsilon)$  and  $A^{s,1}$  for any  $1 \leq s \leq S$  by  $A_{ij}^{s,1}(\varepsilon) = A_{ij}^s - b_j^0(\varepsilon)A_{ii_1}^s$  and  $A_{ij}^{s,1} = A_{ij}^s - b_j^0A_{ii_1}^s$  for any  $i, j \neq i_1$  (remember that  $b_{i_1}^0 = 1$ ). We see that  $A^{s,1}(\varepsilon) \rightarrow A^{s,1}$  as  $\varepsilon \rightarrow 0$ . From (2.11) and the definition of the matrix  $J^1(\varepsilon)$ , we have

$$J^1(\varepsilon) = \varepsilon^{-p_1}A^{1,1}(\varepsilon) + \dots + \varepsilon^{-p_S}A^{S,1}(\varepsilon).$$

Since the matrices  $A^{s,1}(\varepsilon)$  are uniformly bounded as  $\varepsilon \rightarrow 0$ , one gets  $\varepsilon^{p_1}J^1(\varepsilon) \rightarrow A^{1,1}$  as  $\varepsilon \rightarrow 0$ .

By induction, and up to normalization and extraction of some subsequences, we can similarly define  $i_2, \dots, i_{I'_1}$  and  $b^1, \dots, b^{I'_1-1}$  such that  $b_{i_k}^{k-1}(\varepsilon) = 1 = \max_{i \neq i_1, \dots, i_{k-1}} |b_i^{k-1}(\varepsilon)|$  and  $b^{k-1}(\varepsilon) \rightarrow b^{k-1}$  as  $\varepsilon \rightarrow 0$  for each  $2 \leq k \leq I'_1$ . Set  $A_{ij}^{1,k} = A_{ij}^{1,k-1} - b_j^{k-1}A_{ii_k}^{1,k-1}$  for any  $i, j \neq 1, \dots, i_k$ . With the same arguments as above, it follows that  $\varepsilon^{p_1}J^k(\varepsilon) \rightarrow A^{1,k}$  as  $\varepsilon \rightarrow 0$  for each  $1 \leq k \leq I'_1$  and that  $b^{k-1} \neq 0$  satisfies  $b^{k-1}A^{1,k-1} = \mu_k b^{k-1}$ . Eventually, the matrices  $A^{1,1}, \dots, A^{1,I'_1}$  are the reduced matrices (in the sense of section 2.2) of the matrix  $A^1$ . At the step  $I'_1$ , all the nonzero eigenvalues of the kinetic matrix  $A^1$  have been eliminated. From Theorem 2.1, the matrix  $A^{1,I'_1}$  is still semi-stable and therefore it is the 0 matrix. This means that  $\varepsilon^{p_1}J^{I'_1}(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ , which completes the proof of Theorem 2.4.  $\square$

**Proof of Corollary 2.6.** From Theorem 2.3, there exists  $C_0 > 0$  such that for any  $t_0 \geq \max(0, \max_{1 \leq k \leq R} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|)$ , if the initial condition for  $(S^R)$  satisfies  $|y_i^R(t_0) - y_i(t_0)| \leq \varepsilon^{p_1}$  for all  $i \neq i_1, \dots, i_R$ , then  $|y_i^R(t) - y_i(t)| \leq C_0(\varepsilon^{p_1} + |\Re(\lambda_R)|^{-1})$  for all  $1 \leq i \leq N$  and  $t \geq t_0$ .

Under the assumptions of Corollary 2.6, Theorem 2.4 gives  $R = I'_1$  for  $\varepsilon > 0$  small enough. Take  $\delta > |\Re(\mu_{I'_1})|^{-1}$ . From Lemma 4.1,  $\Re(\lambda_R) = \Re(\lambda_{I'_1}) \sim \varepsilon^{-p_1}\Re(\mu_{I'_1})$  as  $\varepsilon \rightarrow 0$ . Moreover,

$$\max(0, \max_{1 \leq k \leq R=I'_1} |\Re(\lambda_k)|^{-1} \ln |\Re(\lambda_k)|) \sim |\Re(\mu_{I'_1})|^{-1}\varepsilon^{p_1} \ln \varepsilon^{-p_1} < \delta\varepsilon^{p_1} |\ln \varepsilon^{p_1}| \quad \text{as } \varepsilon \rightarrow 0.$$

