Assessment of the suitability of the chemical reaction pathway algorithm as a reduction method for plasma chemistry

S Tadayon Mousavi1,*, J G M Gulpen1, W A A D Graef2, P M J Koelman1, E A D Carbone3,4 and J van Dijk1,*,

1 Applied Physics Department, Eindhoven University of Technology, Eindhoven 5600MB, The Netherlands
2 Plasma Matters B.V., Eindhoven 5612AJ, The Netherlands
3 Max Planck Institute for Plasma Physics, Boltzmannstr. 2, D-85748 Garching, Germany

E-mail: s.tadayon.mousavi@tue.nl and j.v.dijk@tue.nl

Received 13 July 2022, revised 29 September 2022
Accepted for publication 11 October 2022
Published 1 November 2022

Abstract

Determination of chemical pathways, sets of interlinked reactions, is a well-known method to study complex chemistries. In order to have a trustworthy and accurate method for analysis of chemical processes in complicated systems such as low-temperature plasmas, a used algorithm should keep conservation of necessary parameters during all steps of an analysis. In this research, the introduced algorithm by Lehmann (2004 J. Atmos. Chem. 47 45–78) is chosen as a candidate for studying tangled low-temperature plasma chemistry. As a first step, the ability of the algorithm to keep conservation of reactions rates and net production of species is assessed by designing some artificial examples. In that stage, ambiguous parts of the algorithm are clarified, and test cases are provided to verify implementation of the algorithm in any programming infrastructure. Then, the capability of the algorithm as a reduction method is investigated by applying a newly developed semi-automated method based on pathway analysis to two plasma chemistry examples, a H2O–He microwave induced plasma and a pulsed H2 plasma.

Keywords: chemical reduction method, pathway analysis, reduced H2 kinetic set, reduced H2O–He kinetic set, microwave induced plasma H2O–He, pulsed H2 plasma, Non-Equilibrium plasma chemistry reduction method

(Some figures may appear in colour only in the online journal)
1. Introduction

Nowadays, non-equilibrium plasmas become a promising field to solve some of the contemporary problems of mankind such as the demand for renewable energy and proper medical treatment. For instance, using plasmas for efficient dissociation of CO₂ to CO as a first step for the production of carbohydrate fuels [1–5], or using plasma jets to heal diabetes wounds are among the promising applications of non-equilibrium plasmas [6, 7]. Due to the non-equilibrium nature of these types of plasmas, a variety of species can be produced that react with each other. Therefore, obtaining knowledge about underlying chemical mechanisms is crucial for a proper interpretation of physical aspects of plasmas. Usually, the first step to gain insights about the evolution of various species is employing zero-dimensional models [8]. In those types of models, only temporal variation of interesting parameters such as species densities, electron/gas temperatures and energies are modeled. Still, the results of these models can not be interpreted very easily due to the complexity and non-linearity of the chemistry being studied. In addition, a reduced chemistry set that can represent properly the complexity of a real kinetic system is necessary to have computationally feasible 1D, 2D, or 3D plasma models for various applications.

In that respect, several methodologies are tried and applied to different chemical kinetics to both provide appropriate interpretation tools and obtain a proper reduced kinetic set for a desired application. They can be categorized as reduction with low dimensional manifold methods and graph-theory related algorithms.

For the first category we can refer to Principle Component Analysis (PCA) and Intrinsic Low Dimensional Manifold (ILDM) as two methodologies that are used for low-temperature plasma kinetics reduction [9–12]. In PCA, a manifold is generated based on the statistics of training simulations, and it is obtained based on linearization of chemical source terms and removing short time scales of a system in ILDM [12]. Although, some successful examples of these methodologies are reported, there are some restrictions for their applicability under different plasma sources and conditions [12].

ILDM can be problematic for chemical kinetics that have more than one equilibrium point under desired conditions and also for plasmas with pulsed power sources. ILDM method is intended to look at chemical systems near an equilibrium point and can therefore identify characteristic eigenvalues of the coupled balance equations which correspond only to long time scales [13]. PCA manifold needed to be linear to work properly, and non-linear PCA methods demands significantly more computational power [12]. Including more time scales in an analysis (ILDM) or investigating more complex systems require an increase in the dimensionality of the manifolds and subsequently of the generated look-up-tables, which can quickly become unfeasible due to excessive demands for computational resources.

Various methodologies are derived and applied to chemical kinetics with basis in graph theories for visualization, topological statistics calculation of a chemical network and determination of pathways, a sequence of reactions.

Flux balance analysis [14] and extreme pathway analysis [15] are two methods to determine the pathways based on mass and energy balance constraints in analogy to Kirchhoff’s laws for electrical circuits [16]. They are used to determine pathways in metabolic physiology. Reaction route graphs, proposed by Fishtik et al. [17], is another technique which is a combination of reaction route theory, graph theory, and Kirchhoff’s laws. This method is applied to heterogeneous and enzyme catalytic reactions [18] and chlorine electrode reactions [19]. Determination of important pathways by following chemical moiety such as a tagged atom is another method proposed by Bai et al [20, 21]. This method is successfully applied to a set of 20 reactions for hydrogen combustion.

Sakai et al. [22] used graph theory to analyze the plasma chemical network in methane. They assigned a vertex for each species and an edge for each individual reaction. With a directed graph, they connected species. They used two parameters to extract qualitative information about chemical processes, the centrality index and the clustering coefficient. With the first one, they calculated a value for each node which reflects either supports of other species to produce this vertex or impact of this species in production of others. In addition, the clustering coefficient represents how much packed/congested a chemical network can be by including or excluding one or more specific species, for instance electron in their demonstrated example. Their approach can be useful to gain qualitative insights about complex chemistries.

Mizui et al. [23] demonstrated a graphical classification method to deduce information on global structures (macroscopic structures) of medium sized chemical reaction systems besides identification of each species role, such as agents, intermediates and products (microscopic features). They tested their methodology for silane and methane plasma chemistries. They drew diagrams with axes of multi-centrality indices to get quantitative insights about chemical mechanisms. The axes of the diagram are betweenness centrality, which represents information as reaction connection between species, and closeness centrality, that indicates information about a location of species in a network.

Murakami and Sakai [24] used graph-theoretical analysis to obtain insight from He–O₂ and He–humid air complex plasma chemistries. They showed how a reduced set from 20-species to 12-species of a He–O₂ model can be achieved with the help of topological centrality (closeness and betweenness) and the scale-freeness nature of a plasma-induced chemistry. Closeness centrality of a node (species) is defined as a measure of how fast information spreads from a given node to other reachable nodes in a network, and the betweenness centrality represents the importance of a species’ role as an intermediate that directly or indirectly connects the reactants and products. In addition, scale-freeness parameter indicates the probability of having a node with ‘k’ degrees (in and out connected edges) remains the same in a reduction process. They showed their reduced system can keep the main features of a DBD plasma.
by comparing the results of 1D simulations with three chemical kinetics: 12-species and 20-species He–O\textsubscript{2} and He-humid air. Their proposed methodology can be used as a systematic pre-processing approach to reduce the complex plasma chemistries before multi-dimensional simulations.

Sun et al [25] reduced CO\textsubscript{2} chemical kinetics for a gliding arc plasma with coupling so-called directed relation graph method (DRP) and sensitivity analysis. Basically, they set a threshold for contribution of each species in production and consumption of target species. Then, they removed species with contribution less than specified threshold. For higher threshold, the chance to remove species that maybe have important impact of targeted features of plasma increases. Therefore, with the help of sensitivity analysis, they aim to achieve more realistic reduced set of a chemistry for the application of their model. The output of a global model is necessary for their reduction technique.

Holmes et al [26] used graph theory to visualize and investigate electron energy demand for production of O\textsubscript{3} in an atmospheric dielectric barrier discharge plasma. The main difference between their method to apply directional graph theory to plasma chemistries with previous attempts is in assigning a vertex not only to each species but also to each individual reaction. Because reactions are set as both edges and vertices, each reaction is connected to each of its reactants and products. Species only are set as vertices. They showed the new approach can produce a clearer image for the interpretation of complex chemistries. In addition, they suggested an algorithm for the estimation of optimal condition sequence for a target plasma chemical species (OCARINA). Furthermore, they used a connectivity matrix as a tool to get some qualitative information about fast and slow reaction pathways in complex plasma kinetics. In general, the accuracy and predictability of their developed algorithm is highly dependent to the chosen parameters to visualize a reaction network. For instance, in their presented example, reaction rates change linearly with the electron energy in the specified range. In addition, the effect of concentration in reaction rates is not included in the analysis. Therefore, for cases that do not meet those limiting assumptions, their methodology will overlook important effects related to the intrinsic non-linearities of the system.

However, the algorithm for determination of chemical pathways that is presented by Lehmann [27] seems a proper option for complex and non-linear plasma chemistry with more than thousands of reactions [2, 4, 28, 29]. This algorithm is the successor of the method proposed by Johnston and Kinnison [30]. No pre-condition is needed to use the algorithm, it is robust, and it avoids the exponential growth of the number of produced pathways by letting users determine a threshold for the rates of pathways. In addition, the effect of concentrations and reaction rates are directly taken into account in analyses in contrast to most of the pre-processing graph-theory related methodologies [22–24, 26]. There is an open source software based on this algorithm, Pumpkin [31]. We implemented the algorithm in our in-house software infrastructure to assess, verify, and use it more easily in addition to check for any bug (Our analysis found that the number of reactions and species are set equal in Pumpkin implementation).

The output of a zero-dimensional model can be an input to this algorithm. The algorithm is started by sorting the species based on their lifetime, destruction rate divided by the particle density. Then each species, from shortest lifetime to the longest, is set as branching species. The branching species is used to connect the pathways which produce it to pathways that consume it. After all the pathways that either produce or consume the branching species are connected, and their rates are calculated based on branching properties, they are deleted from the list of active pathways.

Lehmann’s algorithm lets users determine a threshold for the rate of pathways. If the rate of a pathway is lower than this specified rate, it can be deleted from the list of active pathways. However, in order to keep conservation of rate and net production of each species, the rate of the insignificant pathways should somehow be stored.

After all pathways are formed for the branching species, and insignificant pathways are deleted from the list, the newly formed pathways should be broken up into elementary pathways to prevent unnecessary complexity. After this step, the algorithm can process the next branching species.

One important point should be clarified here: when pathways are formed by connecting the branching species, the branching species is completely recycled. In addition, it is possible that some other species than the branching species are recycled due to connections that occur in this step. Therefore, if the branching species or any other species that are recycled completely have a non-zero net production, their concentration changes automatically vanish from the analysis. In order to avoid this problem, there is another group of pathways that should be determined with their rates in each step, namely pathways that have contributions to density changes.

In this paper, first, the presented algorithm by Lehmann [27] is explained with examples for each step. The focus lies on clarification of ambiguities in the determination of pathways that have contributions in density changes, calculation of their rates, and ensuring conservation when handling insignificant pathways. In other words, the capability of the algorithm is investigated and conservation of critical parameters during whole analyses is proven by designing artful examples. Figure 1 represents the sequence of the algorithm and devoted sections to each step in this paper. Then, the suitability of the algorithm as a reduction method is assessed by applying the newly developed semi-automated tool based on this algorithm on two complex plasma chemistry models: a H\textsubscript{2}O–He microwave induced plasma [32] and a pulsed H\textsubscript{2} plasma [33]. The main difference at this point between this study and Koelman et al [34] lays in developing a tool for chemical reduction (both species and reactions) which is able to represent proper reduced kinetics for complex non-linear plasmas rather than usage of the algorithm for determination of a reduced set of reactions for uncertainty analysis.
Figure 1. Flowchart of the pathway algorithm and devoted sections in this paper for each step.

2. Algorithm description for the connection of pathways

The input of the pathway algorithm can be the time averaged output of a zero-dimensional model or a snapshot at a specific time. In other words, the algorithm needs to have access to the rate or time averaged rate of each reaction involved the chemistry. Based on the reaction rates, the destructive and productive rates for each species can be calculated. In addition, the net production of each species can be defined since
Figure 2. Schematic of two simple pathways when branching species is B. The first pathway, $P_1$ is formed from the reactions $A + e \rightarrow B + e$ and $B + e \rightarrow C + e$. Both reactions have multiplicity one, and the outcome of the formed pathway is $A \rightarrow C$.

The second pathway, $P_2$ is formed from the reactions $A + e \rightarrow B + e$ and $B + e \rightarrow A + e$. Both reactions have multiplicity one, and the pathway is a Null one, where all species are recycled.

the destructive and productive rates (sink and source terms for each species, or consumptive and productive rates) are known. A pathway is defined as a set of interlinked reactions which can be comprised of one or multiple reactions. The algorithm starts by creating a separate pathway from each individual reaction. Then, the connection mechanism is started. Before connecting, all the branching species are sorted based on their lifetime. The lifetime is calculated as

$$\tau_i = \frac{n_i}{d_i},$$

where $n_i$ stands for the particle density of $S_i$ at a specific point in time or averaged over a certain period. It is also possible that the order of branching species is determined beforehand, or some branching species are skipped. However, in the general case, the algorithm picks the species with the shortest lifetime as branching species.

