

**Sum Over Histories Representation of Chemical Kinetics:
An Interpretive and Predictive Method for Modeling
Chemical Kinetics Using Time-dependent Pathways**

by

Shirong Bai

B.S., University of Science and Technology of China, 2012

M.S., University of Colorado Boulder, 2016

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy, Chemical Physics
Department of Chemistry and Biochemistry

2018

This thesis entitled:

Sum Over Histories Representation of Chemical Kinetics: An Interpretive and Predictive Method
for Modeling Chemical Kinetics Using Time-dependent Pathways

written by Shirong Bai

has been approved for the Department of Chemistry and Biochemistry

Prof. Rex T. Skodje

Prof. Robert Parson

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Bai, Shirong (Ph.D., Chemical Physics)

Sum Over Histories Representation of Chemical Kinetics: An Interpretive and Predictive Method
for Modeling Chemical Kinetics Using Time-dependent Pathways

Thesis directed by Prof. Rex T. Skodje

Chemical kinetics can be viewed as an intricate network of inter-related chemical reactions that work cooperatively to convert reagent species into product species. The network is in general time dependent reflecting the non-steady state nature of the chemistry. When it comes to interpreting and predicting chemical kinetics, the history of chemical moieties can play vital roles. In order to study the histories of chemical substances using time-dependent chemical network, this thesis focuses on developing a Sum Over Histories Representation (SOHR for short) of chemical kinetics. The description of time-dependent chemistry of a reaction network is provided by chemical pathways defined at a molecular level. Using this methodology, the quantitative time evolution of the kinetics is described by enumerating the most important pathways followed by a chemical moiety such as a tagged atom. An explicit formula for the pathway probabilities is derived which takes the form of an integral over a time-ordered product. This expression has a simple and computationally efficient Monte Carlo representation which permits the method to be applied to a wide range of problems.

In SOHR, the history of the chemical moiety can be described by time-dependent pathways. Unlike the static flux methods for path analysis, SOHR includes the explicit time-dependence of pathway probabilities. Using SOHR, the sensitivity of an observable with respect to a kinetic parameter such as a rate coefficient is then analyzed in terms of how that parameter affects the chemical pathway probabilities. This thesis demonstrates that large sensitivities are often associated with rate limiting steps along important chemical pathways or by reactions that control the branching of reactive flux, though they vary with time.

In addition to interpreting chemical kinetics, this thesis studies the practical approach to model-

ing chemical kinetics without solving conventional mass-action ODEs. An iterative framework was introduced that allows the time-dependent pathway probabilities to be generated from a knowledge of elementary rate coefficients. To avoid the pitfall of integrating over the histories of long paths, we proposed a sector-by-sector strategy that shortens the candidate path without losing numerical accuracy. This method was successfully applied to the model Lotka-Volterra system and to a realistic H₂ combustion system.

This thesis culminates with a discussion of the interpretative and predictive applicability of SOHR.

Dedication

To Zhangying Bai

Acknowledgements

Here's where you acknowledge folks who helped. But keep it short, i.e., no more than one page, as required by the Grad School Specifications.

Contents

Chapter

1	Introduction	1
2	Mathematical Development	4
2.1	Introduction	4
2.2	Defining Chemical Pathways	5
2.3	Pathway Probabilities	7
2.4	Constructing Observables	12
2.5	Graph Theoretic Enumeration of Chemical Pathways	13
2.6	Stochastic Pathway Enumeration	17
3	Linear Kinetics	18
3.1	Analytical Expression of Linear Kinetics	18
3.2	Number of Pathway vs. Pathway Length	20
3.3	Demonstration of Convergence of SOHR	21
3.4	Pathway Probabilities Changing Over Time	21
4	Kinetic Sensitivity Analysis Using Sum Over Histories Representation	25
4.1	Introduction	25
4.2	Methodology	28
4.2.1	Sum over Histories Representation	28

4.2.2	Sensitivity Analysis	30
4.3	Sensitivity Analysis of Hydrogen Combustion Chemistry	32
4.4	Description of Chemical Pathways for H ₂ -O ₂ Combustion	39
4.5	Pathway Interpretation of Sensitivity	48
4.6	Conclusions	59
5	Simulating Chemical Kinetics Without Differential Equations	61
5.1	Introduction	61
5.2	The Predictive SOHR Method	64
5.3	Pathway Generation	66
5.4	The Lotka-Volterra System	67
5.5	H ₂ -O ₂ Combustion	69
5.6	Utility of the Method	73
6	Conclusion	78
Bibliography		81
Appendix		
A	The Hydrogen Combustion Model	89

Tables**Table**

4.1 Species Labeling of H ₂ -O ₂ system	33
4.2 Reaction Labeling of H ₂ -O ₂ system	34
5.1 Lotka-Volterra System	77
A.1 Rate coefficients and uncertainty ranges of H ₂ -O ₂ system	91

Figures

Figure

2.1	Two representations of chemical histories	8
3.1	Chemical graph corresponding to an acyclic linear kinetic model	19
3.2	Number of path and path length distribution of linear model	20
3.3	The concentration from SOHR vs. concentration from ODEs	22
3.4	Pathway probabilities vs. time	23
3.5	The average path length vs. time	24
4.1	Concentrations versus time of H ₂ -O ₂ combustion system	35
4.2	Reaction rates versus time of H ₂ -O ₂ combustion system	36
4.3	Species decay rates versus time of H ₂ -O ₂ combustion system	36
4.4	Global sensitivity index W versus time of H ₂ -O ₂ combustion system	38
4.5	Oxygen subgraph of H ₂ -O ₂ combustion system	40
4.6	Pathway probabilities for H ₂ O ₂ of H ₂ -O ₂ combustion system	42
4.7	Graph for the oxygen following paths of H ₂ -O ₂ combustion system	43
4.8	Pathway probabilities for H ₂ O of H ₂ -O ₂ combustion system	44
4.9	Oxygen atom following pathways leading to the formation of the H ₂ O product	46
4.10	Pathway probabilities for HO ₂ of H ₂ -O ₂ combustion system	47
4.11	Oxygen atom following pathways leading to the formation of the HO ₂ product	48
4.12	Sensitivity indices of the pathways leading to the H ₂ O target concentration	51

4.13 Sensitivity indices for the five most important paths leading to the H ₂ O target	52
4.14 Correlation between pathways	55
4.15 1 st order component function of the pathway probabilities a1, b1, and c1	56
4.16 Sensitivity indices for the two dominant paths leading to H ₂ O ₂	58
5.1 A schematic representation of a time-dependent pathway	63
5.2 Kinetic graph for the Lotka-Volterra model	68
5.3 Convergence of the iteration for the predictive SOHR method in the LV system	70
5.4 Directed kinetic graphs for the forward reactions in H ₂ -O ₂ system	72
5.5 Convergence of the SOHR iteration for the H ₂ -O ₂ system	74
5.6 A comparison of the predictive SOHR method with the conventional ODE	75

