Master equation lumping for multi-well potential energy surfaces: a bridge between *ab initio* based rate constant calculations and large kinetic mechanisms

Luna Pratali Maffei[1], Matteo Pelucchi[1]*, Carlo Cavallotti[3], Andrea Bertolino[1,2], Tiziano Faravelli[1]

[1]*CRECK Modelling Lab, Department of Chemistry, Materials and Chemical Engineering "G. Natta", Politecnico di Milano, Milano, Italy*

[2]*Aero-Thermo-Mechanics Laboratory, Université Libre de Bruxelles, Ecole polytechnique de Bruxelles, Bruxelles, Belgium*

[3]*Department of Chemistry, Materials and Chemical Engineering "G. Natta", Politecnico di Milano, Milano, Italy*

Corresponding author:

Dr. Matteo Pelucchi,

Email: matteo.pelucchi@polimi.it

Address: Politecnico di Milano,

20133, Piazza Leonardo da Vinci 32,

Milano, Italy

Abstract

Ab initio transition state theory-based master equation methodologies for the calculation of rate constants have gained enormous popularity in the past decades. Nevertheless, introducing these rate constants into large kinetic schemes is a non-trivial task when large potential energy surfaces (PESs) are investigated. To determine proper phenomenological rate constants it is in fact necessary to account for the formation of all the thermodynamically stable wells considered in the master equation (ME), even if most wells do not exhibit significant secondary reactivity. Moreover, reactions involving intermediates with lifetimes comparable to the rovibrational relaxation timescale can exhibit discontinuities both in the rate constants and in the number of thermodynamically stable wells across the investigated temperature and pressure ranges. In this work, we address these problems with a "master equation-based lumping" (MEL) approach specifically designed to

process the output of ME calculations of multi-well PESs. Simple kinetic simulations allow identifying both intermediate wells with limited lifetime and isomers with similar reactivity. Then, equivalent rate constants for a smaller set of pseudospecies are derived so as to reproduce the kinetics of the detailed mechanism. Our methodology is independent of any experimental data or experience-based assumptions. The power of MEL is demonstrated with three case studies of increasing complexity, namely the PES for CH_3COOH decomposition, and the portions of the C_5H_5OH and $C_{10}H_{10}/C_{10}H_9$ PESs accessed from C_5H_5+OH and C_5H_5+C_5H_5 recombination. This work constitutes the first systematic step addressing the robust integration of rate constants derived from ME simulations into global kinetic schemes and provides a useful approach for the entire chemical kinetics community filling the gap between detailed theoretical investigations of complex PESs and the development of detailed kinetic models.

Highlights:

- Master equation-based chemical lumping (MEL) of potential energy surfaces (PESs)
- MEL allows >90% reduction of species and reactions for a complex C10H10/C10H9 PES
- An open-source software implementing MEL is provided for interested users
- MEL connects theoretical chemistry and the development of complex kinetic models

Keywords: chemical lumping, kinetic modelling, master equation, chemical reaction engineering, combustion chemistry

Nomenclature

c_I	Concentration of species I
I, J	Generic species of the detailed mechanism
k_{IJ}	Rate constant of reaction $I \rightarrow J$
k'_{IJ}	Rate constant of a pseudo uni-molecular reaction, i.e. for bimolecular reactions $I + I_2$, $k'_{IJ} = k_{IJ} c_{I_2}$
L	Generic species of the detailed mechanism contributing to $\tilde{I}$
N	Number of species
n_i	Population of the system in the quantum state i (in the master equation context)
p_{ij}	Probabilities per unit time of a transition from state i to state j (in the master equation context)
p	Pressure
$r_{IJ}(t)$	Time-dependent flux of molecules from configuration I to configuration J
R, P	Generic reactant (present in the system at $t = 0$) or product (possibly unreactive, i.e. infinite sink)
T	Temperature
$\sim$	Accent indicating quantities of the lumped mechanism

Acronyms

BF	Branching Fractions
CSE	Chemically Significant Eigenvalues
IERE	Internal-Energy Relaxation Eigenvalues
ES	Electronic Structure

AITSTME	*ab initio* transition state theory-based master equation
ME	Master Equation
MEL	Master Equation based Lumping
PES	Potential Energy Surface
SM	Supplementary Material

1. Introduction

The use of *ab initio* transition state theory-based master equation (AITSTME) methods for the determination of the kinetics of complex gas-phase reactions has grown exponentially in the past decades [1]. In fact, both the increase in computational power and the development of automatic routines for electronic structure (ES) calculations [2,3], reaction mechanism generation [4,5], potential energy surface (PES) exploration [6,7], and reaction rate constant calculation [8–10], simplified and greatly accelerated the application of AITSTME methodologies. Such theoretical investigations allow exploring new reaction pathways relevant to several fields of application, such as interstellar, atmospheric, and combustion chemistry. For instance, a significant research area in combustion chemistry revolves around the reactivity of polyaromatic hydrocarbons (PAHs), essential to the understanding of the mechanism of soot (i.e. particulate matter) formation. The latest theoretical studies in this field focus on PAHs growth and oxidation pathways, resulting in extremely complex multi-well PESs with several reaction intermediates [11–15]. The corresponding rate constants are determined solving the master equation (ME) with codes such as MESS [8,9], MESMER [10], MultiWell [16], or the MC-RRKM [17] stochastic solver. The output of ME simulations consists of up to $\mathrm{N}(\mathrm{N}-1)$ phenomenological rate constants describing the reactivity of N species involved in the PES at a given temperature and pressure. However, testing the impact of these results on reproducing macroscopic experimental trends requires the inclusion of the determined rate constants into existing kinetic mechanisms, typically containing already a large number of species (10^2-10^3) and reactions (10^3-10^4). Complex PESs theoretically investigated generally involve a large number of reaction intermediates, most of which are not even present in the target kinetic mechanism. Indeed, including new reaction parameters leads to two major problems: firstly, the new rate constants do not always cover the full range of temperature and pressure of ME simulations because of thermodynamic or kinetic instability of some ME reaction channels, as discussed below. Secondly, the inclusion of the full set of intermediates results in a dramatic increase of the number of reactions and species, making the solution of even the simplest 0-D or 1-D reactive system tedious. Such a large number of species also complicates the understanding of the macroscopic role of the investigated pathways in the global reaction network, mining the fundamentally rigorous description of the

target phenomena. A viable pathway to reduce the number of isomers with similar structure and reactivity in a PES is to lump them into pseudospecies.

Lumping has a long tradition in chemical kinetics. Detailed kinetic mechanisms are generally used for kinetic simulations of 0-D or 1-D ideal reactors (e.g. batch, flow and stirred reactors, laminar flames) to reproduce macroscopic experimental targets (e.g. species concentrations, ignition delay times, laminar flame speeds etc.) in model validation processes. These models are typically constituted of a very large set of species (10^3-10^4) and rate constants (10^3-10^5), which in principle should retain chemical significance. The size of these models may be reduced significantly with chemical lumping and mechanism reduction techniques [18] when large scale computations (i.e. 2-D and 3-D reactive computational fluid dynamics) have to be performed, as illustrated qualitatively in Figure 1.

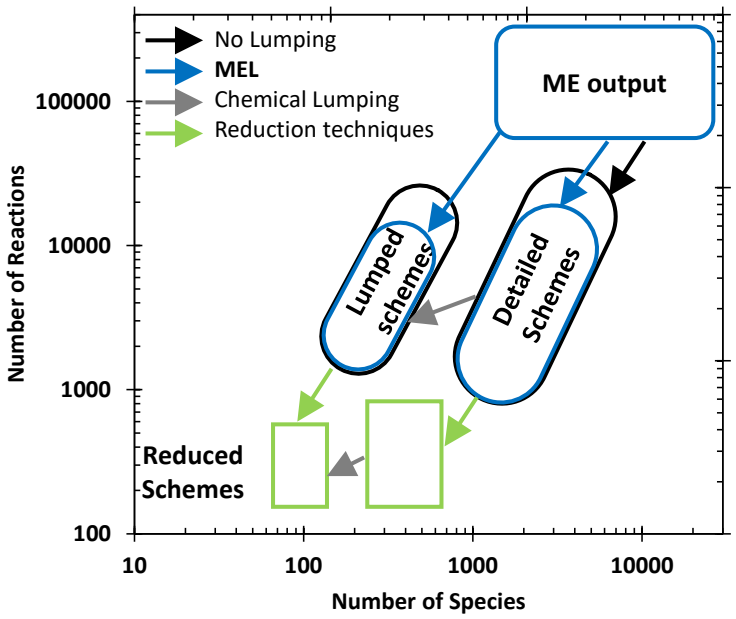

Figure 1: Qualitative representation of the typical number of reactions and species of detailed, lumped and reduced global kinetic schemes. Adapted from Ranzi et al. [18].

Chemical lumping techniques derive equivalent rate constants for "pseudospecies", i.e. subsets of species of the detailed mechanism. The key of chemical lumping is the calculation of the contribution of each isomer to the total concentration of the pseudospecies $\tilde{c}_{\tilde{I}} = \sum_L c_L$ through coefficients $\alpha_L = \frac{c_L}{\sum_L c_L}$, where $0 \leq \alpha_L \leq 1$. The first methodologies proposed and largely applied by Ranzi et al [19–22]. lump highly reactive radicals. α_L is found according to pseudo-steady state approximation (PSSA) at fixed temperature, and the resulting $\tilde{k}_{\tilde{I}\tilde{J}} =$

$\sum_{L,J} \alpha_L k_{LJ}$ are optimized according to the selectivity of primary reaction products. Temperature dependence of product selectivity may be introduced by fitting α_L with Arrhenius expressions, as in Chaos et al [23]. This approach is however limited to highly reactive species with short lifetimes, since the PSSA approximation fails in the presence of significant interactions with the reacting mixture. A more rigorous but mathematically stringent approach was originally developed by Wei and Kuo [24,25]. In this case, the ordinary differential equation (ODE) system describing the chemical species mass balances is transformed into a smaller system using a lumping matrix. Although the solution of the system is exact, the detailed set of rate constants is subject to several formal constraints and it excludes bimolecular reactions. A simplified and less stringent approach was proposed by Huang et al. [26], similar to that adopted by Nagy and Turanyi [27]. The analysis of the Jacobian of the system allows the removal of species and reactions unnecessary to the description of the global chemical reactivity, identified with techniques such as principal component analysis (PCA). One of the downsides of this approach is that lumped groups may lose chemical significance because they are not necessarily constituted of isomers. Chemical significance is recovered in the approach of Pepiot-Desjardins and Pitsch [28], where chemical structure determines the constituents of pseudospecies [29]. Although this approach generally retains good model performances, simulations at several different conditions are required and the results show dependence on the optimization subspace. Finally, the range of applicability of all the described approaches depends on the conditions of the performed kinetic simulations.

Despite the variety of chemical lumping methods widely used for the reduction of global kinetic mechanisms [30], such tools have never been specifically customized nor applied to the theoretical investigation of multi-well PESs. The aim of this work is therefore to develop a general methodology to process the output of ME calculations of such PESs into a smaller set of equivalent rate constants able to retain the accuracy in the description of reactivity and products distribution obtained with a fully detailed approach. We refer to this methodology as “Master Equation-based Lumping” (MEL). The MEL procedure is illustrated qualitatively in Figure 2. Starting from a PES with N species and $N(N-1)$ rate constants k_{IJ} of the elementary reactions $I \rightarrow J$, the number of species is reduced to $\widetilde{N} < N$ pseudospecies by either removing redundant species or grouping together species with similar chemical behavior. The reactivity of this new set of pseudospecies is described by rate constants $\tilde{k}_{\tilde{I}\tilde{J}}$ of lumped reactions $\tilde{I} \rightarrow \tilde{J}$.

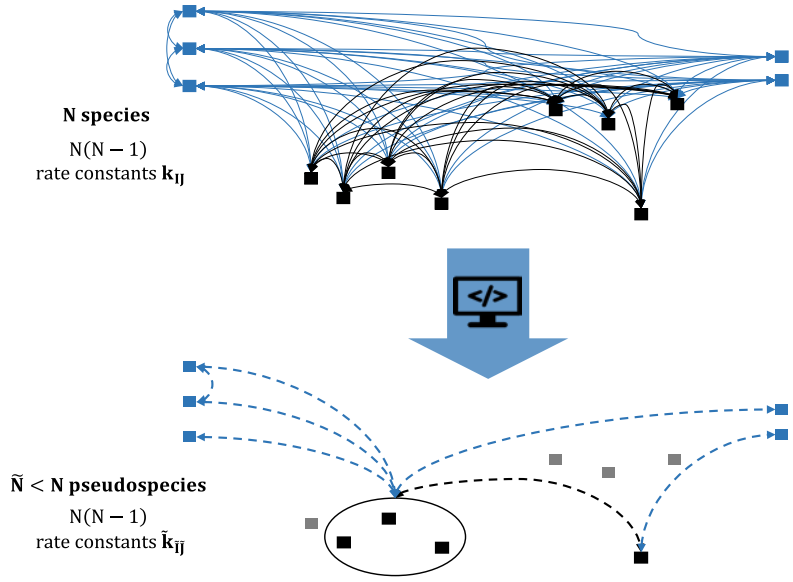

Figure 2: Qualitative representation of MEL. Squares represent the species of the PES and arrows indicate the rate constants of the associated reactivity: solid lines are the original k_{IJ} of ME output, dashed lines are the lumped $\tilde{k}_{\tilde{I}\tilde{J}}$. Shaded squares are redundant species, eliminated in the lumped mechanism; blue squares are bimolecular reactants/products; circles group isomers as one pseudospecies.

The paper is organized as follows. The theoretical background of the methodology is described in Section 2, where Section 2.1 recalls fundamental aspects of ME methods, providing a theoretical framework for the development of MEL, described in Section 2.2. Details of the implementation are reported in Section 2.3, while some proposals for the selection of lumped species are described in Section 2.4. The results (Section 3) include three cases of application of the MEL methodology: CH_3COOH decomposition [31] (Section 3.1); the reactivity of the C_5H_5OH PES [32] (Section 3.2); and the analysis of 328 reactions taking place on the $C_{10}H_{10}/C_{10}H_9$ PESs [33], as accessed from the C_5H_5+C_5H_5 entrance channel (Section 3.3). Further validation is obtained including both the detailed and the lumped kinetic subsets into the global CRECK kinetic mechanism [34]. The conclusions of our work are summarized in Section 4.