For the branching species $S_b$, all pathways that produce this specific species should be connected with pathways that consume it. Assume that pathway $P_p$ produces $S_b$, and pathway $P_c$ consumes this species. By connecting these two pathways, $S_b$ is recycled by the new formed pathway $P_n$. If the stoichiometry of $S_b$ for $P_p$ is $\nu_{bp}$, and for $P_c$ is $\nu_{bc}$, then for connection of these two pathways, $\nu_{bc}$ times $P_p$ should be added to $\nu_{bp}$ times $P_c$. In other words, the multiplicity of the reactions involved in the newly formed pathway $P_n$ is calculated as

$$m_{jn} = |\nu_{bc}| \cdot m_{jp} + \nu_{bp} \cdot m_{jc}.$$  

Equation (5) should be taken into account for all reactions involved in both $P_p$ and $P_c$. In order to have an as simple as possible pathway, we should divide the multiplicity of all reactions involved in $P_n$ by their greatest common divider. In addition, the rate of the pathway should be multiplied by the
The rates of the new pathways are calculated based on the branching probability concept. Imagine a condition that the net production of $S_b$ is $\delta_{S_b} \leq 0$. Whatever is produced by $P_p$ is consumed by destructive pathways, or it changes the density of $S_b$. The probability that the consumption is done by $P_c$ is

$$\Pi_c = \frac{\nu_{bc} \cdot R_c}{d_p}, \quad (6)$$

where $R_c$ stands for the rate of pathway $P_c$. If $\delta_{S_b} > 0$, whatever is produced by $P_p$ is consumed by destructive pathways, or it has a contribution to the concentration change of $S_b$. The probability that the consumption is done by $P_c$ is

$$\Pi_c = \frac{\nu_{bc} \cdot R_c}{d_p + \delta_{S_b}}, \quad (7)$$

We can have the same probability calculation for productive pathways. If $\delta_{S_b} \leq 0$, whatever is consumed by $P_c$ either is produced by productive pathways, or it has a contribution to the concentration change of $S_b$. The probability that the production is from $P_p$ is

$$\Pi_p = \frac{\nu_{bp} \cdot R_p}{p_b + \delta_{S_b}}, \quad (8)$$

If $\delta_{S_b} > 0$, whatever is consumed by $P_c$ is produced by productive pathways. The probability that the production is from pathway $P_p$ is

$$\Pi_p = \frac{\nu_{bp} \cdot R_p}{p_b}. \quad (9)$$

Before going further in the calculation of the pathway rate, the parameter $D$ is defined for each species to simplify rate calculation later on

$$D_i = \max(p_i, d_i) \quad (10)$$

or

$$D_i = \begin{cases} p_i + \delta_i & \text{if } \delta_i \geq 0, \\ d_i + |\delta_i| & \text{if } \delta_i < 0. \end{cases} \quad (11)$$

Now we can go back to the newly formed pathway $P_n$, the rate of whatever is produced by $P_p$ will be consumed by $P_c$ is

$$\Pi_n \cdot \nu_{bp} \cdot R_p \cdot \frac{\nu_{bc} \cdot R_c}{D_{S_b}} \cdot \nu_{bp} \cdot R_p. \quad (12)$$

In addition, we know that the number of particles of $S_b$ that flow between pathway $P_p$ and $P_c$ is $\nu_{bp} \cdot \nu_{bc}$. Therefore, the rate of $\nu_{bp} \cdot \nu_{bc}$ number of particles of $S_b$ that are produced by $P_p$ and will be consumed by $P_c$ can be written as

$$\nu_{bp} \cdot \nu_{bc} \cdot R_n. \quad (13)$$

By combining relations (12) and (13), the rate of the newly pathway $P_n$ is obtained as

$$R_n = \frac{R_p \cdot R_c}{D_{S_b}}. \quad (14)$$

This strategy gives the same result when the procedure is started from the probability of production and using the consumption rate of destructive pathway $P_c$.

After all possible connections between pathways that produce $S_b$ with pathways that consume $S_b$ are done, the parent pathways are deleted from the list of active pathways. At this stage, it is necessary to check whether the net production of each species and rate of each reaction are the same as the initial set. For each reaction we should check the equality of the following relation

$$R_j = \sum_{k} N_{jk} \cdot R_k, \quad (15)$$

where $N_{jk}$ stands for the number of pathways. In other words, the rate of reaction $j$ is distributed over pathways, and each pathway has a portion of this rate based on the multiplicity of this reaction and the rate of the pathway. In addition, for the total production and destruction of each species we have

$$p_i = \sum_{(k|\nu_{jk} > 0)} \nu_{jk} \cdot R_k, \quad (16)$$

$$d_i = \sum_{(k|\nu_{jk} < 0)} \nu_{jk} \cdot R_k.$$

If the algorithm can keep conservation, we should expect the balance between production and consumption is maintained by the formation of pathways. In the appendix A, an example of a chemical equilibrium chemistry set will be presented to show the capability of the algorithm to keep the balance between production and destruction of species, and to assess the performance of the algorithm to keep conservation of rates and net production of species. In addition, one of the most important points, the correct approach to choose $D_{S_b}$ will be discussed.

It should be mentioned that for all the examples given in this paper, the values of the rates are chosen in such a way that one specific capability of the algorithm can be checked. It is almost impossible to design such short and illustrative examples by a real chemistry and rates. As the aim is to produce very special conditions for verification of both algorithm.
and the implementation in each step, these artificial but artful examples are invented. In addition, the order of branching species is designed such that the purpose of the example is satisfied.

It is proven by the presented example that for the equilibrium condition, the rate conservation is satisfied (equation (15)). However, we should not expect to see consistency between the total production and destruction of each species during the algorithm with the initial values because when connections of pathways are made, the branching species are removed from the system, consequently, \( p_i, d_i \) are affected. The net concentration change of species remains zero during the algorithm which is equal to the initial condition (steady state). In addition, in order to satisfy the conservation of reaction rates for the system, \( D_{S_b} \) should be set from the results of the previous step in the algorithm.

In the next section, determination of pathways that have contributions to density changes with their rate calculation will be discussed.

### 3. Determination of pathways that have contributions to density changes

Imagine in contrast to the previous section \( \delta_{S_b} \neq 0 \). By connecting all pathways which produce the branching species \( S_b \) with pathways that consume it, the new set of chemistry has a zero net production for the branching species (because all new formed pathways are recycled branching species) which differs from the initial set. One subtle point should be mentioned here. We just start to connect if we have both pathways that produce and consume the branching species. The rest of this part is based on this assumption, we have both productive and consumptive pathways, we connect them, now the net production of \( S_b \) becomes zero.

In order to avoid this inconsistency in the algorithm, Lehmann [27] introduces another type of pathways—that should be recognized in each step, and their rates should be calculated properly—Paths that have contributions to density changes.

For the case \( \delta_{S_b} > 0 \), all the pathways that produce species \( S_b \) have a contribution to the density change of \( S_b \). The probability of this contribution is defined as

\[
\Pi_{\Delta} = \frac{\delta_{S_b}}{d_{S_b} + |\delta_{S_b}|} = \frac{\delta_{S_b}}{D_{S_b}}.
\]

Therefore, a new rate should be assigned to each pathway that produces \( S_b \)

\[
\Delta R_p = \frac{R_p \cdot \delta_{S_b}}{D_{S_b}}.
\]

In other words, when all possible connections have been made, we should delete all parent pathways from the the list of active pathways (pathways that will be used for the connection of the next branching species), but we should store the productive pathways for the current branching species \( S_b \) in the new list (which is designed for the pathways that have contributions to density changes) with the rate that should be calculated by equation (18).

For the case that \( \delta_{S_b} < 0 \), all consumptive pathways have contributions to the density changes of \( S_b \). The probability of the contribution is

\[
\Pi_{\Delta} = \frac{|\delta_{S_b}|}{p_{S_b} + |\delta_{S_b}|} = \frac{|\delta_{S_b}|}{D_{S_b}}.
\]

This contribution is distributed between all consumptive pathways, the rate of each pathway is

\[
\Delta R_c = \frac{R_c \cdot |\delta_{S_b}|}{D_{S_b}}.
\]

Similar to the case of positive net production, we should store the consumptive pathways in a list of pathways that have a contribution to density changes, but with the new calculated rate, equation (20). In the following section, an example for a non-equilibrium chemical set will be presented, and the role of pathways that have contributions to density changes will be illustrated and discussed.

#### 3.1. Chemical non-equilibrium set: example

For the chemical non-equilibrium case, we consider the same set of reactions as appendix A but with different rates.

As it is noted in appendix A, A is a hypothetical species that can be ionized to \( A^+ \) and electronically excited to different levels \( A^* \) and \( A^{**} \). e represents the electron, and rates of reactions are designed in such way that the purpose of the example is satisfied. In addition, the simplest possible way to write a reaction is used below. For example, reaction \( r_1 \) should have one extra electron on both sides. However, in pathway analysis the net reaction is used.

| No. | Reaction | Rate |
|-----|----------|------|
| \( r_0 \) | \( A \rightarrow A^{**} \) | \( R_0 = 10 \) |
| \( r_1 \) | \( A + A^{**} \rightarrow 2A^* \) | \( R_1 = 15 \) |
| \( r_2 \) | \( A^+ + e \rightarrow A \) | \( R_2 = 20 \) |
| \( r_3 \) | \( A \rightarrow A^+ + e \) | \( R_3 = 10 \) |
| \( r_4 \) | \( 4A^* \rightarrow 4A \) | \( R_4 = 5 \) |

Production, destruction, \( D \) parameter, and net for each species at the beginning of the algorithm have the following values:

| Species | \( p \) | \( d \) | \( D \) | \( \delta \) |
|---------|--------|--------|--------|--------|
| A       | 40     | 35     | 40     | 5      |
| A*      | 30     | 20     | 30     | 10     |
| A**     | 10     | 15     | 15     | -5     |
| A^+     | 10     | 20     | 20     | -10    |
| e       | 10     | 20     | 20     | -10    |
As is clear, we do not have a chemical equilibrium case. The net concentration changes are not zero for all species.

3.1.1. First branching species $A^{**}$. When the pathways which produce and consume $A^{**}$ are connected, the concentration change of this species is removed from the resulted pathways. The formed pathways in this iteration are:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_2$ | $1 \times r_2$ | $A^+ + e \rightarrow A$ | $R_2 = 20$ |
| $P_3$ | $1 \times r_3$ | $A \rightarrow A^+ + e$ | $R_3 = 10$ |
| $P_4$ | $1 \times r_4$ | $4A^* \rightarrow 4A$ | $R_4 = 5$ |
| $P_5$ | $1 \times r_0 + 1 \times r_1$ | $2A \rightarrow 2A^*$ | $R_5 = 10$ |

The resulting production, destruction, net, and $D$ parameter from the above list for each species are:

| Species | $p$ | $d$ | $D$ | $\delta$ |
|---------|-----|-----|-----|-------|
| $A$ | 40 | 30 | 40 | 10 |
| $A^*$ | 20 | 20 | 20 | 0 |
| $A^{**}$ | 0 | 0 | 0 | 0 |
| $A^+$ | 10 | 20 | 20 | −10 |
| $e$ | 10 | 20 | 20 | −10 |

As expected, the branching species $A^{**}$ is no longer on the list. Another interesting thing is the zero net for species $A^*$. This species is not the branching species for this step, but connecting pathways has an influence on the net concentration change of this species too.

If the algorithm is stopped at this stage, the concentration change is not the same as the initial set. Not only $A^{**}$ is recycled due to the connection, but also $A^*$ loses its concentration change, and $A$ is produced more than the initial set.

This is a subtle but important part of the algorithm. Based on Lehmann [27], only for the branching species we should consider the pathways that have impacts on the concentration change. In this case, the concentration change or net of species $A^{**}$ at the beginning of the algorithm is $−5$ (it is consumed in this chemistry). Therefore, based on equation (20), we should assign a new rate for pathways that consume this species. Now we have a new list of pathways that is the summation of the previous one plus the pathway that has a contribution to the concentration change of $A^{**}$:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\Delta P_1$ | $1 \times r_1$ | $A + A^{**} \rightarrow 2A^*$ | $\Delta R_1 = \frac{15 \times 5}{2} = 5$ |
| $P_2$ | $1 \times r_2$ | $A^* + e \rightarrow A$ | $R_2 = 20$ |
| $P_3$ | $1 \times r_3$ | $A \rightarrow A^* + e$ | $R_3 = 10$ |
| $P_4$ | $1 \times r_4$ | $4A^* \rightarrow 4A$ | $R_4 = 5$ |
| $P_5$ | $1 \times r_0 + 1 \times r_1$ | $2A \rightarrow 2A^*$ | $R_5 = 10$ |

After considering the pathway $\Delta P_1$, the concentration change not only for $A^{**}$ but also for $A$ and $A^*$ becomes consistent with the initial set.

Next, lets check the conservation of rates, equation (15). Before considering the pathway $\Delta P_1$, we had the following result:

| No. | Reaction | The RHS of (15)= |
|-----|----------|----------------|
| $r_0$ | $A \rightarrow A^{**}$ | 10 |
| $r_1$ | $A + A^{**} \rightarrow 2A^*$ | 10 |
| $r_2$ | $A^* + e \rightarrow A$ | 20 |
| $r_3$ | $A \rightarrow A^* + e$ | 10 |
| $r_4$ | $4A^* \rightarrow 4A$ | 5 |

Clearly, there is no conservation of rate for reaction $r_1$. After introducing pathway $\Delta P_1$, conservation of rate is satisfied and the RHS of equation (15) becomes equal to 15 which is consistent with the initial rate for this reaction.

3.1.2. Second branching species $A$. For the connection process in this step, we should just connect pathways which are in the list of active pathways, which means pathway $\Delta P_1$ should not be taken into account for connection. The produced pathways in this iteration are:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_6$ | $1 \times r_2 + 1 \times r_3$ | Null | $R_6 = 5$ |
| $P_7$ | $1 \times r_0 + 1 \times r_1 + 2 \times r_2$ | $2A^* + 2e \rightarrow 2A^*$ | $R_7 = 5$ |
| $P_8$ | $4 \times r_3 + 1 \times r_4$ | $4A^* \rightarrow 4A^* + 4e$ | $R_8 = 1.25$ |
| $P_9$ | $2 \times r_0 + 2 \times r_1 + 1 \times r_4$ | Null | $R_9 = 2.5$ |

Production, destruction, net, and $D$ parameter from the above pathway list are:

| Species | $p$ | $d$ | $D$ | $\delta$ |
|---------|-----|-----|-----|-------|
| $A$ | 0 | 0 | 0 | 0 |
| $A^*$ | 10 | 5 | 10 | 5 |
| $A^{**}$ | 0 | 0 | 0 | 0 |
| $A^+$ | 5 | 10 | 10 | −5 |
| $e$ | 5 | 10 | 10 | −5 |

Obviously, branching species $A$ is removed from the list, but the concentration change of $A^*$, $A^+$, and $e$ are also affected. Now we need to complete the list of pathways by adding pathways that have contributions to the generation of $A$ from the previous step. The net of $A$ is positive in the initial case and also after the first branching species $A^{**}$.