Chapter 1

Introduction

Many problems of great practical importance are described by large interconnected networks of elementary chemical reactions.[54, 76, 91] In combustion chemistry, for example, numerical simulations combine chemical mechanisms with computational fluid dynamics to describe the consumption of fuels under engine operating conditions.[99, 101, 72, 9, 102] These mechanisms can grow quite large, e.g. n-heptane combustion can be modelled by a network of 7000 elementary gas phase reactions.[19] Likewise, the CAM-chem model of tropospheric chemistry can potentially consist of thousands of elementary chemical reactions.[56] Reactions on surfaces provide another class of problems modelled by large chemical mechanisms.[90, 35] The accurate performance of these models is necessary for important objectives such as engine design[89] and climate prediction.[55] The construction of accurate chemical mechanisms is typically an iterative endeavour where the model is repeatedly improved after comparison to experiment. This process requires insight into the workings of the chemical network. Unfortunately, the behaviour of complex kinetic mechanisms can become quite obscure making the process of validation and improvement difficult. Even the task of numerical simulation can become computationally prohibitive if the mechanism grows too large especially if spatial transport or multi-phase behaviour needs to be included.

We have recently developed a new approach to modeling the kinetics of chemical networks which we term the sum over histories representation, or SOHR for short.[53, 7, 4, 6] This method replaces the usual concentration based kinetics with a representation based on chemical pathways and can lead to new insight and increased computational capability when properly applied. A chemical

pathway is a sequence of elementary reactions where the product of one reaction becomes the reagent for the next reaction. Of course, the notion of the chemical pathway is quite familiar and is ubiquitous in all areas of chemistry. In kinetics, it has been often employed to provide mechanistic insight into complex networks of reactions.[41, 59, 24, 14] In chemical synthesis, for another example, a substrate molecule can be successively functionalized by a sequence of chemical steps. In atmospheric chemistry, a chlorine atom may be followed through a catalytic cycle that leads to the breakdown of ozone. We note that in these cases we are following a chemical moiety as it migrates from species to species due to chemical reactions. In the SOHR, the idea of the chemical pathway becomes the basis of a quantitative representation of the chemical kinetics. The pathways are explicitly defined by following a tagged atom as it migrates through species space. We find that if all the relevant chemical pathways and their associated probabilities are known, then it is possible to directly compute the value of any kinetic observable. The method draws its motivation from Feynman's 'sum over histories' approach to quantum mechanics.[25] As with Feynman path integrals, chemical pathway theory is often more computationally intensive than traditional approaches but it can provide new insight or numerical advantage for select problems.[68, 13, 11] The key mathematical relation in SOHR is an explicit expression for the probability of an arbitrary chemical path in terms of an integral of a time-ordered product. This high-dimensional integral can be evaluated using an efficient Monte Carlo (MC) integration scheme.

We have found it quite useful to visualise the chemical pathways using the methods of graph theory. On a chemical graph, the species are represented by nodes (or vertices) and the chemical reactions that interconvert the species are represented by edges (or lines).[15, 36, 8, 92, 65] As we follow a tagged atom through the chemical network, the pathway is represented by a sequence of vertices and edges. The graph then allows any pathway to be associated with a unique symbol sequence that labels the vertices and edges. Furthermore, all paths of a given length can be automatically generated by expansions in the symbol sequence. The chemical graphs we use here are weighted since the reaction rates can be used to assign a weight to each edge. This identification opens up the possibility of using graphical search algorithms to locate the most important paths

that contribute to a given process.

One aspect of the present method that distinguishes it from previous attempts to quantify chemical pathways[59, 47] is that SOHR fully embraces the dynamical nature of the chemical kinetics graphs. The time-dependence of edge weights (i.e. reactions rates) are accurately included even though this makes the computation of the probabilities much more challenging. Thus, as a molecule moves through the network its path instantaneously responds to the time varying rates. Other methods use static snapshots of the reaction flux or time-averages of the reaction rates to compute quantities such as the branching ratios in the chemical network. Since the rates of reaction can vary dramatically as a function of time, the passage through the network may be badly approximated by the static approximation.

The mathematical underpinning of the SOHR is discussed in Chapter 2. The computation of pathway probabilities and the methods for pathway enumeration are discussed in detail in this chapter. In Chapter 3 we present the results for a chemical network described by linear kinetics which serve to illustrate some of the general characteristics of the method. For a chemical system composed only of first-order reactions, the pathway probabilities can be computed analytically and is found to reduce to a sum of exponentials in time. It is shown that this linear kinetic problem can be solved with many fewer paths than might be expected. In Chapter 4, the H₂ combustion problem is analyzed using the SOHR method. It is shown that the ignition chemistry for this system can be accurately represented using a small number of chemical pathways. The use of the pathway representation also provides new insight into the mechanism and kinetic sensitivities of this problem. Chapter 5 demonstrates the applicability of SOHR in predicting chemical kinetics. Finally, Chapter 6 presents a short conclusion and some ideas for future research.

Chapter 2

Mathematical Development

Some of the contents of this chapter are reprinted, with permission, from

- [5] Bai, S., & Skodje, R. T. The Sum Over Histories Representation for Chemical Kinetics: A Quantitative Theory Based on Chemical Pathways. *International Reviews in Physical Chemistry*, 35(4), 539-567. **2016**

2.1 Introduction

The application of the SOHR approach requires a well-defined chemical model, and thus we begin with a review of conventional kinetics. The chemistry of a reaction network is traditionally modeled using mass-action kinetics where the concentrations for N -species are solved as functions of time, i.e. $(X_1(t), X_2(t), \dots, X_N(t)) = X(t)$. We shall assume that the chemical model under consideration consists of M elementary chemical reactions in a homogeneous well-stirred reactor

$$\sum_i^{\text{reactants}} \bar{v}_{i,j} S_i = \sum_i^{\text{products}} v_{i,j} S_i \quad j = 1, \dots, M \quad (2.1)$$

involving N distinct chemical species S_i which have concentration X_i . The molecularities, $v_{i,j}$ and $\bar{v}_{i,j}$, enforce the elemental mass conservation for each reaction indexed by j . The rate of each reaction is given by a well-defined rate law, $R_j(\mathbf{X})$. (Nothing prevents the SOHR method from being applied to problems with time-dependent rate coefficients although we do not explicitly include a t -dependence in the formulae.) The continuous concentrations of the distinct species S_i , $X_i(t) \geq 0$,

can be expressed as solutions of the differential rate eqn. 2.1. The rate equations are given by

$$\frac{dX_i(t)}{dt} = F_i(\mathbf{X}) \quad (2.2)$$

The net rate for species S_i , $F_i(\mathbf{X})$, is composed of the source and sink terms from each elementary reaction in the mechanism, i.e.

$$F_i(\mathbf{X}) = \sum_j^{\text{sources}} v_{i,j} R_j(\mathbf{X}) - \sum_j^{\text{sinks}} \bar{v}_{i,j} R_j(\mathbf{X}) \quad (2.3)$$

The solution to eqn. 2.2 can become difficult if the model grows very large or if multiple time-scales yields a high degree of stiffness in the equations.