2. Theoretical methodology

2.1 ME methods and phenomenological rate constants

In order to understand the principles on which the MEL approach is based it is useful to review some fundamental aspects related to the integration of the ME for chemically reactive systems. The time-dependent multi-well ME is written in its primitive form as [35]

$$\frac{dn_i}{dt} = \sum_j \left(p_{ji} n_j - p_{ij} n_i\right), \tag{1}$$

where n_i represents the population of the system in the quantum state i, and p_{ij}, p_{ji} are the probabilities per unit time of a transition between states i and j. In the one-dimensional ME, the population of states is expressed as a function of the energy of the system and the transition probabilities p_{ij}, p_{ji} describe either energy transfer upon collisions with the bath gas or microscopic rate coefficients. Equation (1) can generally be formulated in terms of an eigenvalue problem. The least negative "chemically significant eigenvalues" (CSEs) are relevant to describe the chemical evolution of the system in time (i.e. macroscopic change in species concentration), whereas the remaining "internal-energy relaxation eigenvalues" (IEREs) are related to the change in the rovibrational energy of the molecule. The time evolution of the population of each state may be obtained from the CSE of Eq. (1). Macroscopic populations X_I are then obtained from n_i by integration over the energy for each I chemical species. For the details of ME formulation and solution methods, we refer to the extensive literature published in the past decades [8,17,36–39]. The X_I time dependent profiles must then be properly post-processed to determine rate coefficients that satisfy the phenomenological law:

$$\frac{dX_I}{dt} = \left(-\sum_{J \neq I}^{N} k_{IJ}\right) X_I + \sum_{J \neq I}^{N} k_{JI} X_J \tag{2}$$

written in compact form as $\frac{d\mathbf{X}}{dt} = \mathbf{KX}$. I and J in (2) refer to all wells considered in the ME. For a set of N species, k_{IJ} constitute a unique set of phenomenological rate coefficients, which can be expressed, for example, as a linear combinations of the exponential functions of the CSE eigenvalues [8,37,38]. Well defined phenomenological rate coefficients can be obtained when the timescales of IEREs and CSEs are sufficiently separated [8,40,41].

It is important to highlight two key aspects related to the solution of the ME that are of relevance for the present work and, more in general, for the use of rate constants determined through ME methodologies in kinetic mechanisms:

- the development of a robust procedure to determine the phenomenological rate constants k_{IJ} from the ME solution has required significant effort, with key advancements obtained thanks to the seminal works

of Bartis and Widom [35,40] and of Miller and Klippenstein [8,38,42]. The set of $\mathrm{N}(\mathrm{N}-1)$ rate constants obtained through these procedures include both isomerization and well skipping reactions connecting reversibly all wells in the PES. In a well skipping reaction a species converts directly into an isomer or decomposes to bimolecular products without the collisional stabilization of the visited intermediate wells. It should be noted that though only a limited number of the $\mathrm{N}(\mathrm{N}-1)$ reactions are sufficiently fast to contribute significantly to the system reactivity, it is usually difficult to identify the key reaction steps without a kinetic simulation. Thus, selecting arbitrarily reaction channels that connect two specific wells from ME simulation results, neglecting for example pathways that may originate from other wells, should be done with care or preferably avoided.

- The determination of rate constants for PESs in which are present wells whose lifetime is comparable to the rovibrational relaxation timescale is complicated by the merging of CSEs and IEREs. From the molecular standpoint, this means that the well reacts to form another species before colliding with a bath gas molecule, a condition necessary to attain at least a partial thermal energy distribution. This situation is particularly critical at low pressures and high temperatures. A procedure to describe this process is to merge the highly reactive well into a vicinal more stable well, thus creating a single pseudospecies. The downside of this procedure is that, since the well merging conditions are temperature and pressure dependent, the same functional dependence extends to the total number of chemical species and reactions determined from the ME solution. It should be noted that if wells were not merged when necessary, it becomes possible that rate constants of isomerization reactions among wells change with the PES entrance channel, a situation that would lead to inconsistencies in kinetic simulations if not properly addressed.

2.2 The Master Equation Lumping approach

The aim of the MEL approach is to convert the ME mechanism, constituted by up to $\mathrm{N}(\mathrm{N}-1)$ elementary reactions for a multi-well PES containing N stationary points, into a "lumped mechanism" of $\widetilde{\mathrm{N}}$ psuedospecies and up to $\widetilde{\mathrm{N}}(\widetilde{\mathrm{N}}-1)$ reactions. MEL addresses the two main features of ME estimated rate constants highlighted in the previous section as it: 1) decreases the large number of rate constants derived with AITSTME methodologies which may lead to kinetic schemes of unmanageable sizes, as highlighted in Figure

1; 2) eliminates highly reactive species in the full temperature and pressure ranges considered. MEL has several features in common with chemical lumping, as they both solve the problem of reducing an ODE system for N species:

$$\frac{\mathrm{d}\mathbf{c}}{\mathrm{dt}} = \mathbf{K}'\mathbf{c} \tag{3}$$

to the equivalent lumped system for $\widetilde{\mathrm{N}} < \mathrm{N}$ pseudospecies:

$$\frac{\mathrm{d}\tilde{\mathbf{c}}}{\mathrm{dt}} = \widetilde{\mathbf{K}}'\tilde{\mathbf{c}} \tag{4}$$

In the MEL approach, (3) describes the reactivity of the ME system as an initial value problem imposing initial conditions $\mathbf{c}(\mathrm{t}=0) = \mathbf{c^0}$. For a multi-well PES, $\mathbf{K}$ contains the $\mathrm{N}(\mathrm{N}-1)$ ME rate constants. Apex $'$ indicates that rate constants of bimolecular reactions $\mathrm{I}+\mathrm{I}_2$ are multiplied by the concentration of the excess co-reactant $\mathrm{k}'_{\mathrm{IJ}} = \mathrm{k}_{\mathrm{IJ}}\mathrm{c}_{\mathrm{I}_2}$, thus becoming formally unimolecular and making $\mathbf{K}'$, and consequently $\widetilde{\mathbf{K}}'$, pseudo-unimolecular rate constant matrices, similarly to what described by Wei and Kuo [24,25]. Matrix $\mathbf{K}'$ of pseudo-unimolecular rate constants is:

$$\mathbf{K}' = \begin{bmatrix} -\sum_{\mathrm{J}\neq 1}^{\mathrm{N}} \mathrm{k}'_{1\mathrm{J}} & \mathrm{k}'_{21} & \mathrm{k}'_{31} & \dots & \dots & \dots & \mathrm{k}'_{\mathrm{N}1} \\ \mathrm{k}'_{12} & -\sum_{\mathrm{J}\neq 1}^{\mathrm{N}} \mathrm{k}'_{2\mathrm{J}} & \mathrm{k}'_{32} & \dots & \dots & \dots & \mathrm{k}'_{\mathrm{N}2} \\ \dots & \dots & \ddots & \dots & \dots & \dots & \dots \\ \mathrm{k}'_{1\mathrm{J}} & \mathrm{k}'_{2\mathrm{J}} & \mathrm{k}'_{3\mathrm{J}} & -\sum_{\mathrm{J}\neq \mathrm{I}}^{\mathrm{N}} \mathrm{k}'_{\mathrm{IJ}} & \dots & \dots & \mathrm{k}'_{\mathrm{NI}} \\ \dots & \dots & \dots & \dots & \dots & \ddots & \dots \\ \mathrm{k}'_{1\mathrm{N}} & \mathrm{k}'_{2\mathrm{N}} & \dots & \dots & \dots & \mathrm{k}'_{(\mathrm{N}-1)\mathrm{N}} & -\sum_{\mathrm{J}\neq \mathrm{N}}^{\mathrm{N}} \mathrm{k}'_{\mathrm{NJ}} \end{bmatrix} \tag{5}$$

The time evolution of generic species I is described by:

$$\frac{\mathrm{dc_I}}{\mathrm{dt}} = \left(-\sum_{\mathrm{J}\neq \mathrm{I}}^{\mathrm{N}} \mathrm{k}'_{\mathrm{IJ}}\right)\mathrm{c_I} + \sum_{\mathrm{J}\neq \mathrm{I}}^{\mathrm{N}} \mathrm{k}'_{\mathrm{JI}}\mathrm{c_J} \quad ; \quad \mathrm{c_I}(\mathrm{t}=0) = \mathrm{c_I^0} \tag{6}$$

The MEL lumping procedure is initiated by subdividing the N species of the ME system into two groups, accumulating and non-accumulating species. Only the former species are considered among the lumped $\widetilde{\mathrm{N}}$ pseudospecies. Non-accumulating species include both wells not collisionally equilibrated in ME simulations,

and highly reactive species not accumulating in macroscopic kinetic simulations. Merging isomers of stable species is possible and advised, should they equilibrate very rapidly or show similar reactivity (i.e. characteristic times and product branching fractions) in the considered T and p conditions. The ensemble of the merged and unmerged accumulating species constitutes the set of $\widetilde{\mathrm{N}}$ pseudospecies: pseudospecies profiles $\tilde{c}_{\tilde{I}}$ of system (4) describe either unmerged or merged accumulating species profiles of (3), i.e. $\tilde{c}_{\tilde{I}} = \sum_{L} c_{L}$, where the sum contains a single element in the case of unmerged species ($\tilde{c}_{\tilde{I}} = c_{L}$). A systematic procedure for identifying species to be lumped is described in Section 2.4. Having assigned a specific temperature and pressure, Equation (3) is then solved for $\widetilde{\mathrm{N}}$ times considering for each simulation one of the pseudospecies as reactant and all the other pseudospecies as products. We indicate the reactant and product pseudospecies as $\widetilde{\mathrm{R}}$ and $\widetilde{\mathrm{P}}$, constituted of either unmerged or merged species R and P, respectively: $\tilde{c}_{\tilde{R}} = \sum_{R} c_{R}$, $\tilde{c}_{\tilde{P}} = \sum_{P} c_{P}$. The reactant species R are the only ones with a non-zero initial concentration c_{R}^{0}. The infinite sink approximation is applied to the products P, namely their concentration can only increase in time. This is obtained modifying the $\mathbf{K}'$matrix (5) setting $k_{PJ} = 0$, hence the columns corresponding to all P species are set to 0. This approach, which is similar to that implemented by Barbato et al. to determine phenomenological rate constants from ME simulations [17], results in the exponential decay of the reactant $\widetilde{\mathrm{R}}$. Under the described conditions in fact, reactant $\widetilde{\mathrm{R}}$ with initial concentration $\tilde{c}_{\tilde{R}}^{0}$ decomposes to products $\widetilde{\mathrm{P}}$ according to the phenomenological law:

$$\begin{cases} \frac{d\tilde{c}_{\tilde{R}}}{dt} = \left(-\sum_{\tilde{P}} \tilde{k}'_{\tilde{R}\tilde{P}}\right)\tilde{c}_{\tilde{R}} \;\; ; \; \tilde{c}_{\tilde{R}}(t=0) = \tilde{c}_{\tilde{R}}^{0} \\ \frac{d\tilde{c}_{\tilde{P}}}{dt} = \tilde{k}'_{\tilde{R}\tilde{P}}\tilde{c}_{\tilde{R}} \qquad\quad ; \; \tilde{c}_{\tilde{P}}(t=0) = 0 \end{cases} \tag{7}$$

The analytical integration of Equation (7) describes the exponential decay for $\widetilde{\mathrm{R}}$. The unknown rate constants $\tilde{k}'_{\tilde{R}\tilde{P}}$ of reaction $\widetilde{\mathrm{R}} \rightarrow \widetilde{\mathrm{P}}$ can then be derived from a differential fit of the product concentration profiles of Equation (7). For self-reactions $2\widetilde{\mathrm{R}} \rightarrow \widetilde{\mathrm{P}}$ and reactions with two products $\widetilde{\mathrm{R}} \rightarrow \widetilde{\mathrm{P}} + \widetilde{\mathrm{P}}_{2}$, the evolution of $\tilde{c}_{\tilde{P}}$ is described by $\frac{d\tilde{c}_{\tilde{P}}}{dt} = \tilde{k}'_{\tilde{R}\tilde{P}}\tilde{c}_{\tilde{R}}^{2}$ and $\frac{d\tilde{c}_{P}}{dt} = 2\tilde{k}'_{\tilde{R}\tilde{P}}\tilde{c}_{\tilde{R}}$, respectively. The coefficient of determination R^2 of the linear fits for $\tilde{k}'_{\tilde{R}\tilde{P}}$ provides a first estimate of the validity of the lumped model. Low values of R^2 may be related to a wrong choice of the set of pseudospecies or of the constituents of merged pseudospecies.

A simple example of application of MEL is the lumping of a bimolecular reaction proceeding through a non-accumulating intermediate $\mathrm{R} + \mathrm{R}_2 \leftrightarrow \mathrm{W} \leftrightarrow \mathrm{P} + \mathrm{P}_2$. Figure 3 illustrates the qualitative shape of the PES.

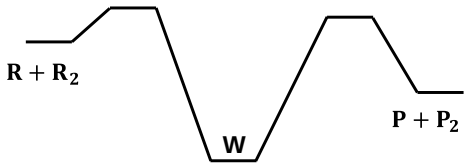

Figure 3: Qualitative PES of the reaction $R + R_2 \leftrightarrow W \leftrightarrow P + P_2$.