Pathways $P_2$ and $P_4$ are producing $A$ in the previous iteration. Another subtle point is how to choose $\delta_{b_1}$ to calculate the rate of pathways that have contributions to density changes. Not only for $D_{b_0}$ but also for $\delta_{b_1}$, the value from the previous step should be chosen instead of from the initial set. In other
words, the net production of A from the previous step’s active pathway list should be used.

### Table 1

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\Delta P_1$ | $1 \times r_1$ | $A + A^{**} \rightarrow$ | $\Delta R_1 = \frac{15 \times 5}{15} = 5$ |
| $\Delta P_2$ | $1 \times r_2$ | $A^+ + e \rightarrow A$ | $\Delta R_2 = \frac{10 \times 20}{40} = 5$ |
| $\Delta P_3$ | $1 \times r_4$ | $4A^* \rightarrow 4A$ | $\Delta R_4 = \frac{10 \times 5}{40} = 1.25$ |
| $P_6$ | $1 \times r_2 + 1 \times r_3$ | Null | $R_6 = 5$ |
| $P_7$ | $1 \times r_0 + 1 \times r_1$ | $2A^+ + 2e \rightarrow$ | $R_7 = 5$ |
| $P_8$ | $4 \times r_3 + 1 \times r_4$ | $4A^* \rightarrow 4A^* + 4e$ | $R_8 = 1.25$ |
| $P_9$ | $2 \times r_0 + 2 \times r_1$ | Null | $R_9 = 2.5$ |

Now with the new set of pathways, we can see that the concentration changes of all species are consistent with the initial set. Next, we should check the conservation of rates. First, we just look at the generated pathways from this step without considering the effect of pathways that have contributions to concentration changes. Therefore, we have:

### Table 2

| No. | Reaction | The RHS of (15) = |
|-----|----------|------------------|
| $r_0$ | $A \rightarrow A^{**}$ | 10 |
| $r_1$ | $A + A^{**} \rightarrow 2A^*$ | 10 |
| $r_2$ | $A^+ + e \rightarrow A$ | 15 |
| $r_3$ | $A \rightarrow A^* + e$ | 10 |
| $r_4$ | $4A^* \rightarrow 4A$ | 3.75 |

After adding the effect of the three $\Delta$ pathways, we see the consistency between the initial set and the current result.

#### 3.1.3. Third branching species $A^*$. The generated pathways in this stage are:

### Table 3

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_{10}$ | $1 \times r_2 + 1 \times r_3$ | Null | $R_{10} = 5$ |
| $P_{11}$ | $2 \times r_0 + 2 \times r_1 + 1 \times r_4$ | Null | $R_{11} = 0.25$ |
| $P_{12}$ | $2 \times r_0 + 2 \times r_1 + 4 \times r_2 + 4 \times r_3 + 1 \times r_4$ | Null | $R_{12} = 1.25$ |

Now the production, destruction, net, and $D$ parameter just from the above pathway list are:

### Table 4

| Species | $p$ | $d$ | $D$ | $\delta$ |
|---------|-----|-----|-----|---------|
| $A$     | 0   | 0   | 0   | 0       |
| $A^*$   | 0   | 0   | 0   | 0       |
| $A^{**}$| 0   | 0   | 0   | 0       |
| $A^+$   | 0   | 0   | 0   | 0       |
| $e$     | 0   | 0   | 0   | 0       |

We have only null pathways, so the above results are expected. Now we should add pathways that have contributions to the concentration change of $A^*$. Because in the initial set the net of this species is positive, and still remains positive (previous branching species $A$), we should look at productive pathways from the previous iteration, which is only pathway $P_7$. Therefore, the list of pathways for this iteration is:

### Table 5

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\Delta P_1$ | $1 \times r_1$ | $A + A^{**} \rightarrow$ | $\Delta R_1 = \frac{15 \times 5}{15} = 5$ |
| $\Delta P_2$ | $1 \times r_2$ | $A^+ + e \rightarrow A$ | $\Delta R_2 = \frac{10 \times 20}{40} = 5$ |
| $\Delta P_3$ | $1 \times r_4$ | $4A^* \rightarrow 4A$ | $\Delta R_4 = \frac{10 \times 5}{40} = 1.25$ |
| $P_6$ | $1 \times r_2 + 1 \times r_3$ | Null | $R_6 = 5$ |
| $P_7$ | $1 \times r_0 + 1 \times r_1$ | $2A^+ + 2e \rightarrow$ | $R_7 = 5$ |
| $P_8$ | $4 \times r_3 + 1 \times r_4$ | $4A^* \rightarrow 4A^* + 4e$ | $R_8 = 1.25$ |
| $P_9$ | $2 \times r_0 + 2 \times r_1$ | Null | $R_9 = 2.5$ |
| $P_{10}$ | $2 \times r_0 + 2 \times r_1 + 4 \times r_2 + 4 \times r_3 + 1 \times r_4$ | Null | $R_{10} = 1.25$ |

Now we can see that not only do we have conservation of rates for each reaction, but also the net production for each species is the same as initial conditions. In conclusion, in chemical non-equilibrium sets, the algorithm is able to satisfy conservation. In order to calculate the rate of active pathways and pathways that have contributions to density changes, $D$ and $\delta$ values should be set from the results of only active pathways of the previous step of the algorithm.

### 4. Three possible ways to handle insignificant pathways

When the pathway algorithm is used for a complex chemistry, a lot of pathways will be produced. From the analysis point of view, it does not yield insights into the underlying chemical mechanisms. From the computational viewpoint, the problem can quickly become too memory intensive. In order to avoid exponential growth of the number of pathways, it is possible to define a threshold for the rate of pathways. This means, if a formed pathway has a rate smaller than the threshold value, it can be deleted from the list of active pathways and added to the list of insignificant pathways.

Now the question is how we should treat these insignificant pathways. Shall we simply remove all of them from the analysis, or shall we take into account their effects on the calculation somehow?

First of all, in order to keep conservation in the algorithm, we should not simply remove insignificant pathways. As suggested by Lehmann, we should keep track of the deleted portion of the rate of each reaction and deleted portion of the production and destruction rates for each species involved in the
insignificant pathways. This information can be tracked by the following equations

\[ \tilde{\bar{R}}_j = \sum_{k=0}^{N_p} \bar{m}_{jk} \cdot \tilde{\bar{R}}_k, \]

\[ \tilde{\bar{p}}_i = \sum_{(k|m_k>0)} \bar{\nu}_{ik} \cdot \tilde{\bar{R}}_k, \]  \hspace{1cm} (21)

\[ \tilde{\bar{d}}_i = \sum_{(j|b_j<0)} |\bar{\nu}| \cdot \tilde{\bar{R}}_j, \]

where the tilde sign stands for insignificant pathways, and summation should be carried out over all insignificant pathways \( \tilde{N}_p \).

In the following subsections, three different ways to incorporate the impact of the insignificant pathways will be explained with illustrative examples for each method.

4.1. Connection between active pathways, active-insignificant pathways, and insignificant pathways

One way to deal with insignificant pathways is to do all possible connections between pathways. This means that not only connections between active pathways are made, but also between active and insignificant pathways, and between insignificant pathways.

This is similar to the situation that we do not have any threshold, so we connect all possible pathways that produce and consume the branching species. The only difference is that we have two different lists, one for active pathways and one for insignificant pathways. In addition, the effect of insignificant pathways which have contributions to the density changes should be taken into account. This is helpful to avoid working with a large number of pathways for post-analysis. However, from a computational point of view, this method is expensive. For complex chemistry, the problem can quickly become too memory intensive.

In addition, in order to handle the insignificant pathways properly and to avoid taking into account the information of deleted pathways in equation (21) twice, whenever a connection between an insignificant pathway with another insignificant pathway or active pathway is made, the information of this insignificant pathway should be subtracted from equation (21). After the new insignificant pathway is formed, the information for species and reactions that are involved in this new insignificant pathway should be used to update equation (21). In appendix B this method is illustrated with an example.

It is shown in the example that the algorithm violates conservation of the net density changes and reaction rates by choosing wrongly the pathways that have impacts on density changes. It is proven that if pathways that have impacts on density changes are chosen based on the sign of the net production of branching species from the results of the last step of the algorithm, the conservation is satisfied.

Therefore, in order to keep conservation, as the algorithm proceeds by generation new set of pathways after each branching species, not only \( D \) and \( \delta \) should be set based on the new set of pathways, but also the sign of the net production of species should be determined based on the new set.

4.2. Connection of only active pathways

It will be shown in the following examples that, it is possible to take a subset of the whole chemistry, start connecting pathways, and reach the final set of pathways while having conservation of rates and density changes. With a correct algorithm, it is possible to connect only active pathways for each branching species. In addition, by keeping track of deleted information by the help of equation (21), we are able to satisfy the conservation of rates and density changes.

The only subtle thing that should be done correctly is choosing the values for \( D_S \) and \( \delta_S \). These values should be obtained from a subset of the chemistry, in this case just from the active pathways that are produced or remained in the list from the previous step.

The advantage of this method is that we can release memory by deleting insignificant pathways and only book-keeping the information of equation (21). However, although conservation of rates for all reactions and conservation of net production for all species are kept, because the algorithm works on a subset of the chemistry (active pathways), the new formed pathways and their rates are not a proper representation of the original chemistry set. In particular, rates are bigger than they should be for a proper representation. An example in appendix C demonstrates this method.

4.3. Connection between active pathways and active-pseudo insignificant pathways

Lehmann mentions in section 5.4 [27] that the connection of pathways should be not only between active pathways but also between active and productive \( (p_{g_S}) \) and destructive \( (d_{g_S}) \) rates of branching species from the insignificant pathways. In addition, the produced pathway from the connection of \( p_{g_S} \) and \( d_{g_S} \) should be taken into account, and the pathways which have contributions to density changes should also be added and their rates should be calculated. After all of these connections are done, equation (21) should be updated in order to keep track of deleted information from the algorithm.

First of all, it is not mentioned clearly how a pathway can be produced from the connection of an active pathway with only a rate (productive or destructive). Although we can calculate the rate of this pathway, it is not clear how to update equation (21) without knowing the multiplicity of reactions and stoichiometries of the species involved in this newly produced pathway.

In general there are two ways to update equation (21):

- To not delete insignificant pathways from memory. Due to the connection of active and insignificant pathways, we produce new pathways with all necessary information that are stored for normal pathways. If we follow this way, already we keep storing all produced pathways, and we will have a memory shortage problem. This way of treating insignificant pathways was explained in section 4.1.
To update the deleted equations only based on the active portion of the newly formed pathways, this method will be discussed in this section.

Second, this part of the algorithm is not explained clearly in Lehmann’s paper [27]. The following mechanism is designed by the authors to present section 5.4 of [27]. It is assumed that for each branching species there are two pseudo pathways, one with the deleted productive rate and the other with the deleted consumptive rate as follows

\[
S_b \rightarrow \tilde{P}_c = \tilde{d}_b S_b, \tag{22}
\]

\[
\rightarrow S_b \leftarrow \tilde{P}_p = \tilde{p}_b S_b.
\]

Therefore, for each branching species, not only do we have a connection between active pathways, but also between active productive pathways and pseudo consumptive pathways \(\tilde{P}_c\) and between active destructive pathways and pseudo productive pathway \(\tilde{P}_p\). Equation (21) should be updated when a pathway is produced from the connection of an active pathway with a pseudo one. As mentioned above, we do not have complete information of insignificant pathways, and we replace them by pseudo pathways. Therefore, we should update the deleted information based on the multiplicity of whatever reactions and stoichiometry of whatever species involved in the active pathways that are used in connections. In addition, the rate of the newly produced pathway (from connection of active and pseudo pathways) should be used in updating necessary information of equation (21).

For the connection of active productive pathways \(R_p\) is the rate of each active pathway that is used for connecting) with the pseudo destructive pathway we have the rate

\[
\tilde{R} = R_p \frac{R_c}{D_{S_b}} = \tilde{d}_b R_p.
\]  

For the connection of destructive active pathways \(R_c\) is the rate of each active pathway that is used for connecting) with the pseudo productive pathway we have the rate

\[
\tilde{R} = R_p \frac{R_c}{D_{S_b}} = \tilde{p}_b R_c.
\]  

In addition, we need to build a pathway that is formed from the connection of deleted destructive rate and deleted productive rate. The rate of this pathway is calculated as equation (25),

\[
\tilde{R} = R_p \frac{R_c}{D_{S_b}} = \tilde{p}_b \tilde{d}_b.
\]  

However, we do not have any information for stoichiometry of species and multiplicity of reactions for this pathway because we delete all insignificant pathways from the memory and we only store equation (21) for species and reactions.

One important point that should be kept in mind is that when we connect an active pathway with a pseudo pathway, we already produce another pseudo insignificant pathway. At the end of the connection process (for a specific branching species), we have zero rate for deleted productive and consumptive pseudo pathways \((\tilde{p}_b = \tilde{d}_b = 0)\). Instead of those two rates, we have a rate from the connection of the pseudo productive pathway with pseudo destructive, equation (25).

The reason for the zero deleted rate is due to the fact that when insignificant pathways with each other or insignificant pathways with active pathways are connected, in order to avoid to take into account the information for the deleted parts twice, the effect of insignificant pathways that are used in these connections from equation (21) (in this case, the rate of pseudo destructive and productive pathways, \(\tilde{d}_b\) and \(\tilde{p}_b\)) should be subtracted. Then, after connections, again equation (21) should be updated based on the newly formed pathways.