Finally, by setting  $C = C_0(1 + 2/|\Re(\mu_{I'_1})|)$ , we conclude that if  $\varepsilon > 0$  is small enough and if  $t_0 \geq \delta\varepsilon^{p_1} |\ln \varepsilon^{p_1}|$ , then one has

$$\forall t \geq t_0, \quad \forall i \in \{1, \dots, N\}, \quad |y_i^R(t) - y_i(t)| \leq C\varepsilon^{p_1}$$

if  $|y_i^R(t_0) - y_i(t_0)| \leq \varepsilon^{p_1}$  for each  $i \neq i_1, \dots, i_R$ . That completes the proof of Corollary 2.6.  $\square$

**Remark 4.3** (About the choice of the indices  $i_k$ ) The fact that  $|b_{i_1}^0|$  is the highest component of  $b^0$  really matters. We see in the following counterexample what happens if  $i_1$  is ill-chosen. Consider the reaction scheme  $\mathcal{A}_4 \xrightarrow{-2} \mathcal{A}_1 \xrightarrow{\varepsilon^{-1}} \mathcal{A}_2 \xrightarrow{1} \mathcal{A}_3$ ,  $\mathcal{A}_1 \xrightarrow{1} \mathcal{A}_3$ . The vector  $(y_1, \dots, y_4)^T$  solves

$$\frac{d}{dt} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} -1 - \varepsilon^{-1} & 0 & 0 & 2 \\ \varepsilon^{-1} & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = J \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

The eigenvalues of  $J$  are  $-\varepsilon^{-1} - 1$ ,  $-2$ ,  $-1$  and  $0$ . A left eigenvector  $b^0$  for the eigenvalue  $\lambda_1 = -\varepsilon^{-1} - 1$  is  $b^0 = (1, 0, 0, -2\varepsilon/(1 - \varepsilon))$ . For  $\varepsilon$  small enough,  $i_1 = 1$  (see Theorem 2.4) and  $y_1$  is the only eliminated species. The reduced system is

$$\frac{d}{dt} \begin{pmatrix} y_2^1 \\ y_3^1 \\ y_4^1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 2/(1 - \varepsilon) \\ 1 & 0 & 2\varepsilon/(1 - \varepsilon) \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} y_2^1 \\ y_3^1 \\ y_4^1 \end{pmatrix} = J^1 \begin{pmatrix} y_2^1 \\ y_3^1 \\ y_4^1 \end{pmatrix}, \quad y_1^1 = \frac{2\varepsilon}{1 - \varepsilon} y_4^1,$$

and one has  $\varepsilon J^1 \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . If we choose to eliminate  $y^4$ , we get the reduced system

$$\frac{d}{dt} \begin{pmatrix} \tilde{y}_1^1 \\ \tilde{y}_2^1 \\ \tilde{y}_3^1 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 0 \\ \varepsilon^{-1} & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{y}_1^1 \\ \tilde{y}_2^1 \\ \tilde{y}_3^1 \end{pmatrix}, \quad \tilde{y}_4^1 = \frac{1 - \varepsilon}{2\varepsilon} \tilde{y}_1^1.$$

This system contains some terms in the scale  $\varepsilon^{-1}$ . This is awkward in the numerical calculations.

**Remark 4.4** Why not reduce  $\varepsilon^{-p_1} A^1$ ? Since  $J = \varepsilon^{-p_1} A^1 + \dots + \varepsilon^{-p_s} A^s$  and the coefficients of the matrices  $\varepsilon^{-p_s} A^s$  ( $s \geq 2$ ) are small compared to the norm of  $\varepsilon^{-p_1} A^1$ , another reduction process could be the following: 1) find an eigenvector  $\underline{b}^0$  of the matrix  $\varepsilon^{-p_1} A^1$  for the eigenvalue  $\varepsilon^{-p_1} \mu_1$ , 2) define  $\underline{i}_1$  to be the smallest integer such that  $|b_{\underline{i}_1}^0| = \max |b_i^0|$ , 3) build the reduced matrix  $\underline{J}^1$  in terms of  $J^0$  and  $\underline{b}^0$  with the same formulæ as in section 2.2, 4) repeat this procedure  $I'_1$  times and eliminate all nonzero eigenvalues  $\varepsilon^{-p_1} \mu_1, \dots, \varepsilon^{-p_1} \mu_{I'_1}$  of  $\varepsilon^{-p_1} A^1$ . However, this method may introduce some errors in the coefficients of the reduced matrices.