In this case, the goal of lumping is to obtain equivalent rate constants $\tilde{\mathrm{k}}_{\tilde{\mathrm{R}}\tilde{\mathrm{P}}}$ and $\tilde{\mathrm{k}}_{\tilde{\mathrm{P}}\tilde{\mathrm{R}}}$ of reaction $\tilde{\mathrm{R}} + \tilde{\mathrm{R}}_2 \leftrightarrow \tilde{\mathrm{P}} + \tilde{\mathrm{P}}_2$. Derivation of $\tilde{\mathrm{k}}_{\tilde{\mathrm{R}}\tilde{\mathrm{P}}}$ requires a set of simulations where R is the reactant and P is the infinite sink product, resulting in the following system of equations:

$$\begin{bmatrix} \frac{\mathrm{dc_R}}{\mathrm{dt}} \\ \frac{\mathrm{dc_W}}{\mathrm{dt}} \\ \frac{\mathrm{dc_P}}{\mathrm{dt}} \end{bmatrix} = \begin{bmatrix} -\mathrm{k_{RP}}\mathrm{c}^0_{\mathrm{R}_2} - \mathrm{c_{RW}} & \mathrm{k_{WR}} & 0 \\ \mathrm{k_{RW}} & -\mathrm{k_{WR}} - \mathrm{k_{WP}} & 0 \\ \mathrm{k_{RP}}\mathrm{c}^0_{\mathrm{R}_2} & \mathrm{k_{WP}} & 0 \end{bmatrix} \begin{bmatrix} \mathrm{c_R} \\ \mathrm{c_W} \\ \mathrm{c_P} \end{bmatrix}, \quad \begin{matrix} \mathrm{c}^0_{\mathrm{R}} = 10^{-5}\mathrm{c}^0_{\mathrm{tot}} \\ \mathrm{c}^0_{\mathrm{R}_2} = \mathrm{c}^0_{\mathrm{tot}} - \mathrm{c}^0_{\mathrm{R}} \\ \mathrm{c}^0_{\mathrm{W}}, \mathrm{c}^0_{\mathrm{P}}, \mathrm{c}^0_{\mathrm{P}_2} = 0 \end{matrix} \tag{8}$$

If the assumption of high reactivity of intermediate W is correct, the solution of (8) will show that W does not accumulate significantly in the system and the profiles of $\tilde{\mathrm{c}}_{\tilde{\mathrm{R}}}$ and $\tilde{\mathrm{c}}_{\tilde{\mathrm{P}}}$ are well described by

$$\begin{cases} \frac{\mathrm{d}\tilde{\mathrm{c}}_{\tilde{\mathrm{R}}}}{\mathrm{dt}} = -\tilde{\mathrm{k}}_{\tilde{\mathrm{R}}\tilde{\mathrm{P}}}\tilde{\mathrm{c}}^0_{\tilde{\mathrm{R}}_2}\tilde{\mathrm{c}}_{\tilde{\mathrm{R}}} & \tilde{\mathrm{c}}^0_{\tilde{\mathrm{R}}} = 10^{-5}\mathrm{c}^0_{\mathrm{tot}} \\ & \tilde{\mathrm{c}}^0_{\tilde{\mathrm{R}}_2} = \mathrm{c}^0_{\mathrm{tot}} - \tilde{\mathrm{c}}^0_{\tilde{\mathrm{R}}} \\ \frac{\mathrm{d}\tilde{\mathrm{c}}_{\tilde{\mathrm{P}}}}{\mathrm{dt}} = \tilde{\mathrm{k}}_{\tilde{\mathrm{R}}\tilde{\mathrm{P}}}\tilde{\mathrm{c}}^0_{\tilde{\mathrm{R}}_2}\tilde{\mathrm{c}}_{\tilde{\mathrm{R}}} & \tilde{\mathrm{c}}^0_{\tilde{\mathrm{P}}} = 0 \end{cases} \tag{9}$$

Where $\tilde{\mathrm{c}}_{\tilde{\mathrm{R}}}$ and $\tilde{\mathrm{c}}_{\tilde{\mathrm{P}}}$ derive from the solution of (8). $\tilde{\mathrm{k}}_{\tilde{\mathrm{R}}\tilde{\mathrm{P}}}$ is determined with a linear differential fit of the profiles of $\frac{\mathrm{d}\tilde{\mathrm{c}}_{\tilde{\mathrm{P}}}}{\mathrm{dt}} = \tilde{\mathrm{k}}_{\tilde{\mathrm{R}}\tilde{\mathrm{P}}}\mathrm{c}^0_{\tilde{\mathrm{R}}_2}\tilde{\mathrm{c}}_{\tilde{\mathrm{R}}}$ in (9). The backward rate constant $\tilde{\mathrm{k}}_{\tilde{\mathrm{P}}\tilde{\mathrm{R}}}$ is determined in a similar way. Differently from this simple example, in most cases each pseudospecies is formed by multiple isomers and it forms multiple products.

The methodology proposed has several advantages. Firstly, it requires only the output of ME simulations, being independent of kinetic simulations of experimental data which usually limit the applicability range of the lumped rate constants. Then, contrary to most chemical lumping approaches, MEL allows simultaneous elimination of redundant species and merging of accumulating isomers with similar reactivity, preserving the

chemical meaning of the mechanism and easily including both unimolecular and bimolecular reaction channels.

2.3 Implementation of MEL methodology

The MEL approach is implemented in an in-house software written in Python (available at https://github.com/lpratalimaffei/MEL) and relies on OpenSMOKE++ [43] for the solution of the ODE system and, optionally, on OptiSMOKE [44,45] for further optimization of the lumped mechanism. The flowchart in Figure 4 depicts the main steps performed by the code. Three main types of simulations may be run: "prescreening" involves the analysis of the reactivity of species in the detailed mechanism; "lumping" allows the derivation of rate constants from the reactivity of the selected sets of accumulating species; "validation" consists in the comparison of the reactivity of the detailed mechanism with that of the lumped mechanism. The main steps are common to all simulation types. Input specifications include the set of reactants R and products P (i.e. the constituents of pseudospecies $\widetilde{\mathrm{R}}, \widetilde{\mathrm{P}}$), as well as the ranges of pressure and temperature of the simulations. For a PES with $\widetilde{\mathrm{N}}$ sets of accumulating species, lumping requires a total of $\widetilde{\mathrm{N}} \times \mathrm{N_p} \times \mathrm{N_T}$ simulations performed at isobaric and isothermal conditions, where $\mathrm{N_p}$ and $\mathrm{N_T}$ are the number of nominal pressures and temperatures at which simulations are performed. The matrix $\mathbf{K}$ of the ME system, namely the values of ME phenomenological rate constants $\mathrm{k_{IJ}}|_{\mathrm{T,p}}$ at the specified conditions, are extracted from ME output files. For lumping simulations, products are treated as infinite sinks, i.e $\mathrm{k_{PJ}} = 0$. Thermodynamics is irrelevant because all reactions are written irreversibly in the forward or backward direction, as also highlighted by Pepiot-Desjardins and Pitsch [28]. When $\tilde{\mathrm{k}}_{\tilde{\mathrm{I}}\tilde{\mathrm{J}}}$ are introduced in a global kinetic scheme that already includes reversible reactions of pseudospecies $\tilde{\mathrm{I}}, \tilde{\mathrm{J}}$, it is recommended to consider either the thermochemistry of the most stable isomer [29], or a weighted average of the thermochemistry of the isomers constituting the pseudospecies [28]. System (3) is then solved with OpenSMOKE++ [43]. Initial conditions for the composition $\mathrm{c_R^0}$ of the reactants are specified as follows:

- Unimolecular reactions $R \rightarrow P$: for single reactant species, $c_R^0 = c_{tot}$; for merged species, $\sum_R c_R^0 = c_{tot}$, with c_R^0 derived from previous simulations or determined iteratively from a random guess. (details in Section 2.4.2).
- Bimolecular reactions $R + R_2 \rightarrow P$: for single bimolecular deficient fragment R, $c_R^0 = 10^{-5} c_{tot}$; for merged fragments, $\sum_R c_R^0 = 10^{-5} c_{tot}$. The concentration of the excess fragment is approximately constant: $c_{R_2}^0 = c_{tot} - \sum_R c_R^0$. For self-reactions $2R \rightarrow P$, the total concentration is kept approximately constant by immersing R in an inert bath gas (e.g. Ar) consistent with that used for ME simulations: $c_R^0 = 10^{-5} c_{tot}$, and $c_{Ar}^0 = c_{tot} - c_R^0$ (see Section 3.3).

The resulting profiles are then processed. Simulation times that lead to less than 1% or more than 99% reactant consumption are excluded, so as to avoid either numerical instabilities or non-exponential behavior due to initial population of the intermediate wells. Examples of the sensitivity of $\tilde{k}_{\tilde{R}\tilde{P}}$ to such cutoff limits are provided in the results (Section 3.1). Then, merged pseudospecies profiles $\tilde{c}_{\tilde{I}} = \sum_L c_L$ (System (4)), where L is a generic constituent of merged pseudospecies $\tilde{I}$, are reconstructed if necessary. The weighted average of the branching fractions (BFs) of constituent L in each merged pseudospecies $\tilde{I}$ are computed as

$$\alpha_L|_{T,p} = \sum_{dt} \frac{c_L}{\tilde{c}_{\tilde{I}}}\Big|_{dt} \cdot \frac{dt}{t_{tot}} \tag{10}$$

where $c_L/\tilde{c}_{\tilde{I}}|_{dt}$ is the contribution of species L to merged pseudospecies $\tilde{I}$ in the timeframe dt, and dt/t_{tot} is the corresponding fraction of time. $\alpha_L|_{T,p}$ may be used for the selection of c_R^0 in successive simulations, as depicted in Figure 4.

The operations described are repeated for the full set of operating conditions T, p specified in the input. For lumping simulations, the resulting $\tilde{k}_{\tilde{R}\tilde{P}}|_{T,p}$ are fitted according to modified Arrhenius expressions. Reiterating these steps for each set of reactants R results in the final lumped mechanism of $\tilde{N}(\tilde{N}-1)\,\tilde{k}_{\tilde{I}\tilde{J}}$, which may be further optimized with tools such as OptiSMOKE++ [44,45].

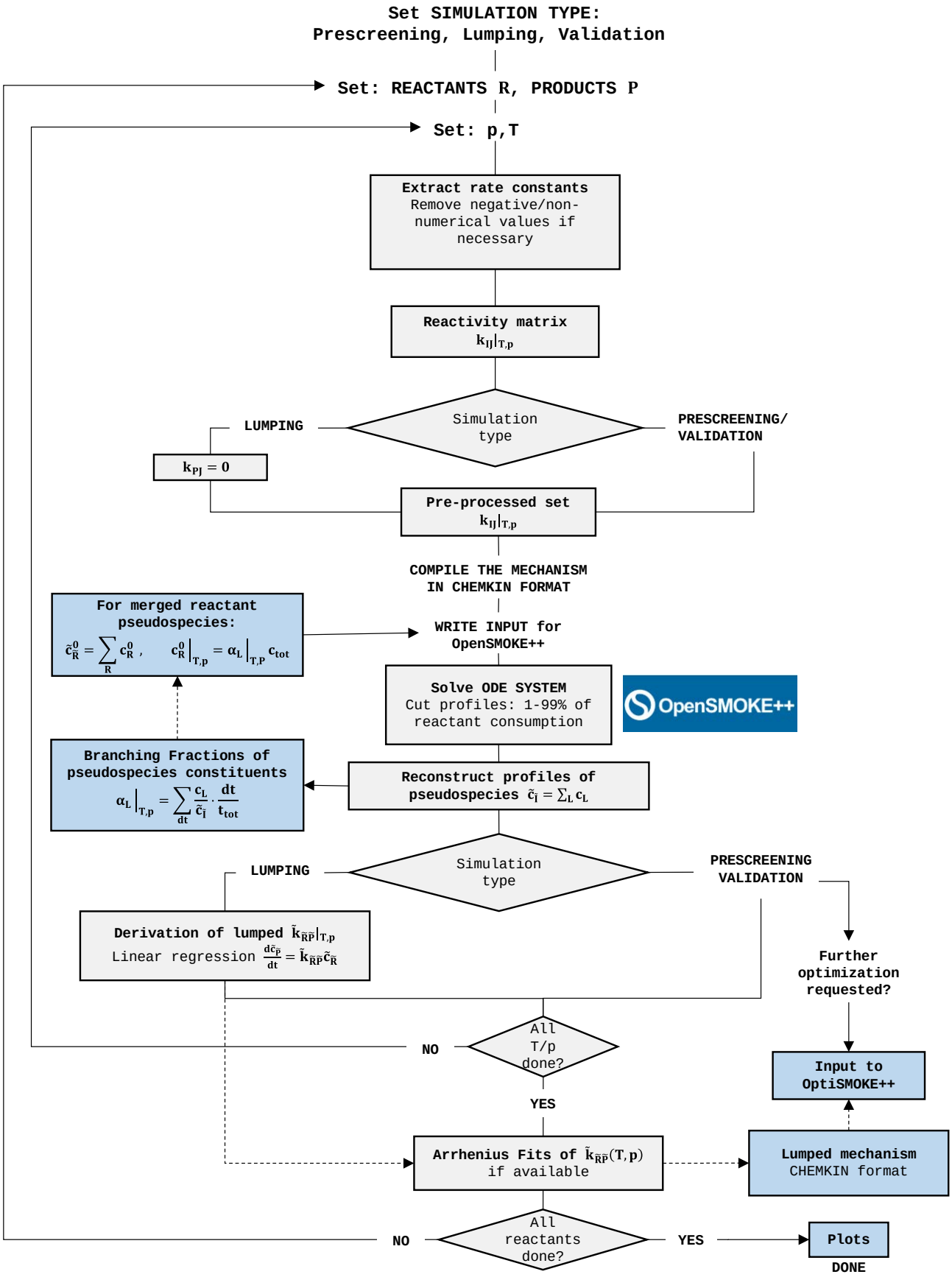

Figure 4: Flowchart of the algorithm used for the implementation of MEL. Blue boxes represent code outputs; dashed arrows indicate connections between portions of the code generated at different steps.

2.4 Species selection

Chemical structure and reactivity guide the selection of pseudospecies, as suggested by Ranzi et al. [21] and Pepiot-Desjardins and Pitsch [28]. Figure 5 summarizes the steps for the definition of accumulating species to include in the lumped mechanism (Figure 5, left) and merged pseudospecies composition (Figure 5, right).