In addition, if \(\delta_{S_b}\) is positive (based on the results of previous branching species), we need to take into account the effect of deleted productive pathways for the density change of the branching species \(S_b\). Therefore, we should assign the following rate for the contributions of insignificant pathways to the density change of the branching species

\[
\Delta \tilde{R} = \frac{\delta_{S_b} \tilde{p}_b S_b}{D_{S_b}}.
\]  

If the sign of net production of the branching species is negative (based on the results of previous branching species), we should take into account the contributions of the consumptive insignificant pathways to the concentration change of the branching species with the following rate

\[
\Delta \tilde{R} = |\delta_{S_b}| \frac{\tilde{d}_b S_b}{D_{S_b}}.
\]  

In section 4.4, an example is presented to illustrate this method.

### 4.4. Connection between active pathways and active-pseudo insignificant pathways: example

For an illustration of this particular connection method, we use the same chemistry set from the appendix C. The difference between this method and the previous one is that not only active pathways are connected, but also insignificant pathways are replaced with pseudo pathways whose rates come from the deleted information equation (21).

The initial set contains the following reactions:

| No. | Reaction | Rate |
|-----|----------|------|
| \(r_0\) | \(A \rightarrow A(v = 1)\) | \(R_0 = 5\) |
| \(r_1\) | \(A(v = 1) \rightarrow A(v = 2)\) | \(R_1 = 10\) |
| \(r_2\) | \(2A(v = 2) \rightarrow A + A(v = 1)\) | \(R_2 = 15\) |
| \(r_3\) | \(A(v = 1) \rightarrow A(v = 3)\) | \(R_3 = 20\) |
| \(r_4\) | \(A(v = 4) \rightarrow A(v = 2)\) | \(R_4 = 80\) |
| \(r_5\) | \(A(v = 2) \rightarrow A(v = 5)\) | \(R_5 = 100\) |
| \(r_6\) | \(A(v = 2) \rightarrow A(v = 3)\) | \(R_6 = 15\) |
Note that A is a hypothetical species that can have different vibrational levels. In addition, the species production, destruction, and their net in the initial step are:

| Species     | $p$ | $d$ | $D$ | $\delta$ |
|-------------|-----|-----|-----|----------|
| A           | 15  | 5   | 15  | 10       |
| A($v = 1$)  | 20  | 30  | 30  | $-10$    |
| A($v = 2$)  | 90  | 145 | 145 | $-55$    |
| A($v = 3$)  | 35  | 0   | 35  | 35       |
| A($v = 4$)  | 0   | 80  | 80  | $-80$    |
| A($v = 5$)  | 100 | 0   | 100 | 100      |

The threshold value is set to 11 for this example.

4.4.1. First branching species A($v = 1$). There is no difference between the results of this method and the previous one in section 4.2 for the first branching species. The list of active pathways is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_4$ | $A(v = 4) \rightarrow A(v = 2)$ | $R_4 = 80$ | |
| $P_5$ | $A(v = 2) \rightarrow A(v = 5)$ | $R_5 = 100$ | |
| $P_6$ | $A(v = 2) \rightarrow A(v = 3)$ | $R_6 = 15$ | |

The list of pathways that have contributions to density changes:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\Delta P_4$ | $A(v = 1) \rightarrow A(v = 2)$ | $\Delta R_4 = 3.33333$ | |
| $\Delta P_5$ | $A(v = 1) \rightarrow A(v = 3)$ | $\Delta R_5 = 6.66667$ | |

Insignificant pathways are produced for this step too, and equation (21) are updated. In this method, we make pseudo pathways from the deleted portion of reactions and net production. In the implementation, no memory is allocated for these pseudo pathways, but here, for illustration of the algorithm, this list is shown. We only keep track of equation (21) in the code implementation.

The list of the pseudo pathways from the deleted portion of reactions and net production of species is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\tilde{P}_0$ | $A$ | $\tilde{R}_0 = 15$ | |
| $\tilde{P}_1$ | $A \rightarrow$ | $\tilde{R}_1 = 5$ | |
| $\tilde{P}_2$ | $A(v = 2)$ | $\tilde{R}_2 = 1.66667$ | |
| $\tilde{P}_3$ | $A(v = 2)$ | $\tilde{R}_3 = 25$ | |
| $\tilde{P}_4$ | $A(v = 3)$ | $\tilde{R}_4 = 13.3333$ | |

where the $\tilde{P}_i$, $\tilde{r}_i$, and $\tilde{R}_i$ represent the pseudo pathways, reactions, and rates, respectively.

4.4.2. Second branching species A($v = 2$). For the second branching species, we connect all possible pairs of active pathways, active with pseudo pathways, and pseudo pathways. The list of active pathways is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_{11}$ | $A(v = 4) \rightarrow A(v = 5)$ | $R_{11} = 57.1429$ | |

The list of active-pseudo pathways formed in this step is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_5$ | $A(v = 4) \rightarrow$ | $R_5 = 14.2857$ | |
| $P_6$ | $A(v = 5)$ | $R_6 = 1.190478$ | |
| $\tilde{P}_7$ | $A(v = 3)$ | $R_7 = 0.17857$ | |

The list of pathways that have contributions to density changes is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\Delta P_1$ | $A(v = 1) \rightarrow A(v = 2)$ | $\Delta R_1 = 3.33333$ | |
| $\Delta P_3$ | $A(v = 1) \rightarrow A(v = 3)$ | $\Delta R_3 = 6.66667$ | |
| $\Delta P_5$ | $A(v = 2) \rightarrow A(v = 5)$ | $\Delta R_5 = 41.6667$ | |
| $\Delta P_6$ | $A(v = 2) \rightarrow A(v = 3)$ | $\Delta R_6 = 6.25$ | |

After connecting active pathways similar to the previous section C, we have one pathway with a rate higher than the threshold, $P_{11}$, and one pathway with the rate lower than the threshold. In this implementation, we do not need to keep track of insignificant pathways, only the deleted information equation (21) are updated. For clarity, the list of insignificant pathways for this step is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_7$ | $A(v = 2)$ | $R_7 = 1.6667$ | |
| $P_8$ | $A(v = 3)$ | $R_8 = 3.33333$ | |
| $P_9$ | $A(v = 2) \rightarrow A$ | $R_9 = 5$ | |
| $P_{10}$ | $A(v = 2)$ | $R_{10} = 10$ | |
| $P_{12}$ | $A(v = 4) \rightarrow A(v = 3)$ | $R_{12} = 8.57143$ | |

Although the threshold value is set to 11, all pathways that are produced due to the connection of active and pseudo pathways do not have rates lower than the threshold (in this case $P_5$). This is expected because there is no guarantee that the rate of deleted production or destruction of branching species becomes lower than the threshold value. Therefore, we should not expect to see a rate of connected active-pseudo pathways lower than the specified threshold, although the information of the connection of active pathways with deleted rate of branching species ($\tilde{p}$ and $\tilde{d}$) automatically are used to update the deleted information equation (21).

In addition to the pseudo-active pathways, we should also take into account the effect of deleted production and destruction rates of branching species in density changes. Because the net production of the current branching species in the previous step is: $p + d + \tilde{p} + \tilde{d} = -58.33333$, the deleted destruction rate has a contribution to the density change of A($v = 2$). The
rate of this contribution is represented with the deleted contribution rate \( \Delta \tilde{d} \) like before, equation (27).

As mentioned in the previous step, the insignificant pathways are substituted with these pseudo pathways and reactions, so the way pathways are connected should satisfy the conservation of these pseudo reactions rates too. For the current branching species, there are both deleted production and destruction rates. Therefore, at the end of this step, also the rate of the pathway which is produced due to the connection of these two deleted rates, equation (25), should be taken into account to prove the conservation for pseudo reactions.

The species net balances with this method are:

| Species   | \( p \) | \( d \) | \( \tilde{p} \) | \( \tilde{d} \) | \( \Delta p \) | \( \Delta d \) | \( \Delta \tilde{p} \) |
|-----------|---------|---------|----------------|----------------|----------------|----------------|----------------|
| A         | 0       | 0       | 15             | 5              | 0              | 0              | 0              |
| A(\( v = 1 \)) | 0       | 0       | 0              | 0              | 0              | 10             | 0              |
| A(\( v = 2 \)) | 0       | 0       | 0              | 0              | 3.33333        | 47.9167        | 0              |
| A(\( v = 3 \)) | 0       | 0       | 22.0833        | 0              | 12.9167        | 0              | 0              |
| A(\( v = 4 \)) | 0       | 57.1429 | 0              | 22.8571        | 0              | 0              | 0              |
| A(\( v = 5 \)) | 57.1429 | 0       | 1.19048        | 0              | 41.667         | 0              | 0              |

Before looking at the balances for reaction rates, another point should be mentioned. As can be seen, the rate of active pathway (\( P_{11} \)) and also insignificant pathway (\( P_{12} \)) at this step are different from the same pathways in the previous section C. This discrepancy is due to the way the \( D \) and \( \delta \) are calculated in the two implementations. In the previous method, those values are based on only active pathways, while in this method, \( D \) and \( \delta \) are based on active pathways and deleted information (a substitution of insignificant pathways). Although both methods satisfy conservation, throwing away the deleted portion of chemistry results in this difference. In the authors’ view, the current method of connection truly represents the reality.

The balances for the real and pseudo reaction rates are:

| No. | \( R \) | \( \tilde{R} \) | \( \Delta R \) | \( R_0 \) |
|-----|--------|----------------|----------------|---------|
| \( r_0 \) | 0 | 5 | 0 | 5 |
| \( r_1 \) | 0 | 6.6667 | 3.33333 | 10 |
| \( r_2 \) | 0 | 15 | 0 | 15 |
| \( r_3 \) | 0 | 13.3333 | 6.6667 | 20 |
| \( r_4 \) | 57.1429 | 22.8571 | 0 | 80 |
| \( r_5 \) | 57.1429 | 1.19048 | 41.667 | 100 |
| \( r_6 \) | 0 | 8.75 | 6.25 | 15 |

| No. | Pseudo | Active-Pseudo | Pseudo-Pseudo | \( \Delta \text{Pseudo} \) | \( \tilde{R} \) |
|-----|--------|---------------|---------------|----------------|---------|
| \( \tilde{r}_0 \) | 15 | 0 | 0 | 0 | 15 |
| \( \tilde{r}_1 \) | 5 | 0 | 0 | 0 | 5 |
| \( \tilde{r}_2 \) | 0 | 1.369048 | 0.297619 | 0 | 1.6667 |
| \( \tilde{r}_3 \) | 0 | 14.2857 | 0.297619 | 10.4167 | 25 |
| \( \tilde{r}_4 \) | 13.3333 | 0 | 0 | 0 | 13.3333 |

In the above table, ‘Pseudo’ represents the pseudo-pathway. ‘Active-Pseudo’ and ‘Pseudo-Pseudo’ stand for pathways that are produced from the connection of active-pseudo pathways and pseudo-pseudo pathways, respectively. \( \Delta \text{Pseudo} \) is a representation of the contributions of deleted rates to density changes. As can be seen, there is conservation of rates for pseudo reactions that are substitutions of the insignificant pathways in the previous step. In the current step, the list of pseudo pathways is updated based on the new deleted productive and destructive rates for each species. As mentioned
before, in this method, we do not need to keep track of insignificant pathways. They are deleted from memory whenever they are produced, and deleted information equation (21) are updated properly. These pseudo pathways and reactions are not created in the implementation, they are shown here just for the illustration of the algorithm.

5. Systematic reduction method for non-equilibrium plasma chemistry based on the pathway algorithm

After the capability of the algorithm is demonstrated in previous sections for keeping conservation of rates and net production of species, the capability of the algorithm to be used as a reduction method is tested in the following sections. In order to assess this capability, a semi-automated tool has been developed. After the global model is run for the full chemistry, the results of the model are used as input for the pathway code. The pathway algorithm in our in-house framework analyzes the chemistry at a specific moment in time. There is no averaging of the reaction rate over the simulation time. It is believed this way of analysis prevents losing important information of reactions that may have only large rates in a small time interval and almost zero rate for the rest of simulation time, but they play an important role in the formation of critical chemical mechanisms. It is also possible averaging over a specified time interval has been done as a post process step over results of a global model to minimize the effect of losing information, and then, the obtained new data are used as an input for the pathway analysis tool. However, in this research, averaging over a time interval is not used, and multiple points of time is chosen to reduce a chemistry set with significant temporal variations, that approach will be discussed in this section.

After a suitable threshold is defined as the input for the pathway code, the code runs the analysis and as a result, two pathway lists are produced with pathways that have rates higher than the threshold. One of the list belongs to the active pathways in the main part of the algorithm and the other is for the pathways which have contributions to density changes, see section 3.

The tool collects the reactions belonging to these lists and puts the rest of the reactions in a reaction Blacklist, meaning a list of reactions that are not important with this specified threshold at the specific time of the simulation according to the pathway analysis. If \( R(L) \) represents the list of reactions in list \( L \), and if the pathway code is run only for one specific moment, we have

\[
R(L_e) = R(L_e) \cup R(L_{Bl}),
\]

(28)

where \( L_e, L_r, \) and \( L_{Bl} \) stand for the reactions present in the full chemistry list, reduced chemistry, and blacklist. The number of points in time that are used for the pathway analysis depends on the chemistry and physics being studied. For example, if steady state results are important, one can run the pathway code once equilibrium is reached. However, for a pulsed power input, the development of plasma chemistry during the whole cycle or only at the ignition time are important. In that case, the pathway analysis should be done over multiple points in time. Another example is when the reduction is desired over a range of power inputs or different concentrations of the feed gases, for instance, \( \text{H}_2\text{O} \) in He [32]. In such cases, it is possible to run the tool with various setting for the pathway analysis either at different points in time of one specific global model, or at a specific time for multiple global models. The tool has two different possibilities for the output, either as a reaction Blacklist, providing the common reactions over all groups:

\[
R_{mL} = R(L_{Bl_1}) \cap R(L_{Bl_2}) \cap R(L_{Bl_i}) \ldots,
\]

(29)

or a combination of all Blacklists:

\[
R_{mL} = R(L_{Bl_1}) \cup R(L_{Bl_2}) \cup R(L_{Bl_i}) \ldots,
\]

(30)

where \( R_{mL_i} \) represents the Blacklist reactions for the set \( i \). Users can define as many sets they like for reduction.