Indeed, consider the matrix

$$J = \begin{pmatrix} -\varepsilon^{-1} - 1 & \varepsilon^{-1} & 1 \\ \varepsilon^{-1} & -\varepsilon^{-1} - 1 & 0 \\ 1 & 1 & -1 \end{pmatrix} = \varepsilon^{-1} \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 1 & -1 \end{pmatrix} = \varepsilon^{-1} A^1 + A^2.$$

The eigenvalues of  $J$  are  $-2\varepsilon^{-1} - 1$ ,  $-2$  and  $0$ . A left eigenvector for the eigenvalue  $-2\varepsilon^{-1} - 1$  is  $b^0 = (1, -1 + \varepsilon^2/2, -\varepsilon/2)$ , and  $i_1 = 1$  for  $\varepsilon$  small enough. In this case, we have  $I'_1 = 1$  and if  $1 > \alpha > 1/2$ , then  $R = 1$  from Theorem 2.4. The final reduced system reads

$$\frac{d}{dt} \begin{pmatrix} y_2^1 \\ y_3^1 \end{pmatrix} = \begin{pmatrix} -1 - \varepsilon/2 & 1/2 \\ 2 - \varepsilon^2/2 & -1 + \varepsilon/2 \end{pmatrix} \begin{pmatrix} y_2^1 \\ y_3^1 \end{pmatrix} = J^1 \begin{pmatrix} y_2^1 \\ y_3^1 \end{pmatrix}, \quad y_1^1 = (1 - \varepsilon^2/2) y_2^1 + \varepsilon/2 y_3^1.$$

By Theorem 2.1, we know that the eigenvalues of the reduced matrix  $J^1$  are  $-2$  and  $0$ .

On the other hand, the eigenvalues of the matrix  $\varepsilon^{-1}A^1$  are  $-2\varepsilon^{-1}$  and 0 (with multiplicity 2). Under the above notations, one has  $\underline{b}_0 = (1, -1, 0)$ ,  $\underline{z}_1 = 1$  and the reduced system is

$$\frac{d}{dt} \begin{pmatrix} \underline{y}_2^1 \\ \underline{y}_3^1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} \underline{y}_2^1 \\ \underline{y}_3^1 \end{pmatrix} = \underline{J}^1 \begin{pmatrix} \underline{y}_2^1 \\ \underline{y}_3^1 \end{pmatrix}, \quad \underline{y}_1^1 = \underline{y}_2^1. \quad (4.4)$$

We then observe that the matrices  $J^1$  and  $\underline{J}^1$  are different. In particular, 0 is an eigenvalue of  $J^1$  and not of  $\underline{J}^1$ . The invariants are not conserved in (4.4):  $(\underline{y}_1^1, \underline{y}_2^1, \underline{y}_3^1)$  converges to  $(0, 0, 0)$  whereas  $y_1^1 + y_2^1 + y_3^1$  is approximately equal to the constant  $y_1(0) + y_2(0) + y_3(0) > 0$ .

## 5 Numerical tests with 2 or 3 scales

In this section, we apply our reduction method to two chemical kinetic systems: one theoretical example with three different scales and one realistic example with two scales.