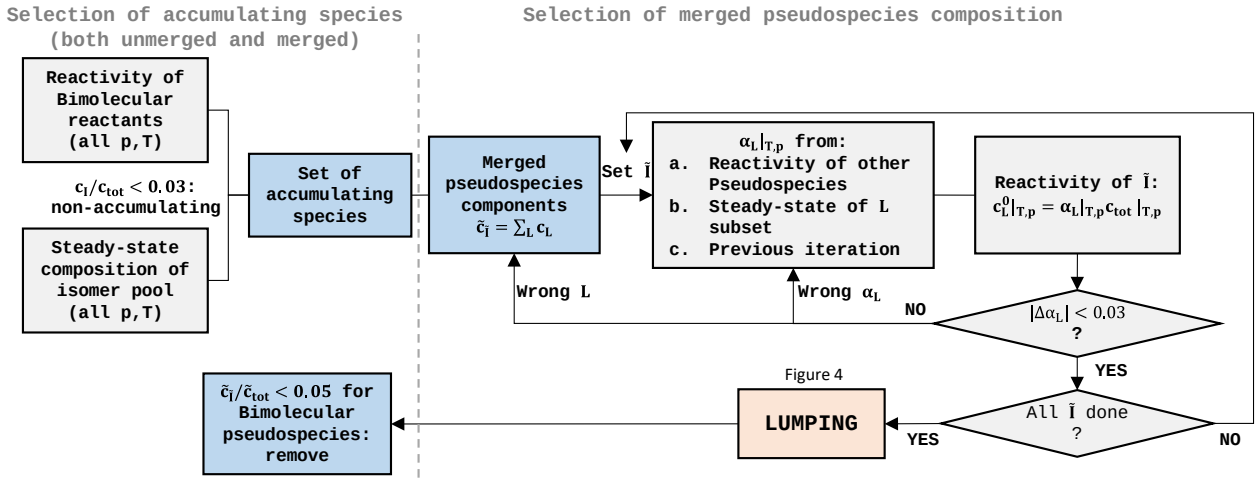

Figure 5: Diagram of the procedure of selection of accumulating species and merged pseudospecies composition. White boxes indicate the steps which require the use of the lumping code. The "Lumping" box contains all the steps illustrated in the flowchart of Figure 4.

2.4.1 Accumulating species

The elimination of non-accumulating species is subject to the constraint that no other accumulation source should be present in the target kinetic scheme. Bimolecular products are always included as pseudospecies in preliminary simulations, however they may be removed from the lumped kinetic mechanism when their production from all the other pseudospecies is negligible (e.g. $\tilde{c}_{\tilde{I}}/\tilde{c}_{tot} < 0.05$ at all investigated operating conditions). An example is provided in Section 3.2.

The investigation of the accumulation of wells is performed in two steps. First, their production from bimolecular reactions is analyzed. Then, the steady-state composition of the set of wells is obtained with the detailed ME mechanism including only isomerization reactions. A well with negligible accumulation at all T,p (in this work: $c_I/c_{tot} < 0.03$) is considered non-accumulating and is therefore not included in any

pseudospecies. The lumping of the PES of CH_3COOH decomposition provides a straightforward example of non-accumulating species elimination (Section 3.1).

Finally, it is noted that a non-accumulating species might be produced by reactions of the target kinetic scheme occurring on a different PES. In these cases, this species may be included among pseudospecies independently of its degree of accumulation. We then envision two possible limit scenarios: if its unimolecular decomposition is extremely fast, it may be appropriate to simply replace the given species with its decomposition products. On the other hand, if its unimolecular reactivity is slow compared to secondary bimolecular reactions, the given species may be safely excluded from the pseudospecies of the selected PES without affecting the macroscopic kinetic behavior of the target scheme.

2.4.2 Pseudospecies composition

Accumulating species with similar chemical structure (i.e. functionality) and reactivity should be merged into lumped groups of isomers (i.e. merged pseudospecies) according to a given composition $\alpha_L|_{T,p}$ [21,28]. The derivation of $\tilde{k}'_{\tilde{I}\tilde{J}}|_{T,p}$ for pseudospecies $\tilde{I}$ constituted of a set of accumulating species L with $\tilde{c}_{\tilde{I}} = \sum_L c_L$ requires $c_L^0|_{T,p} = \alpha_L|_{T,p} c_{tot}|_{T,p}$ for the solution of (3). When isomerizations between L are fast, α_L may be set according to the equilibrium composition of the lumped group. When this is not the case, or the equilibrium composition is unavailable, $\alpha_L|_{T,p}$ is calculated with Equation (10), from the reactivity of the pseudospecies contributing the most to the formation of the selected lumped group. An unphysical choice of the components of $\tilde{I}$ results in non-exponential decays of L (and consequently of $\tilde{I}$), easily identified by looking at the simulation profiles. Similarly, inconsistent composition $c_L^0|_{T,p} = \alpha_L|_{T,p} c_{tot}|_{T,p}$ causes abrupt changes in $c_L(t)$ due to the different characteristic times of the reactivity of L. Wei and Kuo [24,25] suggested that, for unimolecular reactions approaching equilibrium, a correct choice of α_L produces negligible changes in the pseudospecies composition $\alpha_L|_{T,p}(t)$. In MEL, this condition is met by optimizing $\alpha_L|_{T,p}$, selected as described above. In particular, iterative simulations of the reactivity of $\tilde{I}$ using the full ME mechanism are performed until the absolute change in the average $\alpha_L|_{T,p}$ (Equation (10)) with respect to the initial one $\Delta\alpha_L|_{T,p} = \left|\alpha_L - \alpha_L^0\right|_{T,p}$ is below a threshold, here set to 0.03.

Finally, it is recalled that when the components of pseudospecies $\tilde{I}$ are already present in the target kinetic scheme, relevant interactions with the mixture may occur. In these cases, $\alpha_L|_{T,p}$ requires further evaluation according to the global target kinetic scheme, as described in previous works [27,28,30]. In addition, it is possible that a component of a pseudospecies is already present as an independent species in the target kinetic scheme. Consistency between the lumped PES and the target scheme may be ensured in two ways: the selected species may be either considered independent and therefore forced to constitute an unmerged pseudospecies, or it should be lumped also in all the reactions of the target kinetic scheme. In the latter case, the secondary reactivity of the species must be rescaled according to appropriate branching fractions, i.e. the $\alpha_L|_{T,p}$.

3. Results and Discussion

3.1 CH_3COOH

The simplest MEL application of this work consists in the removal of one non-accumulating species in the PES of acetic acid (CH_3COOH) decomposition studied by Cavallotti et al. [31] (Figure 6). ME simulations cover temperature and pressure ranges of 800-2400 K and 0.1-100 atm and are reported in the supplementary material (SM, in *CH₃COOH/base_case/data*). A detailed description of the contents of SM folders is provided in Section S1.

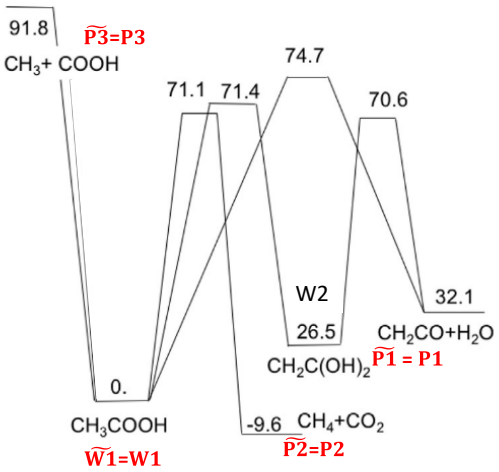

Figure 6: PES of CH_3COOH decomposition reproduced from Cavallotti et al. [31] Energies are in kcal/mol and are inclusive of zero point energies. Bold red names indicate accumulating species, and names with ~ indicate pseudospecies.

The selection of accumulating species starts with prescreening simulations of the reactivity of bimolecular reactions of $\mathrm{P1}, \mathrm{P2}, \mathrm{P3}$ including the full detailed ME mechanism (i.e. without infinite sink species). At every operating condition, the mole fraction of intermediate $\mathrm{W2}$ is below 1%. The same result is obtained with steady-state simulations including only the isomerization reactions between $\mathrm{W1}$ and $\mathrm{W2}$. Since $\mathrm{W2}$ does not accumulate in the system, it is not included among pseudospecies. In this case, pseudospecies $\widetilde{\mathrm{W1}}, \widetilde{\mathrm{P1}}, \widetilde{\mathrm{P2}}, \widetilde{\mathrm{P3}}$ correspond to single (unmerged) species of the detailed ME mechanism, as indicated in Figure 6.

Lumped rate constants are derived as described in Section 2.3 (corresponding plots are provided in the SM: *CH3COOH/base_case/plots_lumping*). All reactant decay profiles are pseudo-exponential, as expected. Variations of cutoff parameters initially set to 1-99% of reactant consumption (see Section 2.3) do not significantly affect $\tilde{\mathrm{k}}_{\tilde{\mathrm{I}}\tilde{\mathrm{J}}}$. In fact, maximum discrepancies of 7% for $\tilde{\mathrm{k}}_{\tilde{\mathrm{I}}\tilde{\mathrm{J}}}|_{1\mathrm{atm}}$ obtained setting the cutoff limits to 0.1-99% or 10-99% highlight only limited formation of $\mathrm{W2}$ at the early stages of the reactivity. Reduction of the lower cutoff limit to 90% resulted instead in differences in $\tilde{\mathrm{k}}_{\tilde{\mathrm{I}}\tilde{\mathrm{J}}}|_{1\mathrm{atm}}$ below 1%. The Arrhenius parameters of $\tilde{\mathrm{k}}_{\tilde{\mathrm{I}}\tilde{\mathrm{J}}}$ obtained in these tests are found in the SM (*CH3COOH/base_case/sensitivity_to_cutoff*).

Table 1 reports Arrhenius parameters for the detailed and lumped rate constants at 1 atm, highlighting the removal of 8 k_{IJ} associated with W2. $\tilde{\mathrm{k}}_{\tilde{\mathrm{I}}\tilde{\mathrm{J}}}$ are validated comparing the performances of the full detailed and lumped mechanisms, thus allowing for non-exponential behavior, accumulation and consumption of pseudospecies.

Table 1: Arrhenius parameters for $k = k_0 T^{\beta}\, exp\left(-\frac{E_A}{RT}\right)$ at 1 atm of the detailed (k_{IJ}) and lumped ($\tilde{k}_{\tilde{I}\tilde{J}}$) mechanisms of CH_3COOH decomposition. The fitting range is 800-2400 K, except for reactions with * (800-2100 K), removed in the lumped mechanism. Arrhenius fits for $\tilde{k}_{\tilde{I}\tilde{J}}$ at different pressures are provided in the SM (CH3COOH/base_case/kin_opt.txt). All R^2 of the fits are higher than 0.99. Units are cm, cal, mol, s.

	Detailed $\mathbf{k_{IJ}}$				**Lumped $\tilde{\mathbf{k}}_{\tilde{\mathbf{I}}\tilde{\mathbf{J}}}$**		
Reaction	$\mathbf{k_0}$	$\boldsymbol{\beta}$	$\mathbf{E_A}$	**Reaction**	$\mathbf{k_0}$	$\boldsymbol{\beta}$	$\mathbf{E_A}$
$\mathbf{W1 \rightarrow W2^*}$	9.88E+46	-10.01	88277.32				
$\mathbf{W1 \rightarrow P1 + H_2O}$	3.12E+46	-9.37	95500.23	$\mathbf{\widetilde{W1} \rightarrow \widetilde{P1} + \widetilde{H_2O}}$	5.00E+41	-8.10	89740.95
$\mathbf{W1 \rightarrow P2 + CO_2}$	8.00E+40	-7.94	87865.73	$\mathbf{\widetilde{W1} \rightarrow \widetilde{P2} + \widetilde{CO_2}}$	1.04E+41	-7.97	87960.13
$\mathbf{W1 \rightarrow P3 + CH_3}$	2.53E+59	-12.76	115873.39	$\mathbf{\widetilde{W1} \rightarrow \widetilde{P3} + \widetilde{CH_3}}$	3.50E+59	-12.80	115973.83

$\mathbf{W2 \rightarrow W1}^{*}$	7.08E+45	-9.84	59774.47				
$\mathbf{W2 \rightarrow P1 + H_2O}^{*}$	7.09E+39	-8.07	56345.70				
$\mathbf{W2 \rightarrow P2 + CO_2}^{*}$	6.69E+38	-7.74	67299.99				
$\mathbf{W2 \rightarrow P3 + CH_3}$	2.31E+45	-9.39	80723.22				
$\mathbf{P1 + H_2O \rightarrow W1}$	3.79E+34	-6.59	58074.71	$\mathbf{\widetilde{P1} + \widetilde{H_2O} \rightarrow \widetilde{W1}}$	5.96E+29	-5.32	52276.85
$\mathbf{P1 + H_2O \rightarrow W2}^{*}$	3.46E+32	-6.46	49535.57				
$\mathbf{P1 + H_2O \rightarrow P2 + CO_2}$	4.59E+24	-3.39	63711.62	$\mathbf{\widetilde{P1} + \widetilde{H_2O} \rightarrow \widetilde{P2} + \widetilde{CO_2}}$	7.12E+22	-2.91	61854.17
$\mathbf{P1 + H_2O \rightarrow P3 + CH_3}$	7.17E+23	-2.80	70897.11	$\mathbf{\widetilde{P1} + \widetilde{H_2O} \rightarrow \widetilde{P3} + \widetilde{CH_3}}$	4.21E+23	-2.74	70645.50
$\mathbf{P2 + CO_2 \rightarrow W1}$	1.82E+32	-5.52	94521.00	$\mathbf{\widetilde{P2} + \widetilde{CO_2} \rightarrow \widetilde{W1}}$	2.09E+32	-5.53	94570.55
$\mathbf{P2 + CO_2 \rightarrow W2}^{*}$	2.50E+34	-6.36	104333.57				
$\mathbf{P2 + CO_2 \rightarrow P1 + H_2O}$	7.43E+27	-3.72	107760.54	$\mathbf{\widetilde{P2} + \widetilde{CO_2} \rightarrow \widetilde{P1} + \widetilde{H_2O}}$	1.01E+26	-3.22	105853.00
$\mathbf{P2 \rightarrow P3 + CH_3}$	1.32E+28	-3.51	115077.65	$\mathbf{\widetilde{P2} + \widetilde{CO_2} \rightarrow \widetilde{P3} + \widetilde{CH_3}}$	1.27E+28	-3.50	115071.80
$\mathbf{P3 + CH_3 \rightarrow W1}$	2.05E+47	-10.21	19107.45	$\mathbf{\widetilde{P3} + \widetilde{CH_3} \rightarrow \widetilde{W1}}$	2.01E+47	-10.21	18927.41
$\mathbf{P3 + CH_3 \rightarrow W2}^{*}$	5.26E+37	-7.95	14494.21				
$\mathbf{P3 + CH_3 \rightarrow P1 + H_2O}$	4.40E+23	-3.01	11552.14	$\mathbf{\widetilde{P3} + \widetilde{CH_3} \rightarrow \widetilde{P1} + \widetilde{H_2O}}$	4.10E+23	-3.00	11534.02
$\mathbf{P3 + CH_3 \rightarrow P2 + CO_2}$	4.24E+24	-3.37	11636.73	$\mathbf{\widetilde{P3} + \widetilde{CH_3} \rightarrow \widetilde{P2} + \widetilde{CO_2}}$	4.63E+24	-3.38	11667.04

$\tilde{k}_{\tilde{I}\tilde{J}}$ were further optimized with OptiSMOKE++ [44,45] on this set of simulations, slightly increasing the accuracy of the predictions. The optimization of the lumped mechanism involved 96 active Arrhenius parameters (pre-exponential factor, temperature exponent and activation energy) in 8 pressure-dependent reactions, using the profiles generated with the full detailed mechanism as targets. The lumped rate constants were allowed to vary by factor of 2 at most. Figure 7 shows profiles of $\widetilde{P1}$ reactivity at 10 atm, where accumulation and consumption of $\widetilde{W1}$ at intermediate temperatures (800-1000 K) and of $\widetilde{P3}$ at higher temperatures (1600-2400 K) are well reproduced by the model.