Another capability of the tool is to determine species that do not have any reactions in the final reduced set. In addition, it is possible that in the reduced set, for some species only productive or destructive reactions remain. In these cases, the tool gives a warning to either remove these species from the list of important species and consequently remove the remaining productive or destructive reactions of these species from the list of the reduced set, or to lower the threshold to have both productive and destruction mechanisms for these specific species in the reduced reaction set. As for the reactions, for species the following relations are valid, if \( S(L) \) stands for the species that are related to each other by the reactions in list \( L \).

\[
S(L_1) = S(L_2) \cup S(L_{Bli}).
\]

(31)

For group reduction, two possible outputs are available, species that are common in all Blacklists (32) or the summation of species in all sets of Blacklists (33):

\[
S_{mL} = S(L_{Bl_1}) \cap S(L_{Bl_2}) \cap S(L_{Bl_i}) \ldots,
\]

(32)

\[
S_{mL} = S(L_{Bl_1}) \cup S(L_{Bl_2}) \cup S(L_{Bl_i}) \ldots.
\]

(33)

In the following sections, two examples of this semi-automated reduction algorithm are given to assess the capability of the pathway algorithm for a systematic reduction method for plasma chemistry.

6. Reduction of a \( \text{H}_2\text{O} \– \text{He} \) mixture in a microwave induced plasma over a range of power inputs

The \( \text{H}_2\text{O}–\text{He} \) mixture described in [32] is used as a test case in which the results at steady state matter. The chemistry set with 622 reactions (elastic collisions are not included in this number) and 56 species is run over a range of input powers from 300 W to 1100 W. The volume of the plasma is assumed constant and equal to \( 6.72 \times 10^{-5} \text{ m}^3 \). No gas heating due to electron dissociation reactions are included in the model. For further details please refer to [32, 35]. Two reduced sets are produced using the reduction tool.
Figure 3. Comparison of the two reduced models of H$_2$O–He to the full simulation at various power inputs for the electron density and the temperature at steady state. In the reduced model 1: 450 reactions out of 622 and 13 species out of 56 are removed, and the reduced model 2: 505 reactions out of 622 and 22 species out of 56 are removed.

- Set 1: 450 reactions out of 622 and 13 species out of 56 are removed.
- Set 2: 505 reactions out of 622 and 22 species out of 56 are removed.

It is shown in our attempts that the quality of the reduced model will be lost by removing more species and/or more reactions. The results of the reduced models are compared to the full model simulations. Figure 3 shows the comparison of the electron density and temperature at steady state between full and reduced models. The relative errors are shown on the right-hand side of the figure.

The pathway analysis is done at steady state. If the temporal behavior of the model is of interest, the pathway analysis should be done at all high gradient moments of a simulation. The maximum relative error for the electron density and temperature in these two reduced models are less than 25% and 3%, respectively. Figure 4 shows the comparison of the gas mixture and one of the most important ions in this chemistry H$_3$O$^+$. Although the electron density and temperature are predicted very well with both reduced models, H$_3$O$^+$ has a maximum relative error of 25% and 144% in the reduced models one and two, respectively. The reason can be due to the removal of all big clusters of H$_2$O and their reactions from the system. It is interesting to note that the variation on the error for the main positive ion density is an order of magnitude higher than the one found for the electrons. This can be explained by the fact that recombination rates for molecular ions, and particularly clusters are all very similar, meaning that the predicted electron loss rates are not affected by a decrease in accuracy of the predicted ions densities ratio.

Figure 5 shows the comparison of H$_2$O$_2$ and O between the full and reduced models. One can see that the increase of error in the prediction of H$_2$O$_2$ densities is large compared to the one of O atoms which are another important species and it indicates that the main pathways for those two species are weakly coupled. If one would be interested to look only at the production of O atoms in this discharge, the model could be further reduced at the cost of a lower accuracy for other species such as the positive ions and H$_2$O$_2$.

The species that are removed from the system for the two reduced sets are listed below:

- Reduced set 1: He$^+$, He$^+_2$, He$^+_3$, O$_2^-$, O$_3$, H$_2$O$_2^-$, H$_3$O$_2^-$, H$_2$O$_3^-$, H$_2$O$_2^+$, H$_2$O$_3^+$, H$_2$(a$^3\Sigma^+$)
- Reduced set 2: reduced set 1 and HeH$^+$, H$_2^+$, H$_3^+$, O$^+$, OH$^+$, H$_2$O$_2^+$, H$_2$(v = 3), O$_2$+, H$_2$O$_3^+$, H$_2$O$_2^+$.

In addition, lists of species and reactions in the reduced model 2 are presented in the appendix D. The steady state densities of H$_2$O$_2$ and radical O are predicted well by the reduced models. The conclusion is that group reduction for H$_2$O–He at constant power microwave induced plasma conditions shows satisfactory results.
Figure 4. Comparison of the steady state mixture temperature and the equilibrium value of $\text{H}_3\text{O}^+$ between the full and the two reduced models in the $\text{H}_2\text{O}$–He mixture. In the reduced model 1: 450 reactions out of 622 and 13 species out of 56 are removed, and the reduced model 2: 505 reactions out of 622 and 22 species out of 56 are removed.

Figure 5. Comparison of the steady state value of $\text{H}_2\text{O}_2$ and O for the $\text{H}_2\text{O}$–He mixture under the microwave induced plasma conditions. In the reduced model 1: 450 reactions out of 622 and 13 species out of 56 are removed, and the reduced model 2: 505 reactions out of 622 and 22 species out of 56 are removed.
Figure 6. Comparison of the electron density variations for a pure H$_2$ chemistry between the reduced set in which 274 reactions and 18 species are removed from the system and the full chemistry. The picture on the right side shows the relative error of the reduced model in comparison with the full model.

7. Reduction of a H$_2$ chemistry for a pulsed power plasma

In order to investigate the performance of the reduction tool in pulsed power systems, a pure H$_2$ chemistry at constant 300 K gas temperature with 306 reactions and 41 species is studied [36–38]. The hydrogen atom up to its fifth electronic excitation state with energy 15.29 eV and hydrogen molecule up to Rydberg sum electronic excitation state with energy 15.2 eV besides its vibrational excitations till level 15 are included in the model. H$^+$, H$_2^+$ and H$_3^+$ are ions and H$^-$ is the only anion in the model. In addition, vibrational–vibrational and vibrational–translational reactions are considered up to the dissociation level (V = 15).

The input power has a sinusoidal form with the following function:

$$P = \begin{cases} \frac{W}{m^3} = 0, & \text{if } \sin(2\pi ft) < 0, \\ \sin(2\pi ft)10^{+7}, & \text{if } \sin(2\pi ft) > 0. \end{cases}$$

where $t$ is time and $f = 4$ kHz is the frequency. Figures 6(b) and 7(b) show some results for two reduction sets for the above mentioned conditions by the developed tool.

- Set 1: 274 reactions out of 306 and 18 species out of 41 are removed.
- Set 2: 260 reactions out of 306 and 14 species out of 41 are removed.

Both reductions are done at 4.35 ms of the simulation when the electron density has the maximum value.

The second reduction shown in the figure 7(a) is better able to capture the peak of the electron density only by adding four more species and 14 more reactions to the first reduced set, shown in figure 6(a).

The species that are removed from the system for the two reduces set are listed below:

- Reduced set 1: H$_2$(b$^3\Sigma_u^+$), H$_2$(G$^1\Sigma_g^+$), H$_2$(D$^1\Pi_u$), H$_2$(B$^{'+\Pi}$), H$_2$(Rydberg), H$_2$(v = 4) – H$_2$(v = 15)
- Reduced set 2: H$_2$(b$^3\Sigma_u^+$), H$_2$(G$^1\Sigma_g^+$), H$_2$(B$^{'+\Pi}$), H$_2$(v = 5) – H$_2$(v = 15).

In addition, lists of species and reactions in the second reduced set are presented in the appendix E. Reactions that are added to the second reduced model in comparison with the first reduced set are shown with asterisk next to the reaction number.
Figure 7. Comparison of the electron density variations for a pure H$_2$ chemistry between the reduced set in which 260 reactions and 14 species are removed from the system and the full chemistry. The picture on the right side shows the relative error of the reduced model in comparison with the full model.

The ignition (first initial cycles) has a higher relative error in the first reduction (274 reactions and 18 species are removed), but after a couple of cycles, the reduced model predicts the behavior of the system quite well. It can be concluded from the reduced sets that only low vibrational states contribute to energy losses of electrons, which is expected for typical DBD conditions. In addition, low vibrational states up to H$_2$(v = 4) play significant role in production of negative ion, H$^-$, so without those reactions and also H$_2$(v = 4) in the first reduced set, the predicted electron density shows higher values (larger relative error) in comparison with the full model’s results. Adding more electronic excitation levels of hydrogen molecule and its direct ionization reaction also help to produce more accurate electron energy balance and distribution in the second reduced set.

In general, reduction for pulsed systems can be more challenging than for constant power systems, but with this example, it is shown that it is possible to reduce systems with high temporal variations with the developed semi-automated tool.

8. Conclusions

The presented algorithm by Lehmann [27] to determine chemical pathways for complex chemistries is analyzed and clarified in this study with more examples. It is shown that important parameters $D$ and $\delta$, which are relevant for the calculation of pathways rates, should be updated based on the results of the previous step, not the initial values. In addition, it is shown that it is possible that the sign of $\delta$ for a species is changed during the algorithm’s steps. This means that at some point during the analysis, for example, the net production of a species can be positive based on active and insignificant pathways of the previous branching species, while the species has a negative net production in the initial step of the algorithm. With a correct update of $D$, $\delta$, and the sign of $\delta$ from the results of the previous step, the algorithm is able to keep both conservation of rates for each reaction and conservation of net production of species in the whole analysis.

Three different ways to handle insignificant pathways are discussed and illustrated with examples. It is shown that, although, the conservation of rate and net production of species are necessary for a suitable algorithm, they are not enough for a proper representation of the initial chemistry set. Handling pathways in such a way that the effects of deleted (insignificant) pathways are taken into account properly in rate calculations is the best way to represent the initial chemical set. This part of the algorithm of Lehmann [27] is explained clearly, and it is demonstrated:
• How conservation can be satisfied even though the insignificant pathways are deleted from memory.
• How to connect active pathways to insignificant pathways without knowing about the stoichiometry and the multiplicity of the insignificant pathways.
• How the deleted information equation (21) must be updated properly after this type of connection (active pathways and insignificant—pseudo—pathways) is made.

In the second part, the capability of this algorithm to be used as a systematic reduction method is investigated by applying the developed semi-automated tool to two chemistry sets: a H₂O–He mixture in a microwave induced plasma over a range of input powers and a pure H₂ chemistry with VV and VT reactions for a sinusoidal input power density. The examples show the capability of the reduction method for non-equilibrium plasma systems where a significant reduction of the number of species and reactions rates are obtained with a minimum loss of accuracy. In addition, we note that we did not observe any impact of the addition of the electron energy balance equation to the pathways analysis, (see Koelman et al [34]), on the results of chemical reduction for the two complex chemistry models tested in this study. This is most likely related to the very large number of processes which are entangled and carry already all the necessary information. The suppression of any important pathway affecting the electron energy balance would in any case strongly change the overall power balance of the system and lead to unphysical results.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgment

This research is supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) and Alliander N.V. in the framework of Project 13581, PlasmaPower2Gas (PP2G): efficient electrical to chemical energy conversion via synergy effects in plasma and catalytic technology. Special thanks to Plasma Matters B.V. for providing their detailed hydrogen model as a test case for the developed reduction tool.

Appendix A. Chemical equilibrium set: example

In this appendix, a simple example consisting of 5 reactions and 5 species in chemical equilibrium is used to assess the capability of the algorithm to satisfy equations (3) and (15). Note that A is a hypothetical species that can be ionized to A⁺ and electronically excited to different levels A** and A***. e represents the electron, and rates of reactions are designed in such way that the purpose of the example is satisfied.

### The reactions with their rates are:

| No. | Reaction | Rate |
|-----|----------|------|
| r₀  | A → A**  | R₀ = 10 |
| r₁  | A + A** → 2A⁺ | R₁ = 10 |
| r₂  | A⁺ + e → A  | R₂ = 10 |
| r₃  | A → A⁺ + e  | R₃ = 10 |
| r₄  | 4A⁺ → 4A   | R₄ = 5  |

Note that the simplest possible way to write a reaction is used above. For example, reaction r₃ should have one extra electron on both sides. However, in pathway analysis the net reaction is used. At the start of the algorithm, we have the following information for net production, destruction, and D for all included species in addition to the rate of each reaction.

| Species | p  | d  | D  | δ  |
|---------|----|----|----|----|
| A       | 30 | 30 | 30 | 0  |
| A⁺      | 20 | 20 | 20 | 0  |
| A**     | 10 | 10 | 10 | 0  |
| e       | 10 | 10 | 10 | 0  |

Clearly, we have a chemical equilibrium case, all net of concentration changes for species are zero.