Let us begin with a theoretical reaction network with 10 species, 13 reactions and 3 time scales  $\varepsilon^{-2}$ ,  $\varepsilon^{-1}$  and 1, which is drawn on Figure 1. Let us choose  $\varepsilon = 10^{-2}$  and  $y_i(t=0) = 1$  for each  $i$ . The exact eigenvalues of the associated matrix  $J$  are:  $-20000$ ,  $-10100$ ,  $-10001$ ,  $-101$ ,  $-100$ ,  $-100$ ,  $-2$ ,  $-1$ ,  $-1$  and 0, and one has  $I_1 = I'_1 = 1$ . After a first reduction, the eliminated “fast” species are  $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ . We solve the reduced system ( $S^3$ ) on the interval  $I_\varepsilon = [1.2 \varepsilon^2 |\ln \varepsilon^2|, 1.2 \varepsilon |\ln \varepsilon|]$  and the evolution of the concentration of the “slow” species  $\mathcal{A}_9$ , which is calculated either with the detailed system, or with our reduced system or with the reduced system obtained from the q.s.s. approximation, is plotted on Figure 2 (left). All the curves for  $\mathcal{A}_9$  are almost identical (as for the other species). The system has been reduced once more to remove the time scale  $\mathcal{O}(\varepsilon^{-1})$ . The eliminated “intermediary” species for this second reduction are  $\mathcal{A}_4, \mathcal{A}_5, \mathcal{A}_6$ . The evolution of the concentrations of the same “slow” species  $\mathcal{A}_9$  after the time  $1.2 \varepsilon |\ln \varepsilon| \simeq 0.055$  is plotted on Figure 2 (right), where  $r2$  (resp.  $QSSA2$ ) means the solution obtained from our reduction method (resp. the q.s.s.a.) applied twice. The computational cost to find the solutions obtained by our method, say on the interval  $I_\varepsilon$ , is close to  $4s$  (including the calculations of the eigenvalues and eigenvectors) on a workstation, whereas it is about  $32s$  for the exact solution and  $4s$  for the q.s.s.a.

The second example arises in biochemistry and modelizes the dynamics of thyroid hormones, see [18]. The chemical network involves 8 species and 14 exchange rates between these species, and, in first approximation and up to a normalization of some volumes, the reaction network is that of Figure 3, under the notation of [18]. The fastest reactions are those whose rates are 5, 2.5, 1 (twice) and the reactants of these reactions are  $\mathcal{A}_1, \dots, \mathcal{A}_4$ . The eigenvalues of this system are  $-5.159$ ,  $-2.517$ ,  $-1.300$ ,  $-1.091$ ,  $-1.231 \cdot 10^{-3}$ ,  $-1.176 \cdot 10^{-5}$  and 0 (twice). Four steps are needed to get the final reduced system. The evolution of the concentration of one of the slow species, say  $\mathcal{A}_6$  ( $T4$ ), after the initial boundary layer and up to  $9 \cdot 10^4 s$ , is plotted on Figure 4 (left) using either our method (total computational cost =  $4s$ ), or the q.s.s.a. ( $4s$ ), or the exact solution ( $584s$  with a Gear scheme). The curves are almost identical. The evolution of one of the fast species, say  $\mathcal{A}_1$  ( $T3F$ ), is plotted on Figure 4 (right), up to  $9 \cdot 10^3 s$ . The total computational costs are  $4s$  for our reduction method and the q.s.s.a., and  $69s$  for the exact solution.

As a conclusion, our reduced method, which is not based on chemical solutions unlike the q.s.s. method, provides very good approximations of the exact solution and it is much faster.

**Acknowledgments.** The authors would like to thank the anonymous referee for making constructive remarks that helped greatly to improve the manuscript. The authors are also grateful to A. Bonnet and Y. Maday for helpful discussions during the 1996 CEMRACS summer school, and to C. Berthon, E. Billette, P. Joly and B. Sportisse.