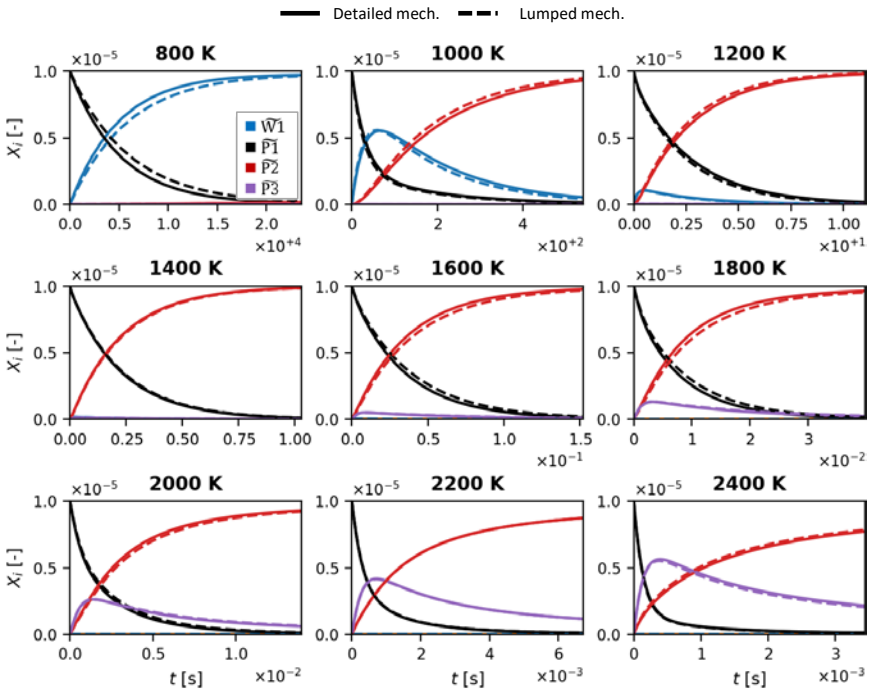

Figure 7: Profiles obtained from the reactivity of $\widetilde{P1}$ at 10 atm including the full detailed (solid lines) and lumped (dashed lines) mechanisms.

Finally, it is instructive to illustrate the simple case where W1 and W2 are lumped together in case of higher thermodynamic stability of W2. Relevant accumulation of W2 was fictitiously induced by setting its energy to 6 kcal/mol (i.e. 20 kcal/mol more stable than the calculated value [31]) in ME simulations (Figure 8a). Consequently, α_{W2} in the steady-state composition of the isomer mixture reaches up to 40% at 2400 K (solid line in Figure 8b), hence W2 must be included in the lumped mechanism.

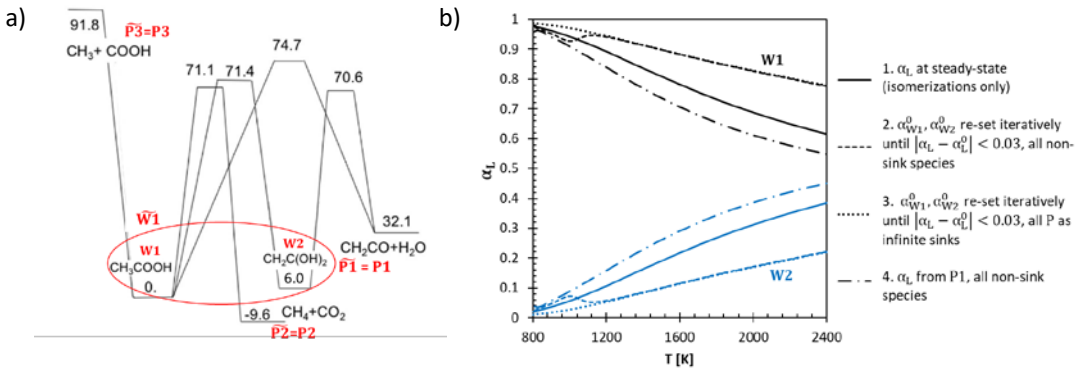

Figure 8: a) PES of CH_3COOH decomposition modified with fictitious stabilization of $W2$ to 6 kcal/mol and b) resulting composition of the pseudospecies $W1 + W2$ at 1 atm. Red names indicate accumulating species, and names with ~ indicate pseudospecies.

Lumping together W1 and W2 ($\tilde{c}_{\widetilde{W}_1} = c_{W1} + c_{W2}$) leads to a reduction of the number of species and reactions equal to the former case. Similar reactivity and product branching fractions obtained analyzing separately the reactivity of W1 and W2 support this choice. As far as the composition of the isomer mixture is concerned, Figure 8b highlights significant variations of α_{W1}, α_{W2} with the type of simulation and reactant composition. Considering the results of lumping according to α_L of cases 1 and 2 in Figure 8b (SM folders: *CH3COOH/stable_W2/plots_validation/Case1* and */Case2*), only $\tilde{k}_{\widetilde{W1}\,\widetilde{P1}}$ and $\tilde{k}_{\widetilde{W1}\,\widetilde{P2}}$ show discrepancies between the two methodologies above 10%, being as high as 33% and 18% at 2400 K, respectively. Notwithstanding these differences, the performances of the models derived according to the two cases are extremely similar, as shown for the profiles of $\widetilde{W1}$ in Figure 9a. Figure 9b highlights that the reactivity of $\widetilde{P1}$ is extremely close to that of the lumped mechanism of Figure 7, however in this case the accumulation of $\widetilde{W1}$ at intermediate temperatures involves a mixture of W1 and W2. It is remarkable that despite the discrepancies in α_L, the profiles simulated with the lumped mechanism well reproduce those of the detailed kinetic scheme even without further optimization.

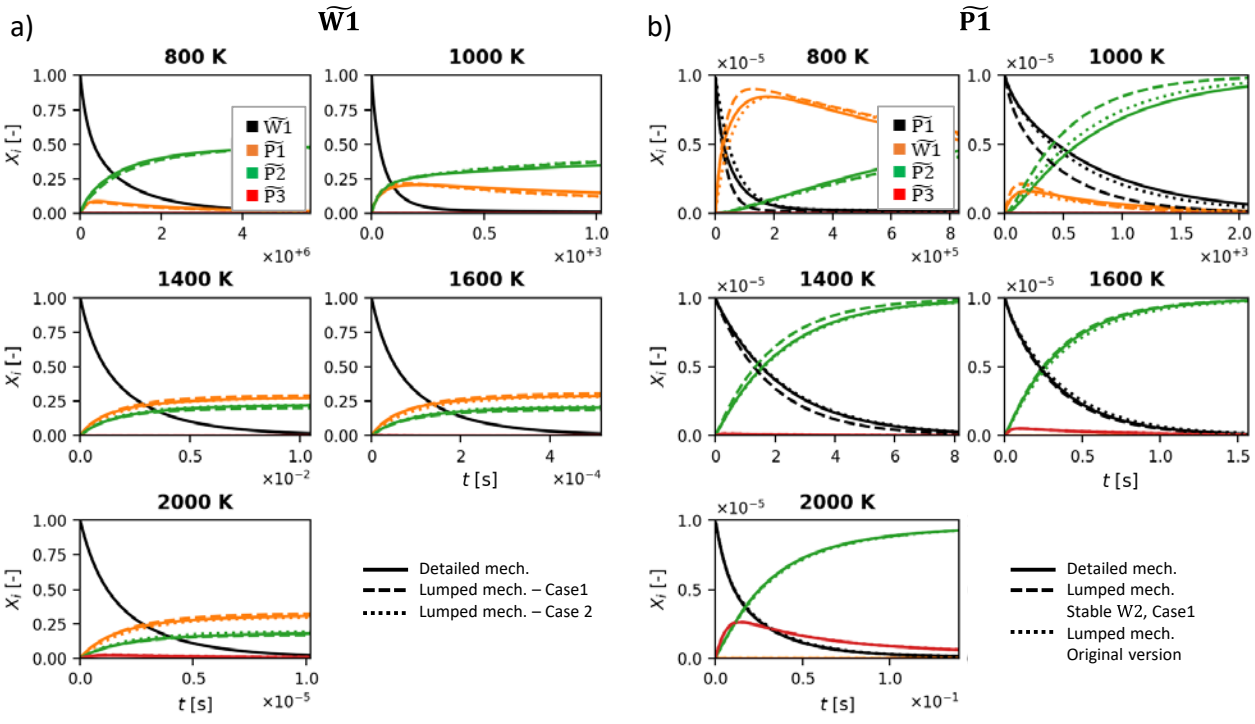

Figure 9: 1 atm profiles obtained with the detailed mechanism with fictitious stabilization of $W2$ and a) with the lumped mechanisms derived setting different α_L^0 for the derivation of rate constants (Case1 and Case2 of Figure 8), b) with the lumped mechanism of Case1 and the original lumped mechanism not including $W2$ stabilization.

3.2 C_5H_5+OH

The application of MEL to C_5H_5OH PES studied by Galimova et al. [32] allows both the extension of the stability range of intermediate wells and the elimination of oscillatory behavior of the associated rate constants, which constitute common major challenges in the attempt of including rate constants derived from first principle calculations into global kinetic schemes. All species and the main reaction channels are depicted in Figure 10 (full PES reproduced in Figure S1). OH addition to C_5H_5 is followed by isomerization reactions between 6 different isomers ($W1-6$). C-H and O-H bond fissions lead to H elimination ($P2, P3, P5, P6$), whereas C-C bond fissions produce either HCO ($P4$) or CO ($P1$). The original paper recommends extra care in including the proposed rate constants into global kinetic schemes, due to the narrow range of thermal stability of wells: at 0.01 atm, $W1$ and $W4$ exist thermodynamically and kinetically only below 1000 K, whereas k_{IJ} for $W5$ are available up to 1900 K. Furthermore, rate constants of the production of such isomers (e.g. $k_{RW5}, k_{RW4}, k_{RW1}$) show significant oscillations. As opposed to Galimova et al. [32], who discussed only relevant reactions starting from OH+C_5H_5, the lumped mechanism of this work implicitly considers all the pathways of the PES.

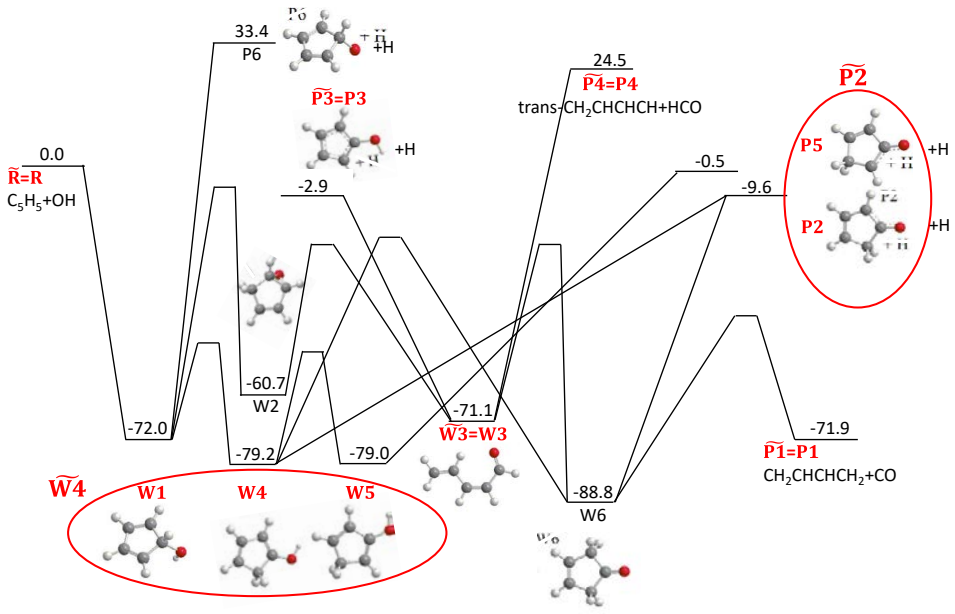

Figure 10: Simplified C_5H_5OH PES from the study of Galimova et al. [32] Energies are in kcal/mol and inclusive of ZPE. Red names in bold indicate accumulating species, and names with ~ indicate pseudospecies.

Species selection required several iterations. The analysis of the reactivity of bimolecular reactants $\mathrm{P1-6}$ shows that $\mathrm{W2}$ does not accumulate significantly in the system and is therefore considered non-accumulating. As suggested by the reaction channels depicted in Figure 10, $\mathrm{W1, W4}$, and $\mathrm{W5}$ have compatible characteristic times and product BFs, therefore they are lumped together as $\widetilde{\mathrm{W4}}$: $\tilde{c}_{\widetilde{W4}} = c_{W1} + c_{W4} + c_{W5}$. On the contrary, $\mathrm{W6}$ has different chemical functionality, hence it is not included into $\widetilde{\mathrm{W4}}$. Since the accumulation of $\mathrm{W6}$ above 600 K is significant only at 100 atm, we decided to consider 600-2500 K range and we excluded $\mathrm{W6}$ from pseudospecies. Finally, $\mathrm{W3}$ constitutes a single unmerged pseudospecies $\widetilde{\mathrm{W3}}$, being the only non-cyclic isomer, and it is produced in relevant amounts only by $\mathrm{P4}$.