We note briefly that another approach to kinetic modeling exists that is based on the methods of stochastic modeling. There, a discrete atomistic representation is adopted in which ensembles of molecules react according to a stochastic model.[96] This viewpoint, pioneered in the early work of McQuarrie[69] and later by Gillespie and coworkers[31, 32, 29], makes use of MC sampling methods to propagate large ensembles of discrete molecules. This method has proved useful in problems, such as those involving biomolecules, where concentrations are very low and hence the fluctuation in the number of molecules can play a role. A mathematical representation for this technique involves the chemical master equation (CME), where one solves for the probability distribution function $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$. The variables $\mathbf{x} = (x_1, x_2, \dots, x_N)$ represent the number of species molecules in a given volume and $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ is the probability that the species are populated by \mathbf{x} at time t given that they were populated as \mathbf{x}_0 at time t_0 . Obviously, the solution to the CME contains vastly more information than does the mean field eqn. 2.2. It is also much more difficult to solve the CME than the concentration based kinetics equations. The most effective numerical techniques for solving the CME involve averaging many stochastic trajectories which simulate the time evolution of a large ensemble of molecules.

2.2 Defining Chemical Pathways

The basic units of description in the SOHR are the chemical pathways themselves. The pathways represent the condition of a chemical moiety as it moves from species to species due to

the action of the elementary chemical reactions. Taken collectively, these pathways describe the entire chemistry of the model. In order to quantitatively apply the SOHR to practical systems, we require an unambiguous procedure to define and enumerate these pathways. One can imagine a reaction pathway as either a sequence of reactions or a sequence of species that propagates the chemical moiety through the chemical network. If the reaction sequence is R_1, R_2, \dots, R_n , then the product molecule from reaction R_i should be a reagent molecule for the reaction R_{i+1} . In terms of species, on the other hand, there will be a sequence of $n + 1$ moiety-containing molecules S_0, S_1, \dots, S_n that are connected by n elementary reactions. For a general non-linear mechanism, it is necessary to specify both the reactions and the species along the reaction pathway in order to achieve uniqueness. Indeed, the same reagents may yield different products or the same products may result from different reagents, depending on which reaction occurs. For example, if we imagine the gas-phase combustion of the methanol molecule, CH_3OH , a hypothetical three-step chemical pathway is $\text{CH}_3\text{OH}(+\text{OH}) \xrightarrow{R_1} \text{CH}_3\text{O}(+\text{O}) \xrightarrow{R_2} \text{CHO}(+\text{OH}) \xrightarrow{R_3} \text{CO}$ which yields the CO product. Obviously the reagents in first step of this path, $\text{CH}_3\text{OH} + \text{OH} \xrightarrow{R_1} \text{CH}_3\text{O} + \text{H}_2\text{O}$, could have equally well yielded the chemically distinct products $\text{CH}_2\text{OH} + \text{H}_2\text{O}$, which must be described by a separate reaction R'_1 . Likewise, the distinct reaction $\text{CH}_3\text{OH} + \text{H} \xrightarrow{R_1} \text{CH}_3\text{O} + \text{H}_2$ is a different reaction route that connects the followed species CH_3OH and CH_3O .

In most numerical applications of the SOHR methodology, we adopt a followed-atom approach for the definition of the chemical pathways. Thus, an atom in a reagent species is tagged at the beginning of the simulation, t_0 , and it is followed as it hops between species due to reaction until the final time t_f . For example, the carbon atom in the above pathway has a history $(t_0 | S_0 R_1 S_1 R_2 S_2 R_3 S_3 | t_f)$ where the C-containing species are $S_0 = \text{CH}_3\text{OH}$, $S_1 = \text{CH}_3\text{O}$, $S_2 = \text{CHO}$, and $S_3 = \text{CO}$, and the reactions are $R_1 = \text{CH}_3\text{OH} + \text{OH} \rightarrow \text{CH}_3\text{O} + \text{H}_2\text{O}$, etc. These histories represent pathways through species space which deliver the chemical moiety (in this case a C-atom) from the reagents to the products. A general pathway that delivers a C-atom from CH_3OH to CO may be arbitrarily long and thus may consist of an arbitrarily long sequence of reactions. The numerical utility of the SOHR method requires that only a finite and manageable number of the pathways are actually

required to converge meaningful results.

From our atomistic viewpoint, the reactions $R_1, R_2 \dots R_n$ are a sequence of random events that occur instantaneously at times $t_1, t_2 \dots, t_n$. Consistent with the time-ordering along the pathway, the times must satisfy the inequalities $t_f > t_n > t_{n-1} \dots > t_1 > t_0$. Hence, we can further specify the path by noting the reaction times along with the initial and final times, i.e. $(t_0 | S_0 R_1(t_1) S_1 R_2(t_2) S_2 R_3(t_3) S_3 | t_f)$. The chemical histories of an atom that goes from an initial to a final species may be, in principle, of any length n passing through $n + 1$ species via n reaction steps occurring at times $t_f > t_n > t_{n-1} \dots > t_1 > t_0$. In Fig. 2.1 we show a schematic of several chemical pathways that deliver a tagged atom from species S_1 to species S_7 . In the upper panel, we show these pathways on a chemical graph. In the lower panel, we show trajectories that follow these paths for a particular set of reaction times.