#### A.1. First branching species A**

For the first branching species, if we choose A**, we will have:

| No.   | Multiplicity | Net    | Rate |
|-------|--------------|--------|------|
| P₂    | 1 × r₂      | A⁺ + e → A | R₂ = 10 |
| P₃    | 1 × r₃      | A → A⁺ + e | R₃ = 10 |
| P₄    | 1 × r₄      | 4A⁺ → 4A  | R₄ = 5  |
| P₅    | 1 × r₀ + 1 × r₁ | 2A → 2A⁺ | R₅ = 10 |

And production, destruction, and net for species:

| Species | p  | d  | D  | δ  |
|---------|----|----|----|----|
| A       | 30 | 30 | 30 | 0  |
| A⁺      | 20 | 20 | 20 | 0  |
| A**     | 0  | 0  | 0  | 0  |
| e       | 10 | 10 | 10 | 0  |

Because A** is chosen as the branching species, this species is removed from the system. Therefore, it is expected that the production and destruction for this species become zero which is not consistent with the starting conditions. This is, however, not a problem, the algorithm should behave that way. For the other species, Dᵢ, pᵢ, and δᵢ are the same as the initial condition. The rate of each reaction is distributed among pathways in such a way that equation (15) is satisfied.
words, we have conservation of rate. Also note that we initialized the algorithm with a chemical equilibrium system.

| No. | Reaction | The RHS of (15) |
|-----|----------|----------------|
| r_0 | A → A^{**} | 10 |
| r_1 | A + A^{**} → 2A^* | 10 |
| r_2 | A^* + e → A | 10 |
| r_3 | A → A^* + e | 10 |
| r_4 | 4A^* → 4A | 5 |

### A.2. Second branching species A

For this iteration we have:

| No. | Multiplicity | Net Rate |
|-----|--------------|----------|
| P_0 | 1 × r_2 + 1 × r_3 | Null |
| P_1 | 1 × r_0 + 1 × r_1 + 2 × r_2 | 2A^* + 2e → 2A^* |
| P_2 | 4 × r_2 + 1 × r_4 | 4A^* → 4A^* + 4e |
| P_3 | 2 × r_0 + 2 × r_1 + 1 × r_4 | Null |

For rate calculation, the D value is set to 30, which is the same as the initial set for A, resulting in:

| Species | p | d | D | δ |
|---------|---|---|---|---|
| A       | 0 | 0 | 0 | 0 |
| A^*     | 6.66667 | 6.66667 | 6.66667 | 0 |
| A^{**}  | 0 | 0 | 0 | 0 |
| A^+     | 6.66667 | 6.66667 | 6.66667 | 0 |
| e       | 6.66667 | 6.66667 | 6.66667 | 0 |

The generated pathways uphold the equilibrium condition (zero net for all species). The total production and destruction for species that have not been branching species are not equal to the initial condition. However, we do have rate conservation:

| No. | Reaction | The RHS of (15) |
|-----|----------|----------------|
| r_0 | A → A^{**} | 10 |
| r_1 | A + A^{**} → 2A^* | 10 |
| r_2 | A^* + e → A | 10 |
| r_3 | A → A^* + e | 10 |
| r_4 | 4A^* → 4A | 5 |

### A.3. Third branching species A^*

For this iteration we have:

| No. | Multiplicity | Net Rate |
|-----|--------------|----------|
| P_6 | 1 × r_2 + 1 × r_3 | Null |
| P_7 | 2 × r_0 + 2 × r_1 + 1 × r_4 | Null |
| P_8 | 2 × r_0 + 2 × r_1 + 1 × r_4 | Null |

The value of parameter D from the previous step is different from the initial step for A^*. Now, the question is which D should be used for rate calculation in order to have conservation of rate. This point is not mentioned clearly by Lehmann [27]. In order to have conservation of rate, we should use the D value of the branching species from the last step. In other words, in each step, we replace the previous chemical set (combination of pathways) with a new set, so we should use the updated values for our next branching species, not the initial values. Production, destruction, and net for species:

| Species | p | d | D | δ |
|---------|---|---|---|---|
| A       | 0 | 0 | 0 | 0 |
| A^*     | 0 | 0 | 0 | 0 |
| A^{**}  | 0 | 0 | 0 | 0 |
| A^+     | 0 | 0 | 0 | 0 |
| e       | 0 | 0 | 0 | 0 |

At this point, all the species are recycled in the pathways, so there is no net production and destruction for any of them. This behavior is expected from the algorithm. The rates for reactions are still conserved based on equation (15).

| No. | Reaction | The RHS of (15) |
|-----|----------|----------------|
| r_0 | A → A^{**} | 10 |
| r_1 | A + A^{**} → 2A^* | 10 |
| r_2 | A^* + e → A | 10 |
| r_3 | A → A^* + e | 10 |
| r_4 | 4A^* → 4A | 5 |

### Appendix B. All possible connections between active and insignificant pathways: example

This is the most memory intensive method among the three introduced ways to handle insignificant pathways. The only advantage of this method is that it can be helpful for post-processing of the data, because pathways are divided into important (active pathways with rates above the defined threshold), and insignificant ones. In addition, another subtle point for keeping conservation will be shown in this example which is not clear from the description of Lehmann [27]. The initial set of chemistry for this example contains the following reactions:

| No. | Reaction | Rate |
|-----|----------|------|
| r_0 | A^- → Ar + e | R_0 = 10 |
| r_1 | A + e → A^* + e | R_1 = 5 |
| r_2 | A^* + e → A + e | R_2 = 20 |
| r_3 | A + e → A^* + 2e | R_3 = 15 |
| r_4 | A^* + 2e → A + e | R_4 = 5 |
| r_5 | A + A^* → A^* + A + e | R_5 = 10 |
| r_6 | A + e → A^{**} + e | R_6 = 18 |
| r_7 | A + e → A^- | R_7 = 6 |
| r_8 | A^- + e → A^-- | R_8 = 10 |
| r_9 | A^- + A^-- → 2A^- | R_9 = 12 |

Note that again, A is a hypothetical species that has positive ion A^+, negative ion A^- and A^{--}, and electronically excited levels A^* and A^{**}, while e represents electron.
Production, destruction, \( D \) parameter, and net for each species at the beginning of the algorithm have the following values:

| Species | \( p \) | \( d \) | \( D \) | \( \delta \) |
|---------|--------|--------|--------|--------|
| A       | 59     | 44     | 59     | 15     |
| A\(^+\) | 5      | 30     | 30     | -25    |
| A\(^*\) | 25     | 17     | 25     | 8      |
| A\(^{**}\)| 18     | 0      | 18     | 18     |
| A\(<\) | 6      | 32     | 32     | -26    |
| A\(<\)\(^*\)| 10     | 0      | 10     | 10     |
| e       | 35     | 21     | 35     | 14     |

The threshold value is set equal to 3.0 for this example.

### B.1. First branching species A\(^*\)

The list of active pathways after this branching species is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| \( P_0 \) | 1 \( \times \) \( r_0 \) | \( A^- \rightarrow A + e \) | \( R_0 = 10 \) |
| \( P_1 \) | 1 \( \times \) \( r_1 \) | \( A \rightarrow e + A^+ \) | \( R_1 = 15 \) |
| \( P_4 \) | 1 \( \times \) \( r_4 \) | \( A^+ + e \rightarrow A \) | \( R_4 = 5 \) |
| \( P_5 \) | 1 \( \times \) \( r_5 \) | \( A \rightarrow A^{**} \) | \( R_5 = 18 \) |
| \( P_7 \) | 1 \( \times \) \( r_7 \) | \( A + e \rightarrow A^- \) | \( R_7 = 6 \) |
| \( P_8 \) | 1 \( \times \) \( r_8 \) | \( A^- + e \rightarrow A^{--} \) | \( R_8 = 10 \) |
| \( P_9 \) | 1 \( \times \) \( r_9 \) | \( A^+ + A^+ \rightarrow 2A \) | \( R_9 = 12 \) |
| \( P_{10} \) | 1 \( \times \) \( r_1 + 1 \times \) \( r_2 \) | Null | \( R_{10} = 3.3333 \) |

Because A\(^*\) has a negative net production in the previous step (initial condition), we should take into account the consumptive pathways that have impact on the concentration changes for this species. They are:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| \( \Delta P_2 \) | 1 \( \times \) \( r_2 \) | \( A^* \rightarrow A \) | \( \Delta R_2 = 16.6667 \) |
| \( \Delta P_3 \) | 1 \( \times \) \( r_3 \) | \( A^* \rightarrow A^+ + e \) | \( \Delta R_3 = 8.3333 \) |

At this step, we also have a pathway with a rate below the specified threshold:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| \( P_{11} \) | 1 \( \times \) \( r_1 + 1 \times \) \( r_3 \) | \( A \rightarrow A^+ + e \) | \( R_{11} = 1.6667 \) |

The balance of production and consumption for each species from all of these three types of pathways are:

| Species | \( p \) | \( d \) | \( \hat{d} \) | \( \Delta p \) | \( \Delta d \) | \( \Delta \hat{d} \) | \( \Delta \delta \) |
|---------|--------|--------|--------|--------|--------|--------|--------|
| A       | 39     | 39     | 0      | 1.6667 | 16.6667 | 0      | 0      | 0      | 15     |
| A\(^+\) | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 25     | -25    |
| A\(^*\) | 15     | 17     | 1.6667 | 0      | 8.3333 | 0      | 0      | 0      | 8      |
| A\(^{**}\)| 18     | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 8      |
| A\(<\) | 6      | 32     | 0      | 0      | 0      | 0      | 0      | 0      | -26    |
| A\(<\)\(^*\)| 10     | 0      | 0      | 0      | 0      | 0      | 0      | 0      | 10     |
| e       | 25     | 21     | 1.6667 | 0      | 8.3333 | 0      | 0      | 0      | 14     |

where \( p, \hat{d}, \tilde{d} \) and \( \tilde{\delta} \) stand for production and destruction rates from active and insignificant pathways, respectively. In addition, \( \Delta p, \Delta d, \Delta \hat{d}, \) and \( \Delta \hat{\delta} \) represent the production and destruction for each species due to the active and insignificant pathways which have contributions to density changes.

In addition, we can check the balance for each reaction involved in the model and the contribution of the three different types of pathways:

| No. | \( R \) | \( \bar{R} \) | \( \Delta R \) | \( \Delta \bar{R} \) | \( R_i \) |
|-----|--------|--------|--------|--------|------|
| \( r_0 \) | 10     | 0      | 0      | 0      | 10   |
| \( r_1 \) | 3.3333 | 1.6667 | 0      | 0      | 5    |
| \( r_2 \) | 3.3333 | 0      | 16.6667| 0      | 20   |
| \( r_3 \) | 15     | 0      | 0      | 0      | 15   |
| \( r_4 \) | 5      | 0      | 0      | 0      | 5    |
| \( r_5 \) | 0      | 1.6667 | 8.3333 | 0      | 10   |
| \( r_6 \) | 18     | 0      | 0      | 0      | 18   |
| \( r_7 \) | 6      | 0      | 0      | 0      | 6    |
| \( r_8 \) | 10     | 0      | 0      | 0      | 10   |
| \( r_9 \) | 12     | 0      | 0      | 0      | 12   |

Where \( R, \bar{R}, \) and \( R_i \) represent the portions of the rate for each reaction due to active and insignificant pathways, and the total rate of this specific reaction, respectively. In addition, \( \Delta R \) and \( \Delta \bar{R} \) stand for portions of the rate for each reaction from active and insignificant pathways which have contributions to density changes.

### B.2. Second branching species A\(^+\)

Before showing the produced pathways after using A\(^+\) as branching species, one important point should be clarified. In the previous sections, it has mentioned that for the calculation of \( D \) and \( \delta \) for the current branching species, the value of these parameters should be set based on the results of active and (in this case) insignificant pathways of the previous step (the effects of pathways which have contributions to density changes should not be taken into account in the calculation of those parameters). However, in all of the previous examples the sign of the net production of the current branching species is the same as its sign in the initial condition. In other words, if a species is produced in the initial condition, its sign always remains positive before it becomes the branching species, although the value changes. It was the case also for species which are consumed in the initial conditions.