## References

- [1] P. Ayoub, A. Bamberger, Z. Benjelloun-Dabaghi, *Méthodes mathématiques et numériques pour la réduction du schéma cinétique de production de l'ozone*, IFP Report **41362**, 1994.
- [2] G. Balakrishnan, F. Mauss, C. Trevino, *The asymptotic structure of Hydrogen-Air diffusion flames*, Comb. Flame **91** (1992), pp 246-256.
- [3] E. Billette, *Dissertation*, IFP-École Polytechnique, 1997.
- [4] A. Blouza, F. Coquel, F. Hamel, *Reduction of a class of kinetic systems with multiple scales*, Preprint Laboratoire d'Analyse Numérique, Paris VI, 1998.
- [5] S.L. Campbell, N.J. Rose, *Singular perturbation of autonomous linear systems*, SIAM J. Math. Anal. **10** (1979), pp 542-551. II, J. Diff. Eq. **29** (1978), pp 362-373, III, Houston J. Math. **4** (1978), pp 527-539.
- [6] J. Carr, *Applications of the centre manifold theory*, Appl. Math. Sci. **35**, Springer-Verlag, 1981.
- [7] J.Y. Chen, *Development of reduced mechanisms for numerical modelling of turbulent combustion*, In: Proc. Workshop on Numerical Aspects of Reduction in Chemical Kinetics, CERMICS-ENPC, 1997.
- [8] G. Dahlquist, *On transformations of graded matrices with applications to stiff ode's*, Numer. Mathematik **47** (1985), pp 363-385.
- [9] P. Duchêne, P. Rouchon, *Kinetic scheme reduction via geometric singular perturbation techniques*, Chem. Eng. Science **51** (1994), pp 4661-4672.
- [10] L. Elsner, *On the variation of the spectra of matrices*, Lin. Alg. Appl. **47** (1982), pp 127-138.
- [11] E. Hessvedt, Ø. Hov, I. Isaacsen, *Quasi-steady state assumption in air pollution modelling*, Intl. J. Chem. Kinetics **10** (1978), pp 971-994.
- [12] F. Hoppensteadt, *Properties of solutions of ordinary differential equations with small parameters*, Comm. Pure Appl. Math. **24** (1971), pp 807-840.
- [13] F. Hoppensteadt, *Analysis of some problems having matched asymptotic expansion solutions*, SIAM Rev. **17** (1974), pp 123-135.
- [14] S.H. Lam, D.A. Goussis, *Basic theory and demonstration of computational singular perturbation for stiff equations*, IMACS, Num. Appl. Math., Part II (Paris 1988), pp 487-492.
- [15] S.H. Lam, D.A. Goussis, *The CSP method for simplifying kinetics*, Intl. J. Chem. Kin. **26** (1994), pp 461-486.
- [16] B. Larrouturou, B. Sportisse, *Some mathematical and numerical aspects of reduction in chemical kinetics*, Comput. Science for the 21st Century, Conf. in honor of Pr. R. Glowinski, J. Wiley and Sons Ed., 1997.
- [17] U. Maas, S.B. Pope, *Simplifying chemical kinetics: Intrinsic low-dimensional manifolds in composition space*, Comb. Flame **88** (1992), pp 239-264.



- [18] P.H. Mak, J.J. DiStefano, *Optimal control policies for the prescription of thyroid hormones*, Math. Biosci. **42** (1978), pp 159-186.
- [19] R.E. O'Malley, *Singular perturbation methods for ordinary differential equations*, Appl. Math. Sci. **89**, Springer-Verlag, 1991.
- [20] M. Nau, A. Wöfert, U. Maas, J. Warnatz, *Applications of a combined pdf/finite volume scheme on turbulent methane diffusion flames*, Proc. 8th Int. Symp. Transp. Phenomena in Combustion, San Francisco, 1995.
- [21] N. Peters, *Systematic reduction of flame kinetics: principles and details*, In: Fluid dynamic aspects of combustion theory, J. Wiley and Sons, 1991.
- [22] N. Peters, B. Rogg, *Reduced reaction mechanisms for applications in combustion systems*, Lect. Notes in Physics, Springer-Verlag **m 15**, 1995.
- [23] N. Peters, F.A. Williams, *The asymptotic structure of stoichiometric methane-air flames*, Comb. Flame **68** (1987), pp 185-207.
- [24] P. Plion, J. Roux, *Analyse mathématique des méthodes de réduction, globalisation des schémas de cinétique chimique, I: Techniques utilisées en combustion, II: Techniques algébriques utilisées en chimie industrielle (regroupement d'espèces)*, EDF Reports, 1994.
- [25] A. Sandu, J.G. Verwer, J.G. Blom, E.J. Spee, G.R. Carmichael, *Benchmarking stiff ODE solvers for atmospheric chemistry problems, II: Rosenbrock solvers*, Atm. Env. **31** (1997), pp 3459-3472.
- [26] L.A. Segel, M. Slemrod, *The quasi-steady state assumption: a case study in perturbation*, SIAM Review **31** (1989), pp 446-477.
- [27] M.D. Smooke (ed.), *Reduced kinetic mechanisms and asymptotic approximations for methane-air flames*, Lecture Notes in Physics **384**, Springer-Verlag, 1991.
- [28] B. Sportisse, *Dissertation*, CERMICS-ENPC, 1998.
- [29] A. Tikhonov, A. Vasil'eva, A. Sveshnikov, *Differential equations*, Springer-Verlag, 1985.
- [30] A.S. Tomlin, T. Turanyi, M.J. Pilling, *Mathematical tools for the construction, investigation and reduction of combustion mechanisms*, In: "Low temperature combustion and auto-ignition", Ed.: M.J. Pilling and G. Hancock, Elsevier, 1997, pp 293-437.
- [31] V. Van Breusegem, G. Bastin, *Reduced order dynamical modelling of reaction systems: a singular perturbation approach*, In: Proc. of the 30th Conf. Decision and Control, Brighton (UK), 1991, pp 1049-1054.
- [32] F. Williams, *Combustion Theory*, Addison-Wesley, Reading MA, 1983.