Lumping together $\mathrm{W1, W4}$ and $\mathrm{W5}$ allows extending the temperature range of validity of $\tilde{k}_{\tilde{I}\tilde{J}}$ with respect to the original k_{IJ}. In fact, k_{IJ} for $\mathrm{W1}$ and $\mathrm{W4}$ are unavailable above 1000 K at low pressures, hence $\widetilde{\mathrm{W4}}$ is constituted only of $\mathrm{W5}$ between 1000 and 1900 K. The composition of $\widetilde{\mathrm{W4}}$ was determined iteratively, as explained in Section 2.4.2, reaching convergence in less than 10 iterations. The final α_L are provided in the SM (folder *C5H5OH/branchings*). The same method was used for the derivation of the composition of $\widetilde{\mathrm{P2}}$, constituted of $\mathrm{P2}$ and $\mathrm{P5}$, consistently with the corresponding $\widetilde{\mathrm{W4}}$. $\widetilde{\mathrm{R}}, \widetilde{\mathrm{P1}}, \widetilde{\mathrm{P3}}$ and $\widetilde{\mathrm{P4}}$ are instead single unmerged pseudospecies. Finally, $\mathrm{P6}$ was excluded from the pseudospecies. In fact, its production is negligible from all species of this PES, and the study of C_5H_5+O PES of Ghildina et al. [46] highlights the instability of $\mathrm{P6}$ with respect to $\mathrm{P2}$ and $\mathrm{P5}$.

MEL finally leads from 13 species and 156 k_{IJ} to a lumped mechanism of 7 pseudospecies and 42 $\tilde{k}_{\tilde{I}\tilde{J}}$, further reduced to 27 by elimination of reaction channels of negligible importance (i.e. product BF <1%). The final optimized mechanism reproduces well the results obtained with the detailed scheme, as shown by the plots in the SM (*C5H5OH/plots_validation* and *plots_validation_opt*). The lower accuracy of the lumped mechanism for $\widetilde{\mathrm{W4}}$ and $\widetilde{\mathrm{P1}}$ reactivity at 600 K and 100 atm derives from the exclusion of W6 from pseudospecies. However, we believe this is an acceptable approximation in light of the good performances of the lumped mechanism at most of the operating conditions investigated. In addition, according to the original paper [32] the reactivity of $\mathrm{P1}$ and $\mathrm{P4}$ is irrelevant to the overall C_5H_6 oxidation kinetics, hence the reactivity of $\mathrm{P1}$, $\mathrm{P4}$, and their main product $\mathrm{W3}$ may be neglected. The mechanism thus obtained consists of

6 pseudospecies (W3 is excluded) and 27 reactions. Both versions of the lumped mechanism are reported in the SM (*C5H5OH folder*) in CHEMKIN format.

The most interesting point of MEL application in this case is the removal of oscillatory behavior of some rate constants of the detailed ME mechanism minimizing the risk of numerical issues. In particular, we depict $\tilde{k}_{\tilde{R}\,\widetilde{W4}}, k_{R\,W1}, k_{R\,W4}, k_{R\,W5}$ in Figure 11.

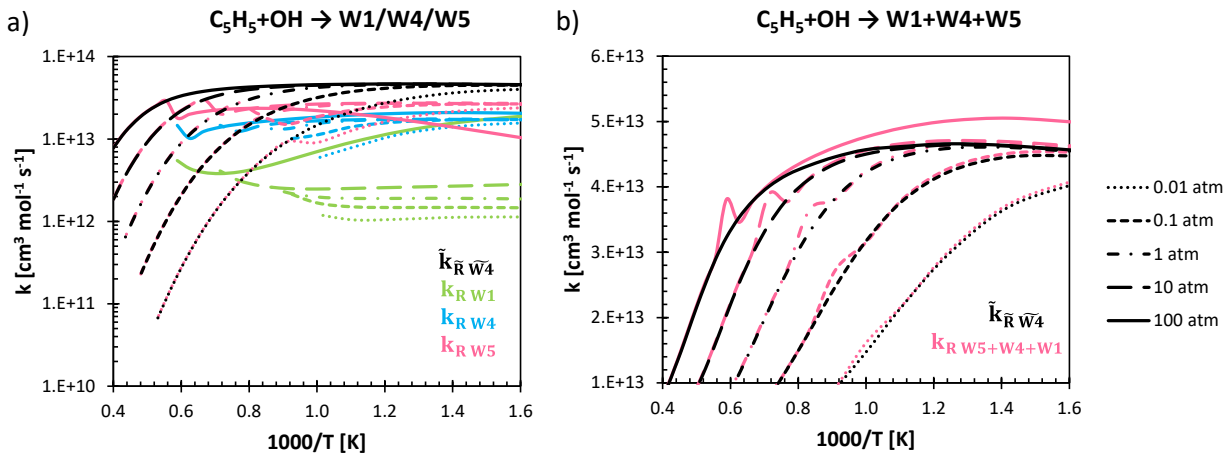

Figure 11: Comparison between the lumped rate constant $\tilde{k}_{\tilde{R}\,\widetilde{W4}}$ and the elementary rate constants $k_{R\,W1}, k_{R\,W4}, k_{R\,W5}$ of MESS output as a function of pressure. a) shows the single elementary rate constants and b) compares the lumped rate $\tilde{k}_{\tilde{R}\,\widetilde{W4}}$ with the sum of the single elementary rates $k_{R\,W1} + k_{R\,W4} + k_{R\,W5}$.

At 0.01 atm, $\tilde{k}_{\tilde{R}\,\widetilde{W4}}$ corresponds approximately to the sum of the rate constants of the elementary reaction channels (Figure 11b). As the pressure increases, $\tilde{k}_{\tilde{R}\,\widetilde{W4}}$ is higher than $k_{R\,W4} + k_{R\,W5}$, and lower than $k_{R\,W1} + k_{R\,W4} + k_{R\,W5}$, showing no oscillations thanks to the implicit inclusion of the reactivity of other intermediate channels. At higher temperatures where intermediates W1 and W4 are kinetically and thermodynamically unstable, $\tilde{k}_{\tilde{R}\,\widetilde{W4}}$ becomes equal to $k_{R\,W5}$, as expected. Additional comparisons are found in Figure S2 of the SM. The removal of the observed oscillatory behavior may also be obtained with the recently introduced WellExtension keyword in mess input, which also allows extending the wells thermodynamic stability to the full range of temperatures. This produces a considerably different composition of the isomer pool $W_1 + W_4 + W_5$, especially at higher temperatures. The reason for this behavior is that in the original simulations, at high temperatures W_5 is actually a combination of $W_1 + W_4 + W_5$, as explained more extensively in Section S2 of the SM. Therefore, the resulting lumped rate constants

$\tilde{k}_{\tilde{R}\,\widetilde{W4}}$ and $\tilde{k}_{\widetilde{W4}\,\tilde{P}_j}$ show excellent agreement with the lumped rate constants presented above, having discrepancies below 15%. Few exceptions of differences up to 50% at higher temperatures regard minor product channels. This result further supports and validate the approach presented in this work.

Finally, it is worth mentioning the implications of including this set of reactions in a target global kinetic scheme. Secondary reactivity of newly introduced pseudospecies should be considered, such as H-atom abstraction reactions on C_5H_6O isomers ($\widetilde{W4}$) to form $\widetilde{P3}$ and $\widetilde{P2}$. The corresponding rate constants may be approximated according to the $\alpha_L$of the constituents of the pseudospecies: $\tilde{k}_{Habs,\widetilde{W4}} = \sum_L \alpha_L k_{Habs,L}$ (α_L are reported in the folder *C5H5OH/branchings* of the SM). Furthermore, unimolecular decomposition of C_5H_5O (i.e. P3 and $\widetilde{P2}$) is relevant to the oxidation kinetics of C_5H_6, as highlighted by Ghildina et al. [46]. This requires lumping C_5H_5O PES [46] keeping the consistency of species common to both mechanisms (i.e. $P2, P3, P5$), as exemplified in the next example (Section 3.3) for $C_{10}H_{10}$ and $C_{10}H_9$ PESs.

3.3 $C_{10}H_{10}$ and $C_{10}H_9$

The detailed kinetic mechanism of $C_{10}H_{10}$ and $C_{10}H_9$ PESs by Long et al. [33] is constituted of 328 reversible reactions and 31 species. MEL was applied to the two interconnected PESs, reducing the kinetic mechanism to 33 irreversible reactions involving 8 pseudospecies; further elimination of all $C_{10}H_9$ isomers, which were found to be extremely reactive, resulted in only 4 lumped reactions involving 3 pseudospecies. The detailed and lumped rate constants were then included in the CRECK kinetic model [34] together with approximate H-atom abstraction reactions for additional validation of the lumped mechanism, obtaining excellent results. Because of its illustrative purposes, we limit the study to the rates at p=1 atm.

Figure 12 shows a schematic representation of the two PESs (reproduction of the original PESs in Figures S3-5). The recombination of two cyclopentadienyl radicals produces 16 $C_{10}H_{10}$ isomers. The prototype structures of the most stable wells, i.e. fulvalanes (6 species) and azulanes (8 species), are shown in Figure 12. Overall, $C_{10}H_{10}$ PES has 23 species and its reactivity is described by 238 reversible reactions. The original work [33] proposes that H-atom abstraction reactions on $C_{10}H_{10}$ are more relevant to the formation of $C_{10}H_9$ radicals (fulvanalyls and azulanyls) than bond-fission reactions at intermediate-to-low temperatures. $C_{10}H_9$ isomers

then decompose via isomerization and beta-scission reactions to fulvalene, azulene, and naphthalene. $C_{10}H_9$ PES has 14 species (6 in common with $C_{10}H_{10}$ PES) and its detailed mechanism has 90 reversible reactions.

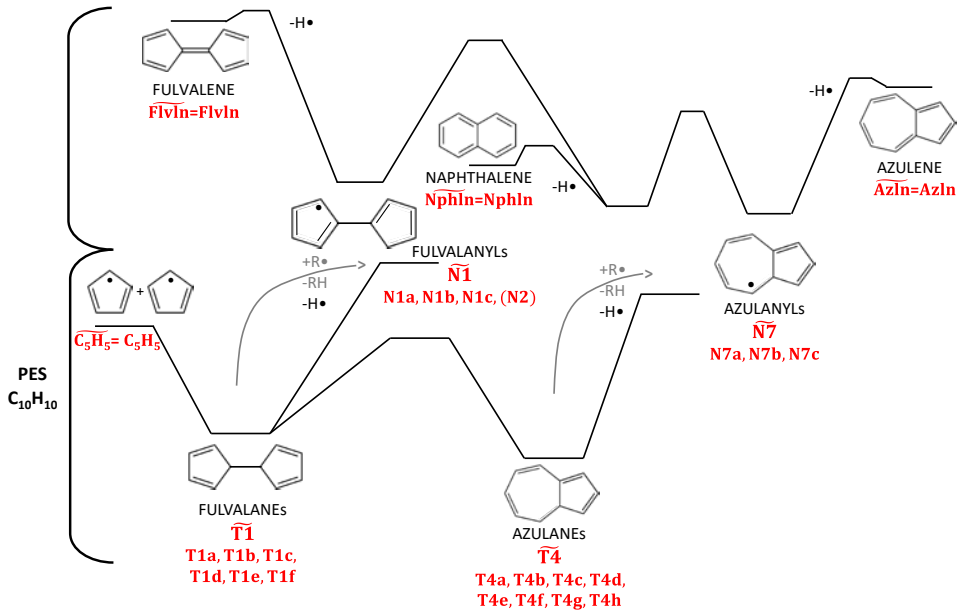

Figure 12: Scheme of the main reaction channels of the $C_{10}H_{10}$ and $C_{10}H_9$ PESs investigated by Long et al. [33]. Red names in italics indicate accumulating species, and names with ~ indicate pseudospecies.

The lumping of $C_{10}H_{10}$ PES starts from the selection of accumulating isomers. The reactivity of bimolecular reactants and the steady-state composition of the isomer pool allows eliminating three species. Cyclopentadienyl constitutes a single pseudospecies. $C_{10}H_{10}$ isomers are grouped into fulvalanes ($\widetilde{T1}$ constituted of 6 wells) and azulanes ($\widetilde{T4}$ constituted of 8 wells) according to their chemical structure. Similarly, bimolecular reactants are lumped into fulvalanyls ($\widetilde{N1}$ constituted of 3 species) and azulanyls ($\widetilde{N7}$ constituted of 3 species). Further lumping of $\widetilde{T1}$ and $\widetilde{T4}$ into a single pseudospecies is discouraged by the large differences in the associated k_{IJ} (see Fig. 5B in the paper of Long et al. [33]). The possible consequences of this incorrect choice are discussed in the SM (Figure S8-9). Concerning pseudospecies composition, steady-state composition was used as guess for α_L^0 of pseudospecies $\widetilde{T1}$ and $\widetilde{T4}$. The guess for α_L^0 of $\widetilde{N1}$ and $\widetilde{N7}$ was instead derived from the simulations of $\widetilde{T1}$ and $\widetilde{T4}$ reactivity, namely the main respective sources of the selected bimolecular products. Convergence of the final α_L (provided in folder *2C5H5/2PES/C10H10/branchings* of the SM) was reached with less than 10 iterative simulations of the full detailed mechanism for each lumped reactant. It is noted that the reactivity of $\widetilde{T1}$ depends only slightly on

the initial composition of the isomer pool α_L^0. In fact, isomerization reactions within $\widetilde{T1}$ are significantly faster than decomposition reactions, such that $\widetilde{T1}$ reaches its steady-state composition much earlier than the initial cutoff time marking 1% of reactant consumption. MEL reduces the mechanism for $C_{10}H_{10}$ PES from 23 species and 238 reversible reactions to 5 pseudospecies and 23 forward reaction channels. The plots of Figure 13 show the excellent results of MEL application to this case: simulation profiles obtained with the detailed and lumped kinetic models are almost superimposed.