In this example, we have a special case, the net change of A\(^+\), the current branching species, from active and insignificant pathways of the previous step is -0.33333. However, in the initial set, the net change for A\(^+\) is 8. This change in sign has an impact on the way we choose pathways which have contributions to density changes. This aspect will be illustrated here. Let us start with showing the results of active and insignificant pathways for this step. The list of active pathways:
In the corrected version, we do not have any insignificant pathways that have contributions to density changes, because the net production of $A^+$ is negative in the previous step, and the only insignificant pathway that is created in the previous step produces $A^+$. The balances for species and reactions are:
The list of active pathways that have contributions to density changes is:

| No. | Multiplicity | Net  | Rate |
|-----|--------------|------|------|
| $\Delta P_2$ | $1 \times r_2$ | $A' \rightarrow A$ | $\Delta R_2 = 16.6667$ |
| $\Delta P_4$ | $1 \times r_4$ | $A^+ \rightarrow A^+ + e$ | $\Delta R_4 = 8.3333$ |
| $\Delta P_5$ | $1 \times r_5$ | $A^+ + e \rightarrow A$ | $\Delta R_5 = 0.0980392$ |
| $\Delta P_9$ | $1 \times r_9$ | $A^- + A^+ \rightarrow 2A$ | $\Delta R_9 = 0.2352294$ |
| $\Delta P_{10}$ | $1 \times r_{10}$ | $A^- + A^+ \rightarrow A + e$ | $\Delta R_{10} = 2.64865$ |
| $\Delta P_{13}$ | $1 \times r_3 + 1 \times r_6$ | $A^- \rightarrow A + e$ | $\Delta R_{13} = 2.80445$ |

In addition, there is an insignificant pathway that has an impact on density changes:

| No. | Multiplicity | Net  | Rate |
|-----|--------------|------|------|
| $\Delta P_{15}$ | $1 \times r_1 + 1 \times r_5 + 1 \times r_9$ | $A^- \rightarrow A + e$ | $\Delta R_{15} = 0.311606$ |

As is clear from the pathways that have contributions to density changes, pathways that produce $e$ are selected, because the sign of the net production of $e$ from active and insignificant pathways from the previous step is positive (for this branching species, the net sign is the same as the initial set). The balances for species and reactions are:

| Species | $\Delta \tilde{d}$ | $\delta$ |
|---------|-----------------|----------|
| $A$     | 0               | 15       |
| $A^+$   | 0               | -25      |
| $A^-$   | 0               | 8        |
| $A^{**}$| 0               | 18       |
| $A''$   | 0               | -26      |
| $A_0$   | 0               | 10       |
| $e$     | 0               | 14       |

B.3. Third branching species $e$

The list of active pathways that have been produced in this step is:

| No. | Multiplicity | Net  | Rate |
|-----|--------------|------|------|
| $P_{10}$ | $1 \times r_6$ | $A \rightarrow A^{**}$ | $R_{10} = 18$ |
| $P_{12}$ | $1 \times r_1 + 1 \times r_2$ | Null | $R_{12} = 4.41176$ |
| $P_{13}$ | $1 \times r_3 + 1 \times r_4$ | Null | $R_{13} = 6.0$ |
| $P_{14}$ | $1 \times r_5 + 1 \times r_6$ | $2A^- \rightarrow A + A^-$ | $R_{14} = 4.59459$ |
| $P_{15}$ | $1 \times r_7 + 1 \times r_8$ | $2A^- \rightarrow A + A^-$ | $R_{15} = 4.86486$ |

The list of insignificant pathways is:

| No. | Multiplicity | Net  | Rate |
|-----|--------------|------|------|
| $P_{14}$ | $1 \times r_7 + 1 \times r_8 + 1 \times r_9$ | Null | $R_{14} = 0.490196$ |
| $P_{16}$ | $1 \times r_9 + 1 \times r_{10}$ | Null | $R_{16} = 2.75676$ |
| $P_{18}$ | $1 \times r_9 + 1 \times r_{10} + 1 \times r_{11}$ | Null | $R_{18} = 2.91892$ |
| $P_{20}$ | $1 \times r_9 + 1 \times r_{10} + 1 \times r_{12}$ | Null | $R_{20} = 0.324324$ |
| $P_{21}$ | $1 \times r_9 + 1 \times r_{10} + 1 \times r_{11}$ | $2A^- \rightarrow A + A^-$ | $R_{21} = 0.540541$ |
| $P_{23}$ | $1 \times r_9 + 1 \times r_{10} + 1 \times r_{12}$ | Null | $R_{23} = 0.66667$ |

The list of active pathways that have contributions to density changes is:

| No. | Multiplicity | Net  | Rate |
|-----|--------------|------|------|
| $r_0$ | $4.59459$ | $2.75676$ | $2.64865$ | $0$ |

As expected, this method maintains conservation of rates and density changes. The disadvantage of this type of
implementation is memory usage. Since the insignificant pathways are not deleted from the memory (the memory is not released), this algorithm can easily result in a lack of memory for complex chemistries.

Appendix C. Connection of only active pathways: example

For this method, another set of chemistry is designed since it allows a more clear explanation of the method 4.2. The initial set contains the following reactions:

| No. | Reaction | Rate |
|-----|----------|------|
| $r_0$ | $A \rightarrow A(v=1)$ | $R_0 = 5$ |
| $r_1$ | $A(v=1) \rightarrow A(v=2)$ | $R_1 = 10$ |
| $r_2$ | $2A(v=2) \rightarrow A + A(v=1)$ | $R_2 = 15$ |
| $r_3$ | $A(v=1) \rightarrow A(v=3)$ | $R_3 = 20$ |
| $r_4$ | $A(v=4) \rightarrow A(v=2)$ | $R_4 = 80$ |
| $r_5$ | $A(v=2) \rightarrow A(v=5)$ | $R_5 = 100$ |
| $r_6$ | $A(v=2) \rightarrow A(v=3)$ | $R_6 = 15$ |

Note that A is a hypothetical species that can have different vibrational levels. In addition, the species production, destruction, and their net in the initial step are:

| Species | $p$ | $d$ | $\tilde{p}$ | $\tilde{d}$ | $\Delta p$ | $\Delta d$ | $\delta$ |
|---------|-----|-----|----------|----------|---------|---------|-----|
| A       | 15  | 5   | 15       | 10       | 5       | 0       | 10  |
| A(v=1)  | 20  | 30  | 30       | -10      | 0       | 0       | -10 |
| A(v=2)  | 90  | 145 | 145      | -55      | 0       | 0       | -55 |
| A(v=3)  | 35  | 0   | 35       | 35       | 0       | 0       | 35  |
| A(v=4)  | 0   | 80  | 80       | -80      | 0       | 0       | -80 |
| A(v=5)  | 100 | 0   | 100      | 100      | 0       | 0       | 100 |

The threshold value is set to 11 for this example. We are going to connect only active pathways and store the deleted portion of the rate for each reaction besides the deleted production and destruction rates for each species.

C.1. First branching species $A(v = 1)$

For the first branching species, there is no difference between this method and the others because the whole initial set of the chemistry is known as active pathways. The list of the active pathways for this step is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_4$ | $1 \times r_4$ | $A(v=4) \rightarrow A(v=2)$ | $R_4 = 80$ |
| $P_5$ | $1 \times r_5$ | $A(v=2) \rightarrow A(v=5)$ | $R_5 = 100$ |
| $P_6$ | $1 \times r_6$ | $A(v=2) \rightarrow A(v=3)$ | $R_6 = 15$ |

The list of pathways that have contributions to density changes is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $\Delta P_1$ | $1 \times r_1$ | $A(v=1) \rightarrow A(v=2)$ | $\Delta R_1 = 3.33333$ |
| $\Delta P_2$ | $1 \times r_2$ | $A(v=1) \rightarrow A(v=3)$ | $\Delta R_2 = 6.66667$ |

The advantage of this way of connecting is that there is no need to keep insignificant pathways in memory. As soon as a pathway with a rate lower than the threshold is produced, the deleted information of equation (21) will be updated, and the pathway will be erased from memory. However, here, just for illustration of the method, the list of produced insignificant pathways in each step is shown. The list of insignificant produced pathways is:

| No. | Multiplicity | Net | Rate |
|-----|--------------|-----|------|
| $P_7$ | $1 \times r_1 + 1 \times r_1$ | $A \rightarrow A(v=2)$ | $R_7 = 1.6667$ |
| $P_8$ | $1 \times r_2 + 1 \times r_1$ | $A \rightarrow A(v=3)$ | $R_8 = 3.33333$ |
| $P_9$ | $1 \times r_1 + 1 \times r_2$ | $A(v=2) \rightarrow A$ | $R_9 = 5$ |
| $P_{10}$ | $1 \times r_2 + 1 \times r_1$ | $2A(v=2) \rightarrow A + A(v=3)$ | $R_{10} = 10$ |

Now the balances for species production and reaction rates are:

| Species | $p$ | $d$ | $\tilde{p}$ | $\tilde{d}$ | $\Delta p$ | $\Delta d$ | $\delta$ |
|---------|-----|-----|----------|----------|---------|---------|-----|
| A       | 0   | 0   | 15       | 5        | 0       | 0       | 10  |
| A(v=1)  | 0   | 0   | 0        | 0        | 0       | 0       | -10 |
| A(v=2)  | 80  | 115 | 1.66667  | 25       | 3.33333 | 0       | -55 |
| A(v=3)  | 15  | 0   | 13.3333  | 6.66667  | 0       | 0       | 35  |
| A(v=4)  | 0   | 80  | 0        | 0        | 0       | 0       | -80 |
| A(v=5)  | 100 | 0   | 0        | 0        | 0       | 0       | 100 |

As expected, we have conservation of rates and net production for each species. Nothing is different from the previous method.

C.2. Second branching species $A(v = 2)$

One important thing about the pathway algorithm is that the algorithm starts from the initial set and after each branching species, the previous set is replaced with the new one. Therefore, all calculations for $D, \delta$, and the sign of $\delta$ should be based on the new set.

In this example, we make a new chemistry set after branching species $A(v = 1)$, which satisfies conservation. Now, it is also possible to only look at a subset of this set, namely only the active pathways, and then, to create a new set of pathways based on this subset. In other words, after the first branching species, we have a list of active pathways, pathways that have contributions to density changes, and some deleted equation (21). The combination of these three sets satisfies conservation. Now, if in this step we manipulate the active pathways list, we will have a new list of active pathways, a
list for pathways that have contributions to density changes (which composed of the previous step’s pathways that have an impact on densities and a new list of pathways that have contributions to density changes), and updated equations for the deleted portions of the reaction rates and species production and destruction rates. If we correctly calculate the necessary parameters needed for the calculation of pathway rates and properly choose pathways that have contributions to density changes, we should be able to satisfy the conservation if we only look at a subset of the chemistry for this step.

In order to test and demonstrate this capability, for the current branching species, we only connect active pathways (a subset of the results of the previous step) and check the net balance for species and reaction rates. One important question is how the values for $D$, $\delta$, and the sign of the net production for branching species should be set. In the previous method, those parameters were calculated or chosen based on the active and insignificant pathways from the previous step because the newly produced set of pathways is a substitution for both active and insignificant pathways. In this case, $D$, $\delta$, and the sign of $\delta$ are calculated and chosen only from the active pathways of the previous step, because we are going to manipulate and find a new set for active pathways list and update the necessary information.

The list of active pathway after connection at this step is:

| No. | Multiplicity | Net       | Rate     |
|-----|--------------|-----------|----------|
| $P_{11}$ | $1 \times r_1 + 1 \times r_3$ | A($v = 4$) $\rightarrow$ A($v = 5$) | $R_{11} = 69.5652$ |

In this method, as mentioned in the previous step, we can erase the insignificant pathways from memory. Again, just for illustration, we show the list of insignificant pathways:

| No. | Multiplicity | Net       | Rate     |
|-----|--------------|-----------|----------|
| $P_7$ | $1 \times r_0 + 1 \times r_1$ | A $\rightarrow$ A($v = 2$) | $R_7 = 1.6667$ |
| $P_8$ | $1 \times r_0 + 1 \times r_3$ | A $\rightarrow$ A($v = 3$) | $R_8 = 3.3333$ |
| $P_9$ | $1 \times r_1 + 1 \times r_2$ | A($v = 2$) $\rightarrow$ A | $R_9 = 5$ |
| $P_{10}$ | $1 \times r_2 + 1 \times r_3$ | 2A($v = 2$) $\rightarrow$ A + A($v = 3$) | $R_{10} = 10$ |
| $P_{12}$ | $1 \times r_3 + 1 \times r_6$ | A($v = 4$) $\rightarrow$ A($v = 3$) | $R_{12} = 10.4348$ |

Clearly, although the branching species is A($v = 2$), there is no connection between insignificant pathways: $P_7$ with $P_9$ and $P_7$ with $P_{10}$. In addition, we do not have any connection between active pathways from the previous steps ($P_4, P_5, P_6$) with insignificant pathways. We only connect the active pathways, and if the rate of a new produced pathway is lower than the specified threshold, it will be deleted from memory, and deleted equation (21) will be updated. In this example, two pathways are produced from the list of active pathways, $P_{11}$ and $P_{12}$, one of them has a rate lower than the threshold, so it is added to the list of insignificant pathways, and deleted equation (21) are updated. In addition, based on the sign of the net change of A($v = 2$) from the list of active pathways ($p - d = -35$), the list of pathways that impact on density changes is updated.

The list of pathways that have an impact on density changes is:

| No. | Multiplicity | Net Rate |
|-----|--------------|----------|
| $\Delta P_1$ | $1 \times r_1$ | A($v = 1$) $\rightarrow$ A($v = 2$) | $\Delta R_1 = 3.3333$ |
| $\Delta P_3$ | $1 \times r_1$ | A($v = 1$) $\rightarrow$ A($v = 3$) | $\Delta R_3 = 6.6667$ |
| $\Delta P_4$ | $1 \times r_5$ | A($v = 2$) $\rightarrow$ A($v = 5$) | $\Delta R_5 = 30.4348$ |
| $\Delta P_6$ | $1 \times r_6$ | A($v = 2$) $\rightarrow$ A($v = 3$) | $\Delta R_6 = 4.5652$ |

The balances for reaction rates and net production of species are:

| Species | $p$ | $d$ | $\dot{p}$ | $\dot{d}$ | $\Delta \rho$ | $\Delta d$ | $\delta$ |
|---------|-----|-----|-----------|-----------|-------------|-----------|------|
| A       | 0   | 0   | 15        | 5         | 0           | 0         | 10   |
| A($v = 1$) | 0   | 0   | 0         | 0         | 10 - 0      | 10 - 0    | 0    |
| A($v = 2$) | 0   | 0   | 1.66667   | 25        | 33.3333    | 35 - 55   | 0    |
| A($v = 3$) | 0   | 0   | 23.7681   | 0         | 11.2391    | 0         | 0    |
| A($v = 4$) | 0   | 0   | 10.4348   | 4         | 35 - 80    | 0         | 0    |
| A($v = 5$) | 69.5652 | 0   | 0         | 30.4348   | 0         | 100    | 0    |

As is evident from the balances, conservation is satisfied. We can conclude that it is possible to take a subset of a chemistry set to find important pathways and also have conservation of rates and density changes. However, for the calculation of $D$, $\delta$, and sign of $\delta$, we have to take into account the effect of the subset that is going to be manipulated. Another subtle point is that although this method keeps both conservation of rate for each reaction and net production for each species, the calculated rates for connection of active pathways have a higher value than what should be for a set that is a proper representation of the initial chemistry set.