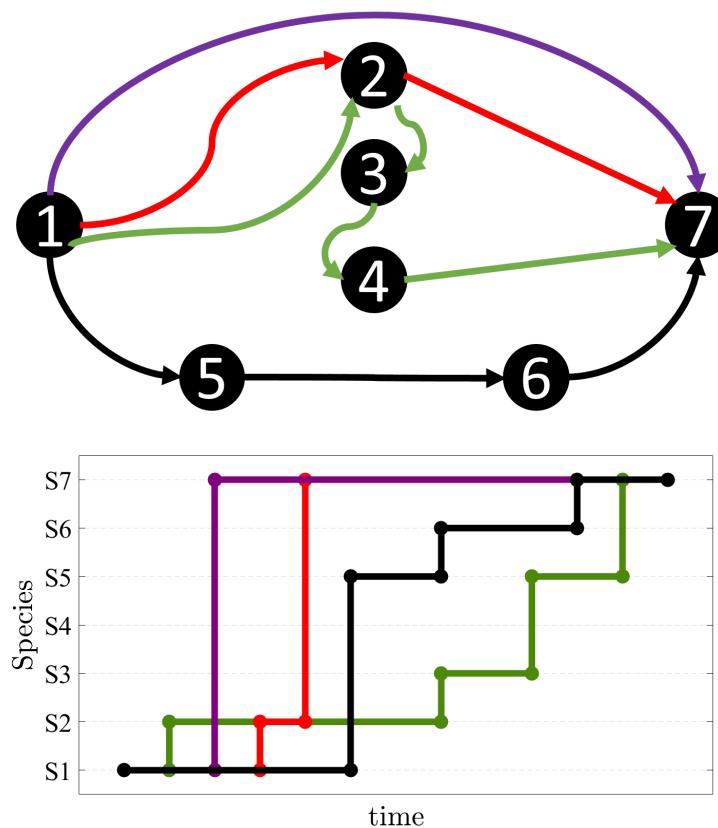
2.3 Pathway Probabilities

The essential step that enables the SOHR to be used as a quantitative model of the reaction kinetics is the assignment of a numerical probability for a given chemical pathway. We consider a general pathway consisting of n -reactions and $n + 1$ -species, i.e. $S_0 \xrightarrow[R_1]{t_1} S_1 \xrightarrow[R_2]{t_2} S_2 \dots S_{n-1} \xrightarrow[R_n]{t_n} S_n$ where the followed-atom hops from S_i to S_{i+1} at time t_{i+1} due to elementary reaction R_{i+1} . We define a *pathway probability density function*, *pPDF* for short,

$$\begin{aligned} \rho(t_0, t_1, \dots, t_n, t_f) \equiv & \text{the probability density, starting at time } t_0, S_0 \text{ undergoes} \\ & \text{reaction } R_1 \text{ to form } S_1 \text{ at time } t_1, \text{ then undergoes reaction} \\ & \text{reaction } R_2 \text{ at } t_2 \text{ to form } S_2, \text{ and so forth. The final species } S_n \text{ that} \\ & \text{is created at } t_n \text{ must survive until the final observation time } t_f. \end{aligned} \quad (2.4)$$

The probability density has units of $1/\text{time}^n$. Since the occurrence of reactions R_1, \dots, R_n , constitute a sequence of independent random events, i.e. a Markov chain, then $\rho(t_0, t_1, \dots, t_n, t_f)$ is a product of the appropriate decay probability densities for each step along the chain. The total pathway probability, $P_{\text{path}}(t_0, t_f)$, is obtained by integrating the probability density over all times

Figure 2.1: Two representations of chemical histories. In the upper panel, four chemical pathways followed by a tagged atom are displays on a chemical graph. The species are represented as vertices and reactions as lines (or edges). In the lower panel, four analogous stochastic trajectories are plotted as species number vs. time. The time-dependence of the trajectory corresponds to a Markov chain of random reactions.



(t_1, \dots, t_n) consistent with the time-ordering, i.e.

$$P_{path}(t_0, t_f) = \int_{t_0}^{t_f} dt_n \int_{t_0}^{t_n} dt_{n-1} \cdots \int_{t_0}^{t_2} dt_1 \rho(t_0, t_1, \dots, t_n, t_f) \quad (2.5)$$

As a very simple illustration, consider a one-step pathway $S_0 \rightarrow S_1$ governed by a mechanism consisting of a single first-order reaction with rate law $k[X_0]$. The quantity $\rho(t_0, t_1, t_f)$ is the probability density that a molecule of type S_0 , which was created at time t_0 , decays at time t_1 to form S_1 which then survives until time t_f . Clearly, in this case $\rho(t_0, t_1, t_f) = k \times e^{-k(t_1-t_0)}$ since the species S_1 is stable. Integrating over t_1 gives the result $P_{path} = 1 - e^{-k\Delta t}$, with $\Delta t = t_f - t_0$. Notice for this system there are only two possible paths that can be followed by a tagged atom, viz. the one-step path, $S_0 \rightarrow S_1$, and the trivial zero-step path where S_0 does not decay. For the latter path, the pathway probability is the non-decay probability $e^{-k\Delta t}$ and thus the sum of the pathway probabilities is 1. A slightly less trivial case would be a mechanism with a sequence of two first-order reactions, $S_0 \rightarrow S_1$ and $S_1 \rightarrow S_2$, with rate laws $k_1[X_0]$ and $k_2[X_1]$, respectively. Again S_2 is a stable species. Notice in this case there are three possible paths emanating from S_0 : a two-step path $S_0 \rightarrow S_1 \rightarrow S_2$, a one-step path $S_0 \rightarrow S_1$, and a zero-step path S_0 . The probability density for the two-step path $S_0 \rightarrow S_1 \rightarrow S_2$ is

$$\rho(t_0, t_1, t_2, t_f) = k_1 e^{-k_1(t_1-t_0)} k_2 e^{-k_2(t_2-t_1)} \quad (2.6)$$

and the total integrated pathway probability is

$$\begin{aligned} P_{path}(2-step) &= \int_{t_0}^{t_f} \int_{t_0}^{t_2} dt_1 \rho(t_0, t_1, t_2, t_f) \\ &= 1 + \frac{k_1 e^{-k_2 \Delta t}}{k_2 - k_1} - \frac{k_2 e^{-k_1 \Delta t}}{k_2 - k_1} \end{aligned} \quad (2.7)$$

where $\Delta t = t_f - t_0$. The probability density for the one-step path, $S_0 \rightarrow S_1$, is

$$\rho(t_0, t_1, t_f) = k_1 e^{-k_1(t_1-t_0)} e^{-k_2(t_f-t_1)} \quad (2.8)$$

Note that in eqn. 2.8, we have inserted the non-decay probability of species S_1 , i.e. $e^{-k_2(t_f-t_1)}$, which is created at time t_1 and must survive until t_f . The corresponding pathway probability is

$$\begin{aligned} P_{path}(1 - step) &= \int_{t_0}^{t_2} dt_1 \rho(t_0, t_1, t_f) \\ &= -\frac{k_1 e^{-k_2 \Delta t}}{k_2 - k_1} + \frac{k_1 e^{-k_1 \Delta t}}{k_2 - k_1} \end{aligned} \quad (2.9)$$

The zero-step pathway probability is the non-decay probability that S_0 is created at time t_0 and survives until time t_f , i.e.

$$P_{path}(0 - step) = e^{-k_1 \Delta t} \quad (2.10)$$

It is easily seen that the sum of eqns. 2.7, 2.9, and 2.10 is one.