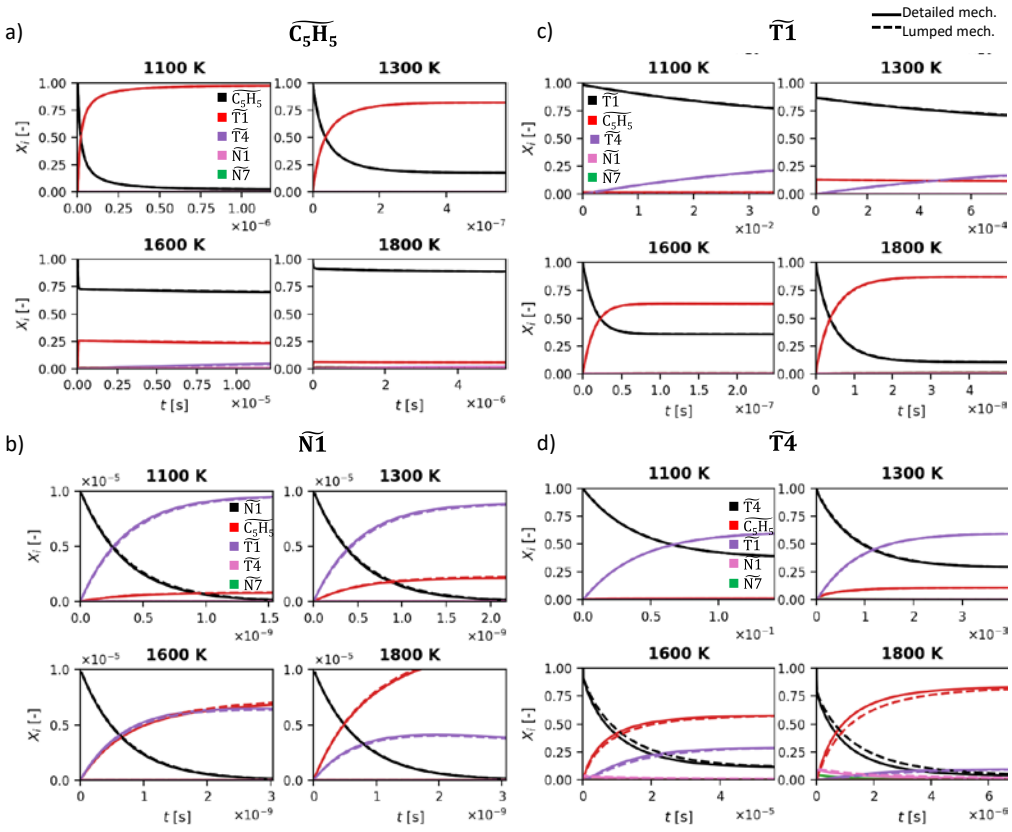

Figure 13: Comparison between the performances of the detailed (solid lines) and lumped (dashed lines) mechanisms in the 1000-1900 K range of the reactivity of pseudospecies $\widetilde{C_5H_5}$ a), $\widetilde{N1}$ b), $\widetilde{T1}$ c), and $\widetilde{T4}$ d). Complete plots are provided in the SM (2C5H5/2PES/C10H10/plots_validation).

The lumping of $C_{10}H_9$ PES was performed independently of $C_{10}H_{10}$ PES. Nevertheless, fulvalanyls and azulanyls pseudospecies ($\widetilde{N1}$and $\widetilde{N7}$) are common to both PESs, hence their lumping should be consistent. Preliminary simulations determined four non-accumulating isomers (see Figure S7 of the SM). Fulvalene, azulene and naphthalene were included as single unmerged pseudospecies. The remaining accumulating species were merged according to their chemical structure. $\widetilde{N7}$ has the same components as in $C_{10}H_{10}$ PES, and is mostly

constituted of N7b, like in the former PES. The behavior of fulvalanyl group shows instead significant discrepancies with $\widetilde{N1}$ of $C_{10}H_{10}$ PES. In fact, not only the isomer pool composition is different, but the correct reproduction of the profiles of the detailed mechanism requires the inclusion of one further species absent in the $C_{10}H_{10}$ PES (N2 in Figure 12 and Figure S7). The consistency of this choice was verified with kinetic simulations of the $C_{10}H_9$ lumped model in comparison with profiles of the detailed mechanism starting from $\widetilde{N1}$ composition derived in $C_{10}H_{10}$ PES. The performances of the two models are almost identical, as shown in the SM (*2C5H5/2PES/C10H9/plots_validation/N1N7_from _C10H10*).

Overall, MEL application to $C_{10}H_9$ PES decreases the size of the associated mechanism from 14 species and 90 reversible reactions to 5 species and 20 irreversible reactions, further reduced to 13 when neglecting minor channels. Figure 14 reports comparisons between the detailed and lumped model performances for some pseudospecies: the profiles obtained are almost indistinguishable, supporting the accuracy of the proposed approach.

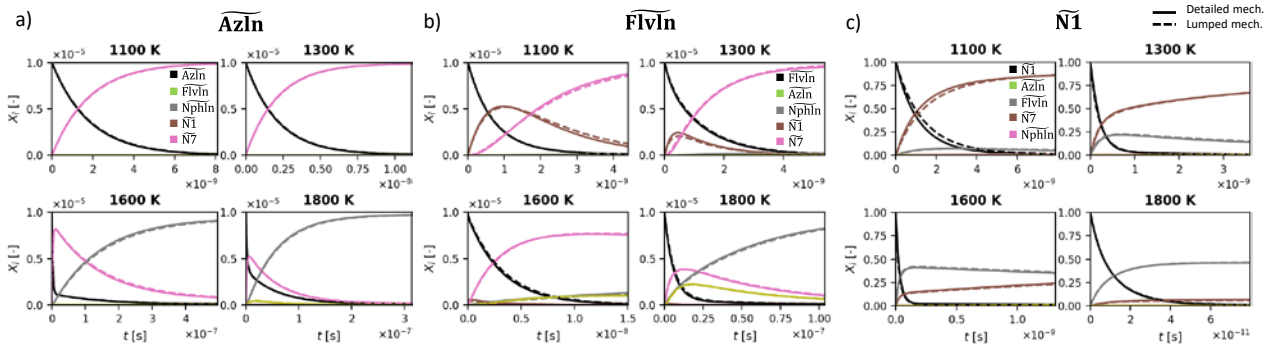

Figure 14: Comparison between the performance of the detailed (solid lines) and lumped (dashed lines) mechanisms in the 1100-1800 K range of the reactivity of pseudospecies $\widetilde{Azln}$ a), $\widetilde{Flvln}$ b), $\widetilde{N1}$ c), considering the full reactivity (no infinite sink species). Pseudospecies composition and the complete plots are provided in the SM (2C5H5/2PES/C10H9/branchings and plots_validation).

Finally, the new rate constants were integrated into the CRECK kinetic model [34]. Both cyclopentadienyl (C_5H_5) and naphthalene ($C_{10}H_8$) were already present in the target kinetic scheme, whereas the remaining 6 pseudospecies were added as $FC_{10}H_{10}$ ($\widetilde{T1}$), $AC_{10}H_{10}$ ($\widetilde{T4}$), $FC_{10}H_9$ ($\widetilde{N1}$) and $AC_{10}H_9$ ($\widetilde{N7}$), Flvln ($\widetilde{Flvln}$) and Azln ($\widetilde{Azln}$). Their thermodynamics corresponds to that of the most stable isomer constituting the pseudospecies. H-atom abstraction reactions on $FC_{10}H_{10}$ and $AC_{10}H_{10}$ by the main radicals were introduced using analogy rules from cyclopentadiene. We then simulated available experimental data of cyclopentadiene pyrolysis

[47–49] with the target model updated with both the detailed kinetic mechanism and its lumped version. $AC_{10}H_{10}$, all $C_{10}H_9$ isomers, azulene and fulvalene do not accumulate in relevant fractions in any of the cases tested, highlighting fast decomposition of $C_{10}H_{10}$ to naphthalene+2H and suggesting that the $C_{10}H_9$ reactivity may be entirely neglected. Hence, a further reduction was performed: $C_{10}H_{10}$ PES was lumped considering only $FC_{10}H_{10}$ ($\widetilde{T1}$) and a single group of $C_{10}H_9$+H bimolecular reactants (details in Section S4 of the SM and in folder 2C5H5/1PES). $C_{10}H_9$+H were always considered as infinite sinks due to fast decomposition to products. Replacing $C_{10}H_9$+H with $C_{10}H_8$+2H in the final model allowed complete removal of $C_{10}H_9$ isomers. The final lumped mechanism consists of only 3 species and 4 $\tilde{k}_{\tilde{I}\tilde{J}}$, an impressive reduction with respect to the full detailed mechanism of 31 species and 656 k_{IJ}. The performances of this lumped mechanism are almost identical to the first lumped version of 8 species and 33 reactions. Figure 15 shows experimental data of C_5H_6 pyrolysis in a PFR in comparison with model performances of the original CRECK kinetic model, and of the model updated with the reactivity of $C_{10}H_{10}$/$C_{10}H_9$ PESs [33]. The simulation profiles of the lumped and detailed kinetic schemes are almost superimposed, supporting the validity of our approach. Figures S10-12 in Section S5 of the SM show the validation with other sets of experimental data.

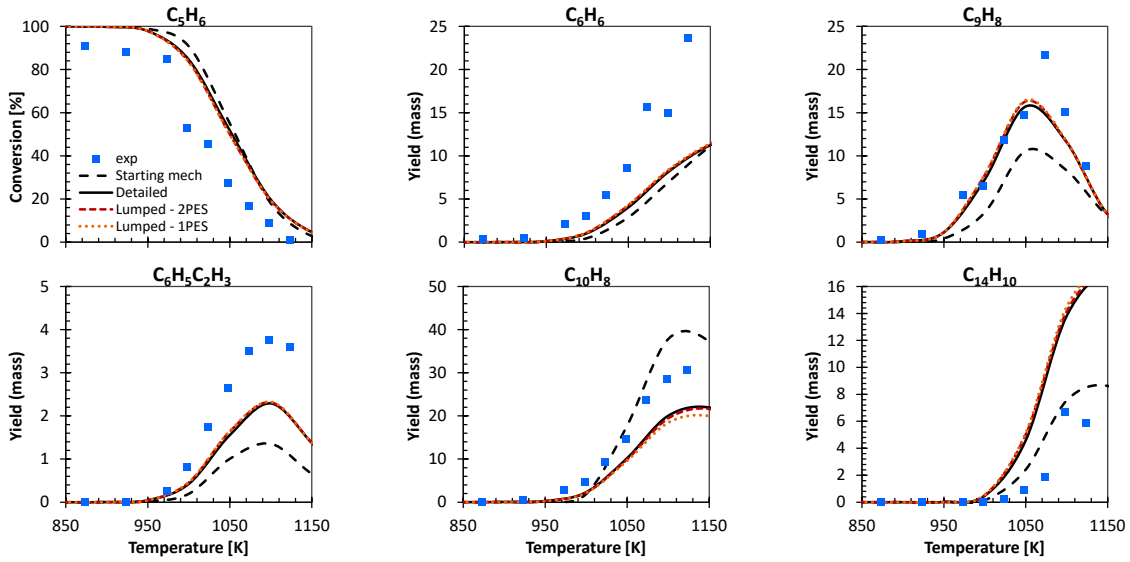

Figure 15: Experimental data of C_5H_6 pyrolysis in a plug flow reactor [47] compared to the profiles obtained from kinetic simulations of CRECK kinetic model [34], of updated detailed $C_{10}H_{10}$/$C_{10}H_9$ reactivity from Long et al. [33] (31 species, 656 reactions), and of updated lumped reactivity (2PES: 8 species, 33 reactions; 1PES: 3 species, 4 reactions).

4. Conclusions and future perspectives

In this work we introduce master-equation based lumping (MEL), the first systematic approach specifically designed to process rate constants derived from master equation simulations and integrate them into global kinetic schemes. Our methodology addresses two main issues: first, the introduction of several intermediate species often unnecessary to the macroscopic representation of the reactivity. Second, discontinuities in rate constants and number of species originating from reactions involving intermediates whose lifetime is comparable to the rovibrational relaxation timescale for ME models that use the well merging methodology, as well as dependence of rate constants on the entrance well that are found when other approaches are used.

MEL retains features of lumping methodologies for both ME simulations and global kinetic schemes. The main idea is similar to chemical lumping of unimolecular reactions: a simple ODE system of equations describing time evolution of species concentration according to a detailed mechanism is reduced to an equivalent lumped system of a smaller set of pseudospecies. Pseudospecies may be constituted of either unmerged or multiple merged species of the ME mechanism. The derivation of lumped rate constants for pseudospecies is instead inspired by the solution procedure of the non-conservative ME: simulations are run for each pseudospecies imposing that all products are non-reactive, so that the reactant decays exponentially. Differential fits of the profiles of products concentration allow obtaining lumped rate constants. The constituents of pseudospecies are selected according to their stability, chemical structure, and reactivity: highly reactive species not accumulating significantly in the system are not included in the lumped mechanism and the composition of each merged pseudospecies is derived iteratively such that it remains approximately constant in time. The final lumped mechanism may be further optimized considering kinetic simulations of the full system reactivity. All the operations listed above are performed with an in-house Python code relying on OpenSMOKE++ [43] and OptiSMOKE++ [44,45] for kinetic simulations and optional mechanism optimization, respectively.

The approach proposed has several advantages. Firstly, contrary to most chemical lumping approaches, it allows both species elimination and merging of species with similar structure and reactivity. Secondly, it does

not require any experimental data, hence the applicability range of the mechanism corresponds to that of ME simulations. Thirdly, it includes both bimolecular and unimolecular reaction channels. Future developments of MEL include full automatization of pseudospecies selection and lumping of multiple PESs. The validity of MEL is supported by three cases of application with increasing complexity. Lumping of the CH_3COOH PES shows how to remove a non-accumulating intermediate from the mechanism, still retaining implicitly its reactivity. The main challenges of including the reactivity of the C_5H_5+OH PES into a global kinetic scheme are the limited range of validity of the rate constants of the detailed mechanism, mostly related to the partial stability of the wells, and their strong non-Arrhenius behaviors. We solve these issues by lumping together isomers with similar reactivity, obtaining very similar performances of the lumped and detailed mechanisms. MEL thus extends the validity range of the resulting rate constants and eliminates their oscillatory behavior. Finally, lumping $C_{10}H_{10}$ and $C_{10}H_9$ PESs allows decreasing dramatically the initial set of 31 species and 328 reversible reactions to only 3 pseudospecies and 4 irreversible reactions, achieving excellent performances of the lumped kinetic mechanism. The integration of both the detailed and the lumped rate constants into an existing global target kinetic scheme also considered the secondary reactivity of pseudospecies. The almost identical profiles of kinetic simulations of real experimental data obtained with the detailed and lumped kinetic models fully support the validity of our approach.