Appendix D. H$_2$O–He reduced model (set 2)

The list of species remained in the model after reduction is as follow, for more information about the model, please refer to [32]:

- He, H$_2$, H, H(n = 2), H(n = 3), H(n = 4), H$^+$, H$^-$, H$_2$(rot),
- H$_2$(v = 1), H$_2$(v = 2), H$_2$($b^1\Sigma^+_u$),
- O$_2$, O$_2$(v = 1), O$_2$(v = 2), O$_2$(v = 3), O$_2$(v = 4), O$_2$(a),
- O$_3$(b), O, O$^+$, O(1D), O(1S),
- OH, OH$^-$, OH(A),
- H$_2$O, H$_2$O(010), H$_2$O(100 + 001), HO$_2$, H$_2$O$_2$, H$_2$O$^+$, H$_2$O$^-$, e.
The list of reactions after reduction is as

| No. | Reaction |
|-----|----------|
| 1.  | H⁺ + He → H⁺ + e |
| 2.  | He + H(n = 3) → H⁺ + He |
| 3.  | He + O(1D) → O + He |
| 4.  | He + H + H → H₂ + He |
| 5.  | He + H + O → OH + He |
| 6.  | He + H + O₂ → He + HO₂ |
| 7.  | He + O₂(v = 1) → He + O₂ |
| 8.  | He + O₂(v = 2) → He + O₂ |
| 9.  | He + O₂(v = 3) → He + O₂ |
| 10. | He + O₂(v = 4) → He + O₂ |
| 11. | He + H₂(b³Σg⁺) → He + H₂ |
| 12. | He + H₂(v = 1) → He + H₂ |
| 13. | He + H₂(v = 2) → He + H₂ |
| 14. | He + H₂(rot) → He + H₂ |
| 15. | + H → H⁺ + e + e |
| 16. | + H₂O → H₂O(010) + e |
| 17. | + H₂O → H₂O(100 + 001) + e |
| 18. | + H₂O → H₂ + O + H + e |
| 19. | + H₂O → H₂ + O(1D) + e |
| 20. | + H₂O → H + O(H(A) + e |
| 21. | + H → H(n = 2) + e |
| 22. | + H → H(n = 3) + e |
| 23. | + H → H(n = 4) + e |
| 24. | + H₂ → H₂(rot) + e |
| 25. | + H₂(rot) → H₂ + e |
| 26. | + H₂ → H₂(v = 1) + e |
| 27. | + H₂(v = 1) → H₂ + e |
| 28. | + H₂ → H₂(v = 2) + e |
| 29. | + H₂ → H₂(b³Σg⁺) + e |
| 30. | + H₂(b³Σg⁺) → H₂ + e |
| 31. | + H₂ → H + H + e |
| 32. | + H(n = 4) → H(n = 3) + e |
| 33. | + O → O(1D) + e |
| 34. | + O(1D) → O + e |
| 35. | + O → O(1S) + e |
| 36. | + O₂ → O₂(v = 1) + e |
| 37. | + O₂(v = 1) → O₂ + e |
| 38. | + O₂ → O₂(v = 2) + e |
| 39. | + O₂(v = 2) → O₂ + e |
| 40. | + O₂ → O₂(v = 3) + e |
| 41. | + O₂ → O₂(v = 4) + e |
| 42. | + O₂ → O₂(a) + e |
| 43. | + O₂(a) → O₂ + e |
| 44. | + O₂ → O₂(b) + e |
| 45. | + O₂(b) → O₂ + e |
| 46. | + O₂ → O + O + e |
| 47. | + O₂ → O(1D) + O + e |
| 48. | + O₂(v = 1) → O₂(a) + e |
| 49. | + O₂(v = 1) → O₂(b) + e |
| 50. | + O₂(v = 2) → O₂(a) + e |
| 51. | + O₂(v = 3) → O₂(a) + e |
| 52. | + O₂(v = 4) → O₂(a) + e |
| 53. | + O₂(a) → O₂(b) + e |
| 54. | + O₂(b) → O₂(a) + e |
| 55. | + O₂(a) → O + O + e |
| 56. | + O₂(a) → O + O(1D) + e |

(Continued.)
Appendix E. H₂ reduced model (set 2)

The list of species remained in the model after reduction is as follow, for more information about the model, please contact [36]:

\[
\begin{align*}
H, H^+, H^-, H(2p), H(2s), H(3), H(4), H(5), \\
H_2, H_2 (B^1\Sigma_u^+), H_2 (c^1\Pi_u), H_2 (a^3\Sigma_g^+), \\
H_2 (C^1\Pi_u), H_2 (E^1\Sigma_u^+/F^1\Sigma_u^+), H_2 (e^3\Sigma_g^+), \\
H_2 (B^1\Sigma_u^+), H_2 (d^3\Pi_u), H_2 (D^3\Pi_u), H_2 (D^1\Pi_u), \\
H_2 (Rydberg), \\
H_2^+, H_3^+ H_2(v = 1), H_2(v = 2), H_2(v = 3), H_2(v = 4), \text{ etc.}
\end{align*}
\]

The list of reactions after reduction is as follow, the reactions with asterisk close to their numbers are added to the second reduction list in comparison with the first set.

| No. | Reaction |
|-----|----------|
| 1.  | e + H → e + H(2p) |
| 2.  | e + H → e + H(2s) |
| 3.  | e + H → e + H(3) |
| 4.  | e + H → e + H(4) |
| 5.  | e + H → e + H(5) |
| 6.  | e + H → 2e + H^+ |
| 7.  | e + H → e + 2H |
| 8.  | e + H → e + H_2 (B^1\Sigma_u^+) |
| 9.  | e + H_2 → e + H_2 (c^1\Pi_u) |
| 10. | e + H_2 → e + H_2 (a^3\Sigma_g^+) |
| 11. | e + H_2 → e + H_2 (C^1\Pi_u) |
| 12. | e + H_2 → e + H_2 (E^1\Sigma_u^+/F^1\Sigma_u^+) |
| 13. | e + H_2 → e + H_2 (e^3\Sigma_g^+) |
| 14. | e + H_2 → e + H_2 (B^1\Sigma_u^+) |
| 15. | e + H_2 → e + H_2 (d^3\Pi_u) |
| 16. | e + H_2 → e + H_2 (D^3\Pi_u) |
| 17. | e + H_2 → e + H_2 (Rydberg) |
| 18. | e + H_2 → e + H + H(2p) |
| 19. | e + H_2 → e + H + H(3) |
| 20. | e + H_2 → e + H + H(4) |
| 21. | e + H_2 → e + H + H(5) |
| 22. | e + H_2 → 2e + H^+ |
| 23. | e + H_2 → e + H_2 (v = 1) |
| 24. | e + H_2 → e + H_2 (v = 2) |
| 25. | e + H_2 → e + H_2 (v = 3) |
| 26. | e + H_2 → e + H_2 (v = 4) |
| 27. | e + H^+ → H + H(2s) |
| 28. | e + H^+ → e + H + + |
| 29. | e + H → 2e + H |

(Continued.)

ORCID iDs

S Tadayon Mousavi  https://orcid.org/0000-0002-6064-4938
W A A D Graef  https://orcid.org/0000-0002-5471-5912
P M J Koelman  https://orcid.org/0000-0002-9987-6875
E A D Carbone  https://orcid.org/0000-0003-3455-0708
J van Dijk  https://orcid.org/0000-0001-8002-5459

References

[1] Aerts R, Martens T and Bogaerts A 2012 Influence of vibrational states on CO₂ splitting by dielectric barrier discharges J. Phys. Chem. C 116 23257–73
[2] Kozák T and Bogaerts A 2014 Splitting of CO₂ by vibrational excitation in non-equilibrium plasmas: a reaction kinetics model Plasma Sources Sci. Technol. 23 045004
[3] Kozák T and Bogaerts A 2015 Evaluation of the energy efficiency of CO₂ conversion in microwave discharges using a reaction kinetics model Plasma Sources Sci. Technol. 24 015024
[4] Koelman P, Heijkers S, Mousavi S T, Graef W, Mihailova D, Kozák T, Bogaerts A and van Dijk J 2016 A comprehensive chemical model for the splitting of CO₂ in non-equilibrium plasmas Plasma Process. Polym. 14 1600155
[5] Berthelot A and Bogaerts A 2017 Modeling of CO₂ splitting in a microwave plasma: how to improve the conversion and energy efficiency J. Phys. Chem. C 121 8236–51
[6] Boselli M, Colombo V, Ghedini E, Gherardi M, Laurita R, Linguori A, Sanibondi P and Stancampiano A 2014 Schlieren high-speed imaging of a nanosecond pulsed atmospheric pressure non-equilibrium plasma jet Plasma Chem. Plasma Process. 34 853–69
[7] Zhang S, Sobota A, van Veldhuizen E M and Bruggeman P J 2015 Gas flow characteristics of a time modulated appj: the effect of gas heating on flow dynamics J. Phys. D: Appl. Phys. 48 015203
[8] Graef W A A D 2012 Zero-dimensional models for plasma chemistry PhD Thesis Eindhoven University of Technology, The Netherlands.
[9] Peerenboom K, Parente A, Kozák T, Bogaerts A and Degrez G 2015 Dimension reduction of non-equilibrium plasma kinetic models using principal component analysis Plasma Sources Sci. Technol. 24 025004

[10] Rehman T, Kemaneci E, Graef W and van Dijk J 2016 Simplifying plasma chemistry via ILDM J. Phys.: Conf. Ser. 682 012035

[11] Rehman T 2018 Studies on plasma-chemical reduction PhD Thesis Eindhoven University of Technology

[12] Koelman P M J 2019 Chemical aspects of CO2 plasma modelling PhD Thesis Technical University of Eindhoven

[13] Kemaneci E, Carbone E, Graef W, van Dijk J and Kroesen G M W 2015 Evaluation of a measure on the quasi-steady state assumption of collisional radiative models via intrinsic low dimensional manifold technique (arXiv:1511.09237)

[14] Varma A and Palsson B O 1994 Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110 Appl. Environ. Microbiol. 60 3724–31

[15] Schilling C H, Schuster S, Palsson B O and Heinrich R 1999 Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era Biotechnol. Prog. 15 296–303

[16] Price N D, Famili I, Beard D A and Palsson B O 2002 Extreme pathways and Kirchhoff’s second law Biophys. J. 83 2879–82

[17] Fishitik I, Callaghan C A and Datta R 2004 Reaction route graphs. I. Theory and algorithm J. Phys. Chem. B 108 5671–82

[18] Fishitik I, Callaghan C A and Datta R 2004 Reaction route graphs. II. Examples of enzyme- and surface-catalyzed single overall reactions J. Phys. Chem. B 108 5683–97

[19] Vijay P, Tadé M O, Fishitik I and Datta R 2013 A graph theoretical approach to the elucidation of reaction mechanisms: analysis of the chlorine electrode reaction Comput. Chem. Eng. 49 94

[20] Bai S, Zhou D, Davis M J and Skodje R T 2015 Sum over histories representation for chemical kinetics J. Phys. Chem. A 119 9582–91

[21] Bai S and Skodje R T 2016 The sum over histories representation for chemical kinetics: a quantitative theory based on chemical pathways Int. Rev. Phys. Chem. 35 539–67

[22] Sakai O, Nobuto K, Miyagi S and Tachibana K 2015 Analysis of weblike network structures of directed graphs for chemical reactions in methane plasmas AIP Adv. 5 107140

[23] Mizui Y, Kojima T, Miyagi S and Sakai O 2017 Graphical classification in multi-centrality-index diagrams for complex chemical networks Symmetry 9 309

[24] Murakami T and Sakai O 2020 Rescaling the complex network of low-temperature plasma chemistry through graph-theoretical analysis Plasma Sources Sci. Technol. 29 115018

[25] Sun S R, Wang H X and Bogaerts A 2020 Chemistry reduction of complex CO2 chemical kinetics: application to a gliding arc plasma Plasma Sources Sci. Technol. 29 025012

[26] Holmes T D, Rothman R H and Zimmerman W B 2021 Graph theory applied to plasma chemical reaction engineering Plasma Chem. Plasma Process. 41 531–57

[27] Lehmann R 2004 An algorithm for the determination of all significant pathways in chemical reaction systems J. Atmos. Chem. 47 45–78

[28] van Gaens W and Bogaerts A 2013 Kinetic modelling for an atmospheric pressure argon plasma jet in humid air J. Phys. D: Appl. Phys. 46 275201

[29] Liu D, Sun B, Iza F, Xu D, Wang X, Rong M and Kong M G 2017 Main species and chemical pathways in cold atmospheric-pressure Ar + H2O plasmas Plasma Sources Sci. Technol. 26 045009

[30] Johnston H and Kimsson D 1998 Methane photooxidation in the atmosphere: contrast between two methods of analysis J. Geophys. Res.: Atmos. 103 21967–84

[31] Markosyan A H, Luque A, Gordillo-Vázquez F J and Ebert U 2014 Pumpkin: a tool to find principal pathways in plasma chemical models Comput. Phys. Commun. 185 2697–702

[32] Tadayon Mousavi S, Carbone E A D, Wolf A J, Bongers W A and van Dijk J 2021 Two-temperature balance equations implementation, numerical validation and application to H2O–He microwave induced plasmas Plasma Sources Sci. Technol. 30 075007

[33] Lieberman M A and Lichtenberg A J 2005 Principles of Plasma Discharges and Materials Processing (New York: Wiley)

[34] Koelman P, Yordanova D, Graef W, Mousavi S T and van Dijk J 2019 Uncertainty analysis with a reduced set of input uncertainties selected using pathway analysis Plasma Sources Sci. Technol. 28 075009

[35] Tadayon Mousavi S 2020 Modeling of microwave induced plasmas: application to the production of solar fuels PhD Thesis Applied Physics, Proefschrift

[36] PLASIMO (available at: https://plasimo.phys.tue.nl/)

[37] van Dijk J, Peerenboom K, Jimenez M, Mihailova D and van der Mullen J 2009 The plasma modelling toolkit plasma J. Phys. D: Appl. Phys. 42 194012

[38] Gulpen J G M, Tadayon Mousavi S and van Dijk J 2019 A chemical reduction method based on pathway analysis MScThesis Technische Universiteit Eindhoven