To generalize the probability density for a time-resolved pathway to an arbitrary kinetic mechanism, we find it convenient to define two quantities: the species survival probability and the reaction branching ratio. The species survival probability, $\mathcal{P}_i(t_a, t_b)$, is the probability that a molecule of type S_i , which is created at t_a , undergoes no reaction in the time interval (t_a, t_b) where $\mathcal{P}_i(t, t) = 1$. Unlike the simple first-order reactions above, where $P(t_a, t_b) = \exp(-k(t_b - t_a))$, the survival probability is temporally inhomogeneous for general mechanisms, i.e. $\mathcal{P}_i(t_a, t_b) \neq \mathcal{P}_i(t_a + \Delta, t_b + \Delta)$. However it is straightforward to compute the survival probability for any species S_i by integrating the species decay probability obtained from the sum of sink reactions defined in the rate eqn. 2.2

$$\begin{aligned} \mathcal{P}_i(t_a, t_b) &= \exp \left(- \int_{t_a}^{t_b} \sum_j^{\text{sink}} A_{i,j}(t) dt \right) \\ &\equiv \exp \left(- \int_{t_a}^{t_b} A_i(t) dt \right) \end{aligned} \quad (2.11)$$

The quantities $A_{i,j}(t)$ are similar to pseudo-first-order rate coefficients (with dimensions of 1/time) and are defined by the time-dependent expressions

$$A_{i,j}(t) \equiv \frac{\bar{v}_{i,j} R_j(\mathbf{X})}{X_i(t)} \quad (2.12)$$

where i labels the species and j labels the reactions. The concentrations $\mathbf{X}(t)$ are evaluated along the solution to eqn. 2.2. The instantaneous decay rate for species S_i at time t_b , assuming that S_i was created at t_a , is $-d\mathcal{P}_i(t_a, t_b)/dt_b$. A second useful quantity is the time-dependent branching

ratio, $\Gamma_i(t)$, which measures the instantaneous fraction of the reactive flux that follows the correct branch when species S_i reacts to form products. The branching ratio is simply the reactive flux along the selected ' J^{th} -branch' divided by the total reactive flux of species S_i , i.e.

$$\Gamma_i(t) \equiv \frac{A_{i,J}(t)}{A_i(t)} \quad (2.13)$$

In this notation we drop the explicit J-index on Γ . The stoichiometric factors are not included in the numerator of eqn. 2.13 since we are following a single tagged atom. Using these expressions in eqn. 2.11 and 2.13, $pPDF$ defined in eqn. 2.4 can be written as

$$\rho(t_0, t_1, \dots, t_n, t_f) = (-1)^n \prod_{i=1}^n \left(\frac{d\mathcal{P}_{i-1}(t_{i-1}, t_i)}{dt_i} \Gamma_{i-1}(t_i) \right) \mathcal{P}_n(t_n, t_f) \quad (2.14)$$

In eqn. 2.14, $(-1) \frac{d\mathcal{P}_{i-1}(t_{i-1}, t_i)}{dt_i}$ term represents, given species S_{i-1} formed at time t_{i-1} , the *probability density* that S_{i-1} will react at time t_i ; $\Gamma_{i-1}(t_i)$ is attributed to the instantaneous branching ratio that S_{i-1} reacts by R_J and forms S_i ; $\mathcal{P}_n(t_n, t_f)$ guarantees terminating species S_n survives from time t_n till t_f . Conclusively, plug eqn. 2.14 into eqn. 2.5, it shows that the pathway probability can be written as an integral of a time-ordered product.[7, 4]

$$P_{path}(t_0, t_f) = (-1)^n \int_{t_0}^{t_f} dt_n \int_{t_0}^{t_n} dt_{n-1} \cdots \int_{t_0}^{t_2} dt_1 \prod_{i=1}^n \left(\frac{d\mathcal{P}_{i-1}(t_{i-1}, t_i)}{dt_i} \Gamma_{i-1}(t_i) \right) \mathcal{P}_n(t_n, t_f) \quad (2.15)$$

This expression is similar to Dyson's formula for the amplitude in quantum mechanics.[83, 73] Unlike the quantum expression, however, the integrand of eqn. 2.15 is positive definite and thus phase cancellation is not an issue.

The evaluation of the path probability using a numerical quadrature for eqn. 2.15 can become laborious when the dimension, n , is large. However, this multidimensional integral is well suited for MC evaluation since the integrand is positive. In the numerical implementation of the MC approach, we find that the wide range of time scales present in realistic kinetic problems leads to stiffness and inefficient sampling of the time intervals. This technical difficulty can be overcome by using an importance sampling that employs the integration variables P_i , rather than t_i . In this method, a

random number y_0 is first chosen for species S_0 in the interval $(P_{min}^0(t_0), 1)$; $P_{min}^0(t_0) = \mathcal{P}_0(t_0, t_f)$ is the minimum value of the survival probability of S_0 which occurs at $t = t_f$. Then, the expression $y_0 = \mathcal{P}_0(t_0, t_1)$ is inverted to obtain t_1 , the time of the reaction R_1 . This time is used in eqn. 2.7 for Γ_1 . Then, for species S_1 a second random number, y_1 , is chosen in the interval $(P_{min}^1(t_1), 1)$. As before, the reaction time t_2 is obtained and branching ratio Γ_2 is computed. This process is continued for all the species along the pathway, which thus requires a string of n random numbers. We have for the final MC expression

$$P_{path}(t_0, t_f) = \frac{1}{N} \sum_{q=1}^N \mathcal{P}_n(t_n^q, t_f) \prod_{k=1}^n \left(\Gamma_{k-1}(t_k^q) \left(1 - P_{min}^{k-1}(t_k^q) \right) \right) \quad (2.16)$$

where the index q is used to specify each string of n -random numbers where, in this equation only, N is the size of the MC sampling. It is important to note that, except for linear or steady state systems, the SOHR apparently requires input from the traditional kinetics representation since the pseudo-first-order rate coefficients $A_{i,j}(t)$ are computed using the kinetic solutions $\mathbf{X}(t)$. The reference trajectory $\mathbf{X}(t)$ is an input when treating non-linear kinetic systems. Developing a method to detach the SOHR from traditional kinetics, and from need for a reference trajectory by using an iterative solution technique will be discussed in chapter 5.[6]