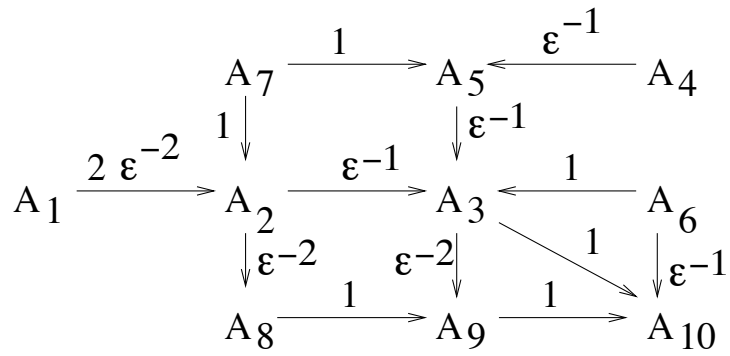


Figure 1: Theoretical reaction scheme with 3 time scales

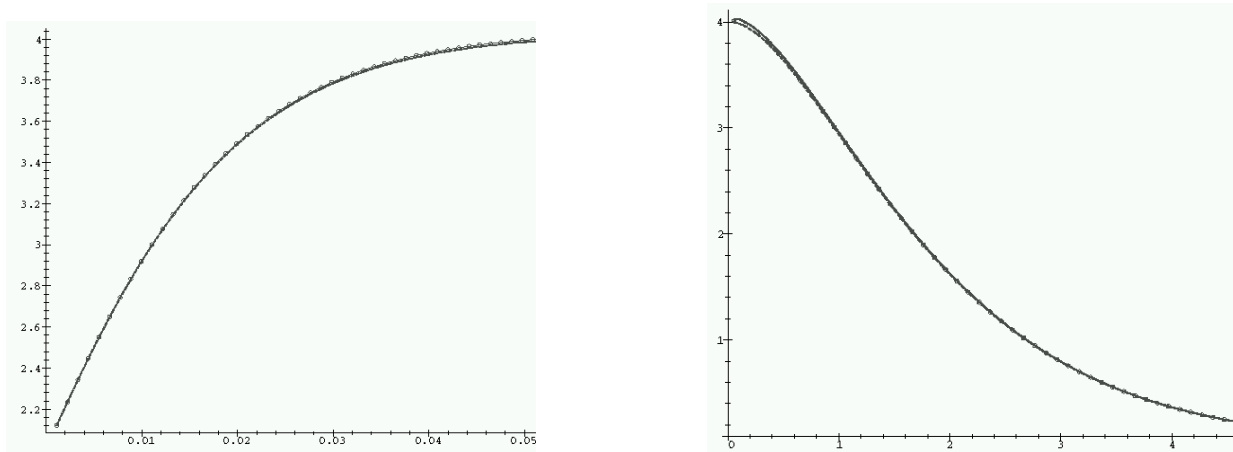


Figure 2: Theoretical network, “slow” species  $\mathcal{A}_9$ : —,  $y_9(t)$ ;  $\circ$ ,  $y_9^r(t)$ ; — —,  $y_9^{QSSA}(t)$ , after first (left) and second (right) reduction