MEL approach constitutes an important step towards the systematic and rigorous integration of AITSTME-derived rate constants into global kinetic schemes, effectively bridging the gap between the needs of kinetic model parameters of increasing accuracy and the necessity of producing predictive, reliable, physically meaningful, manageable and human readable kinetic models. The same methodology can be used to process existing kinetic models rather than ME outputs, for example lumping complex chemical kinetics such as those involved in low temperature combustion of alkanes, as will be reported in upcoming publications.

Supplementary Material

- SM.docx: Additional information and plots of the lumping procedure in the cases tested.

- Folders CH3COOH.zip, C5H5OH.zip, 2C5H5.zip containing simulation input profiles, plots, and kinetic mechanisms of the lumping cases analyzed in this work. Detailed description of the contents of the folder is provided in Section S1.

Acknowledgments:

Conflicts of interest: the authors have no conflicts of interest to declare.

References

[1] S.J. Klippenstein, C. Cavallotti, Ab initio kinetics for pyrolysis and combustion systems, in: Math. Model. Gas-Phase Complex React. Syst. Pyrolysis Combust., 2019: pp. 115–167.

[2] M. Frisch, G. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, Gaussian 09, revision D. 01, (2009).

[3] H.J. Werner, P. Knowles, R. Lindh, F.R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. Amosand, A. Bernhardsson, MOLPRO 2010, (2010).

[4] C.W. Gao, J.W. Allen, W.H. Green, R.H. West, Reaction Mechanism Generator: Automatic construction of chemical kinetic mechanisms, Comput. Phys. Commun. 203 (2016) 212–225. https://doi.org/10.1016/j.cpc.2016.02.013.

[5] N.M. Vandewiele, K.M. Van Geem, M.-F. Reyniers, G.B. Marin, Genesys: Kinetic model construction using chemo-informatics, Chem. Eng. J. 207–208 (2012) 526–538. https://doi.org/10.1016/j.cej.2012.07.014.

[6] R. Van De Vijver, J. Zádor, KinBot: Automated stationary point search on potential energy surfaces, Comput. Phys. Commun. 248 (2020) 106947. https://doi.org/10.17632/hsh6dvv2zj.1.

[7] C. Cavallotti, M. Pelucchi, Y. Georgievskii, S.J. Klippenstein, EStokTP: Electronic Structure to Temperature- and Pressure-Dependent Rate Constants—A Code for Automatically Predicting the Thermal Kinetics of Reactions, J. Chem. Theory Comput. 15 (2019) 1122–1145. https://doi.org/10.1021/acs.jctc.8b00701.

[8] J.A. Miller, S.J. Klippenstein, Master Equation Methods in Gas Phase Chemical Kinetics, J Phys Chem

A. 110 (2006) 10528–10544. https://doi.org/10.1021/jp062693x.

[9] Y. Georgievskii, J.A. Miller, M.P. Burke, S.J. Klippenstein, Reformulation and solution of the master equation for multiple-well chemical reactions, J. Phys. Chem. A. 117 (2013) 12146–12154. https://doi.org/10.1021/jp4060704.

[10] D.R. Glowacki, C.-H. Liang, C. Morley, M.J. Pilling, S.H. Robertson, MESMER: An Open-Source Master Equation Solver for Multi-Energy Well Reactions, J. Phys. Chem. A. 116 (2012) 9545–9560. https://doi.org/10.1021/jp3051033.

[11] A.M. Mebel, A. Landera, R.I. Kaiser, Formation Mechanisms of Naphthalene and Indene: From the Interstellar Medium to Combustion Flames, J. Phys. Chem. A. 121 (2017) 901–926. https://doi.org/10.1021/acs.jpca.6b09735.

[12] N. Hansen, M. Schenk, K. Moshammer, K. Kohse-Höinghaus, Investigating repetitive reaction pathways for the formation of polycyclic aromatic hydrocarbons in combustion processes, Combust. Flame. 180 (2017) 250–261. https://doi.org/10.1016/J.COMBUSTFLAME.2016.09.013.

[13] A.M. Mebel, Y. Georgievskii, A.W. Jasper, S.J. Klippenstein, Pressure-dependent rate constants for PAH growth: formation of indene and its conversion to naphthalene, Faraday Discuss. 195 (2016) 637–670. https://doi.org/10.1039/c6fd00111d.

[14] A. Raj, M.J. Al Rashidi, S. Ho Chung, S. Mani Sarathy, PAH Growth Initiated by Propargyl Addition: Mechanism Development and Computational Kinetics, J. Phys. Chem. A. 118 (2014) 2865–2885. https://doi.org/10.1021/jp410704b.

[15] A. Raj, G.R. da Silva, S.H. Chung, Reaction mechanism for the free-edge oxidation of soot by O 2, Combust. Flame. 159 (2012) 3423–3436. https://doi.org/10.1016/j.combustflame.2012.06.004.

[16] J.R. Barker, Multiple-Well, multiple-path unimolecular reaction systems. I. MultiWell computer program suite, Int. J. Chem. Kinet. 33 (2001) 232–245. https://doi.org/10.1002/kin.1017.

[17] A. Barbato, C. Seghi, C. Cavallotti, An *ab initio* Rice-Ramsperger-Kassel-Marcus/master equation investigation of SiH4 decomposition kinetics using a kinetic Monte Carlo approach, J. Chem. Phys.

130 (2009) 074108. https://doi.org/10.1063/1.3077561.

[18] E. Ranzi, A. Frassoldati, A. Stagni, M. Pelucchi, A. Cuoci, T. Faravelli, Reduced kinetic schemes of complex reaction systems: Fossil and biomass-derived transportation fuels, 2014. https://doi.org/10.1002/kin.20867.

[19] E. Ranzi, M. Dente, S. Pierucci, G. Biardi, Initial Product Distributions from Pyrolysis of Normal and Branched Paraffins, 1983. https://pubs.acs.org/sharingguidelines (accessed June 26, 2020).

[20] E. Ranzi, T. Faravelli, P. Gaffuri, A. Sogaro, Low-temperature combustion: Automatic generation of primary oxidation reactions and lumping procedures, Combust. Flame. 102 (1995) 179–192. https://doi.org/10.1016/0010-2180(94)00253-O.

[21] E. Ranzi, M. Dente, A. Goldaniga, G. Bozzano, T. Faravelli, Lumping procedures in detailed kinetic modeling of gasification, pyrolysis, partial oxidation and combustion of hydrocarbon mixtures, Prog. Energy Combust. Sci. 27 (2001) 99–139. https://doi.org/10.1016/S0360-1285(00)00013-7.

[22] E. Ranzi, A. Frassoldati, S. Granata, T. Faravelli, Wide-Range Kinetic Modeling Study of the Pyrolysis, Partial Oxidation, and Combustion of Heavy n-Alkanes, Ind. Eng. Chem. Res. 44 (2005) 5170–5183. https://doi.org/10.1021/ie049318g.

[23] M. Chaos, A. Kazakov, Z. Zhao, F.L. Dryer, A High-Temperature Chemical Kinetic Model for Primary Reference Fuels, Int. J. Chem. Kinet. 39 (2007) 399–414. https://doi.org/10.1002/kin.

[24] J. Wei, J.C. W Kuo, A lumping analysis in monomolecular reaction systems - Analysis of Exactly Lumpable System, I&EC Fundam. 8 (1969) 114–123.

[25] J.C. W Kuo, J. Wei, A lumping analysis in monomolecular reaction systems - Analysis of Approximately Lumpable System, I&EC Fundam. 8 (1969) 124–133.

[26] H. Huang, M. Fairweather, J.F. Griffiths, A.S. Tomlin, R.B. Brad, A systematic lumping approach for the reduction of comprehensive kinetic models, Proc. Combust. Inst. 30 (2005) 1309–1316. https://doi.org/10.1016/j.proci.2004.08.001.

[27] T. Nagy, T. Turányi, Reduction of very large reaction mechanisms using methods based on simulation

error minimization, Combust. Flame. 156 (2009) 417–428. https://doi.org/10.1016/j.combustflame.2008.11.001.

[28] P. Pepiot-Desjardins, H. Pitsch, An automatic chemical lumping method for the reduction of large chemical kinetic mechanisms, Combust. Theory Model. 12 (2008) 1089–1108. https://doi.org/10.1080/13647830802245177.

[29] S.S. Ahmed, F. Mauß, G. Moreácz, T. Zeuch, A comprehensive and compact n-heptane oxidation model derived using chemical lumpingw, Phys. Chem. Chem. Phys. 9 (2006) 1107–1126. https://doi.org/10.1039/b614712g.

[30] P. Pepiot, L. Cai, H. Pitsch, Model reduction and lumping procedures, in: Comput. Aided Chem. Eng., 2019: pp. 799–827. https://doi.org/10.1016/B978-0-444-64087-1.00016-4.

[31] C. Cavallotti, M. Pelucchi, A. Frassoldati, Analysis of acetic acid gas phase reactivity: Rate constant estimation and kinetic simulations, Proc. Combust. Inst. 37 (2019) 539–546. https://doi.org/10.1016/j.proci.2018.06.137.

[32] G.R. Galimova, V.N. Azyazov, A.M. Mebel, Reaction mechanism, rate constants, and product yields for the oxidation of Cyclopentadienyl and embedded five-member ring radicals with hydroxyl, Combust. Flame. 187 (2018) 147–164. https://doi.org/10.1016/j.combustflame.2017.09.005.

[33] A.E. Long, S.S. Merchant, A.G. Vandeputte, H.-H. Carstensen, A.J. Vervust, G.B. Marin, K.M. Van Geem, W.H. Green, Pressure dependent kinetic analysis of pathways to naphthalene from cyclopentadienyl recombination, Combust. Flame. 187 (2018) 247–256. https://doi.org/10.1016/j.combustflame.2017.09.008.

[34] W. Pejpichestakul, E. Ranzi, M. Pelucchi, A. Frassoldati, A. Cuoci, A. Parente, T. Faravelli, Examination of a soot model in premixed laminar flames at fuel-rich conditions, Proc. Combust. Inst. 37 (2019) 1013–1021. https://doi.org/10.1016/j.proci.2018.06.104.

[35] B. Widom, Molecular transitions and chemical reaction rates, Science (80-.). 148 (1965) 1555–1560. https://doi.org/10.1126/science.148.3677.1555.

[36] S.J. Klippenstein, J.A. Miller, From the time-dependent, multiple-well master equation to phenomenological rate coefficients, J. Phys. Chem. A. 106 (2002) 9267–9277. https://doi.org/10.1021/jp021175t.

[37] J.A. Miller, S.J. Klippenstein, From the multiple-well master equation to phenomenological rate coefficients: Reactions on a C3H4 potential energy surface, J. Phys. Chem. A. 107 (2003) 2680–2692. https://doi.org/10.1021/jp0221082.

[38] J.A. Miller, S.J. Klippenstein, Determining phenomenological rate coefficients from a time-dependent, multiple-well master equation: "species reduction" at high temperatures, Phys. Chem. Chem. Phys. 15 (2013) 4744–4753. https://doi.org/10.1039/c3cp44337j.

[39] D. Polino, A. Barbato, C. Cavallotti, Theoretical investigation of germane and germylene decomposition kinetics, Phys. Chem. Chem. Phys. 12 (2010) 10622–10632. https://doi.org/10.1039/c002221g.

[40] J.T. Bartis, B. Widom, Stochastic models of the interconversion of three or more chemical species, J. Chem. Phys. 60 (1974) 3474–3482. https://doi.org/10.1063/1.1681562.

[41] O.K. Rice, Relation between Equilibrium Constant and Rate Constants, J. Phys. Chem. 65 (1961) 1972–1976.

[42] J.A. Miller, S.J. Klippenstein, The Recombination of Propargyl Radicals and Other Reactions on a C6H6 Potential, J Phys Chem A. 107 (2003) 7783–7799. https://doi.org/10.1021/jp030375h.

[43] A. Cuoci, A. Frassoldati, T. Faravelli, E. Ranzi, OpenSMOKE++: An object-oriented framework for the numerical modeling of reactive systems with detailed kinetic mechanisms, Comput. Phys. Commun. 192 (2015) 237–264. https://doi.org/10.1016/j.cpc.2015.02.014.

[44] M. Fürst, A. Bertolino, A. Cuoci, T. Faravelli, A. Frassoldati, A. Parente, OptiSMOKE++: A toolbox for optimization of chemical kinetic mechanisms, Comput. Phys. Commun. 264 (2021) 107940. https://doi.org/10.1016/j.cpc.2021.107940.

[45] A. Bertolino, M. Fürst, A. Stagni, A. Frassoldati, M. Pelucchi, C. Cavallotti, T. Faravelli, A. Parente, An

evolutionary, data-driven approach for mechanism optimization: theory and application to ammonia combustion, Combust. Flame. 229 (2021) 111366. https://doi.org/10.1016/j.combustflame.2021.02.012.

[46] A.R. Ghildina, A.D. Oleinikov, V.N. Azyazov, A.M. Mebel, Reaction mechanism, rate constants, and product yields for unimolecular and H-assisted decomposition of 2,4-cyclopentadienone and oxidation of cyclopentadienyl with atomic oxygen, Combust. Flame. 183 (2017) 181–193. https://doi.org/10.1016/j.combustflame.2017.05.015.

[47] D.H. Kim, J.A. Mulholland, D. Wang, A. Violi, Pyrolytic Hydrocarbon Growth from Cyclopentadiene, J. Phys. Chem. A. 114 (2010) 12411–12416.

[48] M.R. Djokic, K.M. Van Geem, C. Cavallotti, A. Frassoldati, E. Ranzi, G.B. Marin, An experimental and kinetic modeling study of cyclopentadiene pyrolysis: First growth of polycyclic aromatic hydrocarbons, Combust. Flame. 161 (2014) 2739–2751. https://doi.org/10.1016/J.COMBUSTFLAME.2014.04.013.

[49] A.J. Vervust, M.R. Djokic, S.S. Merchant, H.-H. Carstensen, A.E. Long, G.B. Marin, W.H. Green, K. Van Geem, Detailed Experimental and Kinetic Modeling Study of Cyclopentadiene Pyrolysis in the Presence of Ethene, Energy & Fuels. 32 (2018) 3920–3934. https://doi.org/10.1021/acs.energyfuels.7b03560.