2.4 Constructing Observables

The SOHR can be used to compute the values of observables if a sufficient number of chemical pathways are employed. Consider the fundamental task in kinetics of computing the concentration of a species S_i at time t_f , $[X_i(t_f)]$, given the values of the initial reagent concentrations $[X(t_0)]$. Taking the simplest case first, assume that there is no more than one followed-atom, lets say a carbon atom, that exists in any species. Furthermore, assume that the target species S_i contains one carbon atom and that only one of the reagent species contains any carbon, say S_r . This would be the case, e.g. if we were following a carbon atom from CH_4/O_2 reagents to the CO_2 product in a pure C_1 mechanism. Using the pathway index j , we enumerate all important paths that deliver the carbon atom from the reagent S_r at t_0 to the target species S_i at time t_f . The pathway probabilities

for each path j , computed using eqn. 2.15 or 2.16, are abbreviated by P_j . Then, it is clear that the concentration of product is given by

$$[X_i(t_f)] = \sum_j [X_r(t_0)] P_j(t_f) \quad (2.17)$$

Of course this represents only a special case since in general there can be more than one followed-atom in each species and there may be more than one initial reagent species containing followed-atoms. We use the notation μ_i to represent the number of chemically equivalent followed-atoms in species S_i . We generally have that the followed-atom for the target species S_i lies in a multiplet of μ_i chemically equivalent atoms. Likewise, the followed-atom may potentially originate from one of several initial species, S_k , each of which may contain a multiplet of μ_k chemically equivalent followed-atoms. The pathways are labeled by ' j_k ' where k specifies the species of origin for the path of the followed-atom and the associated probabilities are denoted by P_{j_k} . Notice that the pathway probabilities are defined using species to species reaction probabilities and thus the numbers μ_i are only required at the beginning and the end of the path. The concentration of S_i is thus seen to be

$$\mu_i [X_i(t_f)] = \sum_{j_k} \mu_k [X_k(t_0)] P_{j_k}(t_f) \quad (2.18)$$

Most other kinetic observables can be expressed in terms of the species concentrations, and hence the pathway representation can be employed to analyse most kinetic problems.

2.5 Graph Theoretic Enumeration of Chemical Pathways

While eqn. 2.18 provides a means to represent kinetic observables given a sufficient number of paths, we must still devise a method to generate those paths. For simple mechanisms that contain a small number of species and reactions, it is possible to construct the chemical pathways by inspection or with chemical insight. When the mechanism is complicated, however, it becomes necessary to employ a more systematic algorithm. A useful approach involves the use of graph theoretic methods.[8] We construct the chemical graph using two sorts of elements: vertices and edges. The vertices, represented by dots on the graph, are species which contain the followed-atom.

The edges (or lines) on the graph connect the vertices and correspond to elementary reactions that deliver the followed-atom from one species (vertex) to another. The kinetic graphs employed here have certain specific characteristics which prove useful. First, the graphs are directed since the elementary reactions have a sense of direction. Second, the graphs are weighted since the edges are quantified by the pseudo-first-order rate coefficients, $A_{i,j}$. Third, the kinetic graphs are multigraphs since there may be more than one edge connecting two given vertices. Fourth, the graphs are dynamical since the weights are generally time-dependent. A unique chemical pathway of length n is represented by a specific directed sequence of $n + 1$ -vertices and n -edges on the graph.

Elementary graph theory can be used to enumerate those chemical pathways that are sufficiently short. Consider first the case of simple graphs where at most one edge connects any two vertices and no vertex is connected by an edge to itself. This might describe, e.g. the case of a totally linear kinetic network. We define the adjacency matrix, \mathbf{M} , as an $N \times N$ matrix (where N is the order of graph, i.e. the number of vertices) which has the matrix elements given by

$$M_{i,j} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are connected by an edge} \\ 0, & \text{if } i \text{ and } j \text{ are not connected by an edge} \end{cases} \quad (2.19)$$

For a simple graph, it is not hard to show that the total number of paths of length n between species i and j , $K(n; i, j)$, is given by

$$K(n; i, j) = (\mathbf{M}^n)_{i,j} \quad (2.20)$$

Furthermore, we can explicitly generate these $K(n; i, j)$ paths using matrix multiplication. Define the closely related S-matrix to be $S_{i,j} = M_{i,j}s_j$, where the dummy variable s_j is the species label (i.e. a symbol). Then the sum of paths of length n may be obtained by expanding out the (i, j) th matrix element in the product

$$s_i \times (\mathbf{S}^n)_{i,j} \quad (2.21)$$

These $K(n; i, j)$ symbol sequences of species labels uniquely define paths on a simple graph since only one possible edge can connect any two vertices.

The situation is more complicated for a general kinetic multigraph where more than one reaction (i.e. edge) can deliver a tagged atom between two species. In this case we need to specify which edge is being followed by the tagged atom. For a multigraph, define the generalized adjacency matrix, \mathbf{M} , to be

$$M_{i,j} = \begin{cases} m_{i,j}, & \text{if } i \text{ and } j \text{ are connected by } m_{i,j} \text{ edge} \\ 0, & \text{if } i \text{ and } j \text{ are not connected by an edge} \end{cases} \quad (2.22)$$

Then, again, the number of paths of length n connecting species i and j is given by the matrix element of \mathbf{M}^n specified in eqn. 2.20. To explicitly generate these paths, we generalize the \mathbf{S} -matrix above to the \mathbf{SR} -matrix that represents both the species and reactions involved in the path. We define

$$(SR)_{i,j} = \begin{cases} \sum r(i,j) \times s_j, & \text{if } M_{i,j} \neq 0 \\ 0, & \text{if } M_{i,j} = 0 \end{cases} \quad (2.23)$$

where $r(i,j)$ are the $m_{i,j}$ dummy variables (i.e. symbols) labeling the elementary reactions connecting species i and j , directed as $i \rightarrow j$, and s_j are again species symbols. The sum of symbol sequences for paths of length n are obtained expanding the matrix product

$$s_i \times (SR^n)_{i,j} \quad (2.24)$$

The most straightforward approach to the pathway enumeration problem in the SOHR is to use the expansion in eqn. 2.24 to compute all the paths of length up to $n = n_{max}$, and then converge the observables by progressively increasing this maximum length.

Obviously the number of potential paths can grow tremendously large as the path length n , or graph order N , increases. Even with the efficient MC-algorithm for computing the pathway weights, eqn. 2.16, the SOHR method can bog down if the number of paths required becomes too large. We point out, however, that it is common to find that many of the paths enumerated in eqn. 2.24 have very low probabilities and make a negligible contribution in the expansion of the

observables. Hence, it is often advisable to use a method that selects the most important pathways rather than simply employing a full expansion of all paths up to a given length. This brings us to the topic of path search algorithms.

It is a classic problem in computer science to locate optimal paths on a weighted graph.[60] For example, there have been a number of algorithms developed to systematically find the shortest path between two vertices where the edge weight, w_i , is identified as a length. In that problem, the path length, i.e. the sum of lengths of individual edges $L = \sum w_i$, is minimised by varying the path itself. The Dijkstra algorithm[21] is a commonly used procedure to solve this problem. The Bellman-Ford[10] and A*-search algorithms[39] are also well-known methods and provide somewhat more computationally efficient approaches. The present kinetics system is different from the traditional shortest path problem in several ways. First, we generally need to find many important paths rather than a single optimal path. This is not a major problem since several of the traditional optimal path methods have been generalised to obtain the 'K-shortest paths' on a weighted graph[22]. Second, the chemical paths we seek will maximise a probability rather than minimising a length. The total probability of a chemical pathway is related to the product of weights for each edge rather than the sum of edge weights used for computing path length. This distinction is easily overcome by taking the logarithm of the pathway weights which converts the product into a sum. Third, and most problematic, as a dynamical graph[38] the weights of the kinetic graph are time-dependent and thus the optimization process is itself time-dependent. This is because the weights of the edges are computed from the time-dependent pseudo-first-order rate coefficients $A_{i,j}(t)$. One simple shortcut around this problem would be to hold fixed the values of $A_{i,j}(t)$ at some particular time value and carry out the optimization for the resulting static graph. Alternatively, one might carry out the search using weights averaged over some time interval. The time-independent weights might also be updated during the run according to some criterion. While the static approximations to the dynamical graph problem may prove useful for some problems, it is easy to imagine scenarios where the true high probability pathways are not correctly identified. Therefore, we have adopted an approach based on stochastic sampling which is discussed in the next section.

2.6 Stochastic Pathway Enumeration

We have devised a simple method to enumerate the important chemical pathways on a dynamical graph using a stochastic sampling method closely related to the simulation algorithm developed by Gillespie and coworkers[32] and others[94]. Imagine that we want to identify important paths that carry a tagged atom from an initial species, S_0 , to a final species S_f . We can locate such pathways by propagating individual molecules through the reaction network using a kinetic MC simulation. An individual molecule of type S_i carrying the tagged atom has an instantaneous transition probability (not *PDF*) for going to species S_k by reaction R_j given by

$$TP(S_i \rightarrow S_k; R_j) = A_{i,j}(t) \times \gamma_k^j \quad (2.25)$$

Where γ_k^j indicates the branching ratio the 'tagged' atom flows into species S_k given reaction R_j (γ_k^j is typically a constant tensor). Using importance sampling based on the species survival probability, as above, the tagged atom can be propagated efficiently through the network via a random walk. For one step of that walk, a single random number determines the reaction time while a second random number selects the product branching. Those paths that go from S_0 to S_f are recorded. Using a modest number of such MC trajectories, a list of the important paths along with their approximate ranking is obtained. While this method is not appropriate for locating rare paths, we have found it very useful for many of the kinetic systems we have studied. Since this method involves probabilistic sampling, it is naturally possible that paths can be missed if too few random walks are employed.

Chapter 3

Linear Kinetics

Some of the contents of this chapter are reprinted, with permission, from

- [5] Bai, S., & Skodje, R. T. The Sum Over Histories Representation for Chemical Kinetics: A Quantitative Theory Based on Chemical Pathways. *International Reviews in Physical Chemistry*, 35(4), 539-567. **2016**

3.1 Analytical Expression of Linear Kinetics

If the chemistry of a system is accurately described by a purely linear kinetic model, then the implementation of the SOHR becomes especially simple. We use this transparent case to illustrate some of the characteristics of the method. Consider a reaction network consisting of N-species, S_i ($i = 1, \dots, N$), which can undergo the first-order reactions $S_i \rightarrow S_j$ described by the rate laws $k_{i,j}[X_i]$. The graph for this linear system is 'simple' since at most one edge (reaction) will connect any two vertices (species) and there are no self reactions, $S_i \not\rightarrow S_i$. Thus, any n-step path can be uniquely represented by the corresponding product of species, $S_0S_1 \cdots S_n$ and it is not necessary to specify the reactions. The quantities $A_{i,j}(t)$ defined in eqn. 2.12 are constants in time, i.e. $A_{i,j}(t) = k_{i,j}$. The species decay rates, $A_i = \sum A_{i,j} = \sum k_{i,j} \equiv k_i$, are likewise time-independent and all the species survival probabilities correspond to pure exponential decay, i.e. $\mathcal{P}_i(t_a, t_b) = \exp(-k_i(t_b - t_a))$. The reaction branching ratios are given by $\Gamma_i = k_{i,J}/k_i$ where J denotes the appropriate channel along the pathway selected. Thus, for linear kinetics the reference trajectory $\mathbf{X}(t)$ from eqn. 2.2 is not required to implement the SOHR. Given these simplifications,

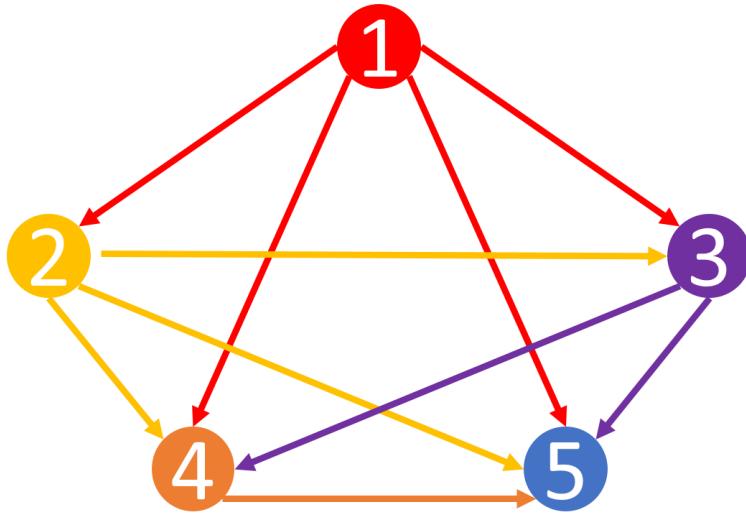
the integral of the time-ordered product, eqn. 2.3, can then be evaluated analytically[7]

$$\left\{ \begin{array}{l} P_{path}(t_0, t_f) = \sum_{i=0}^n D_i^n \exp(-k_i(t_f - t_0)), \\ D_i^n = \frac{\prod_{j=0}^{n-1} k_j \Gamma_j}{\prod_{j=0, j \neq i}^n (k_i - k_j)}, \end{array} \right. \quad \begin{array}{l} (a) \\ (b) \end{array} \quad (3.1)$$

The pathway probability is seen to be related to the integrated rate law for n-sequential first-order reactions.[1] We note that the apparent singularity of this expression, which occurs when two decay rates are equal, $k_i = k_j$, can be removed by making a small displacement of one of the k_i 's.

To illustrate the workings of the SOHR for linear kinetics, we consider a simple 20 species system,

Figure 3.1: The chemical graph corresponding to an acyclic linear kinetic model with 5 species. The edges connect only the species $S_i \rightarrow S_j$ with $j > i$.



i.e. $N = 20$, where the graph is acyclic. In an acyclic graph, no path can form a closed loop and thus all the paths are of finite length. This condition is enforced here by the condition $k_{i,j} = 0$ when $i \geq j$. In Fig. 3.1 we show the chemical graph restricted to the first 5 species. We allow the lower triangle of the rate constant matrix $k_{i,j}$ to be fully occupied. The non-zero rate constants are selected with a random number generator over an order of magnitude range, i.e.

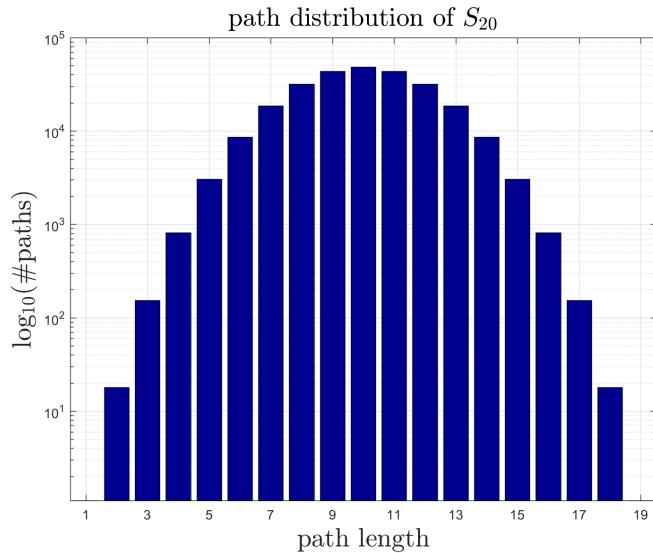
$$k_{i,j} = 1 + (10 - 1) \times \text{random}(0, 1)$$

We find most random sets of rate constants give results that are qualitatively similar.

3.2 Number of Pathway vs. Pathway Length

Using the adjacency matrix, eqn. 2.19, we find that the number of potential pathways can grow quite large even for this simple problem. For example, if we follow species S_1 through the network to species S_{20} there are 262,144 allowed paths that a molecule might take. In Fig. 3.2, we show the distribution of paths vs. path length, L . There is one path of length $L = 1$, viz. $S_1 \rightarrow S_{20}$, and one path of length $L = 19$, viz. $S_1 \rightarrow S_2 \cdots S_{19} \rightarrow S_{20}$. The largest number of paths occur in the middle of the range, $L = 10$. Of course the number of paths will depend on the choice of initial and final species.

Figure 3.2: The number of independent chemical pathways connecting species S_1 with S_{20} in a linear kinetic model with rate coefficients $k_{i,j} \neq 0$ for $j > i$ and $k_{i,j} = 0$ for $i \geq j$. The length L is the number of reactions along the path.



3.3 Demonstration of Convergence of SOHR

In order to assess the numerical performance of the SOHR method, we have computed for comparison the 'exact' solution to the linear system using a 4th-order Runge-Kutta integration of the differential kinetics equations

$$\frac{d[X_i]}{dt} = Sources - Sinks = \sum_{j=1}^{i-1} k_{i,j}[X_j] - k_i[X_i] \quad (3.2)$$

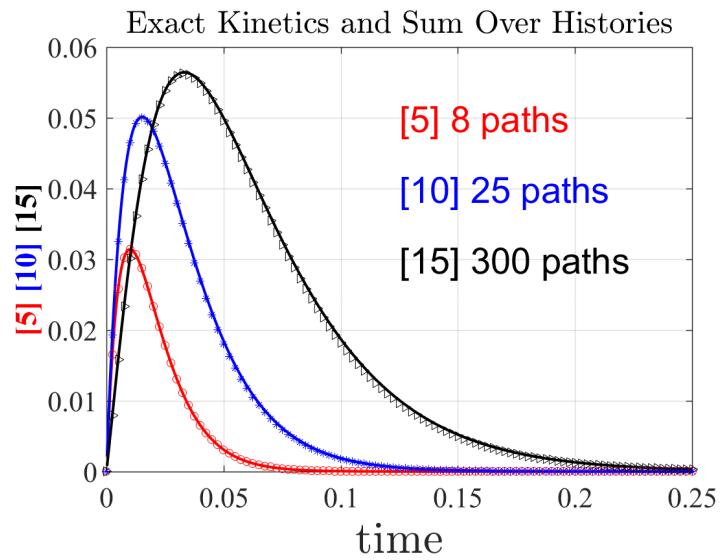
We note that species S_{20} is stable since $k_{20} = 0$. Thus, for the initial condition $[X_i(t = 0)] = \delta_{i,1}$ we know that the solution will tend to $[X_{20}(t)] \rightarrow 1$ as $t \rightarrow \infty$. We shall compare the Runge-Kutta solution of the kinetics originating from the S_1 initial condition to the prediction of the SOHR using the pathways originating with S_1 . We note that if all the chemical paths are included in the SOHR expansion we also obtain the exact result.

Using the stochastic method described in Section 2.6, we have identified a number of candidate pathways that converts species S_1 to each of the other species, S_i . Thus, we have run several thousand MC trajectories and produced a rough ranking of the pathways that connect S_1 to the other species. In Fig. 3.3, we show the concentrations of several intermediate species computed using eqn. 3.1 and the pathway expansion $[X_i(t_f)] = \sum_j P_j [X_1]$. To achieve better than 1% accuracy, we used 8 paths for S_5 , 25 paths for S_{10} , and 300 paths for S_{15} . Clearly, the pathway representation reaches practical convergence for the concentration long before the all of the potential paths are included. Indeed, the rapid convergence of the expansion with respect to the number of paths is a criterion for the practical utility of the SOHR as a computational method.

3.4 Pathway Probabilities Changing Over Time

It is instructive to investigate the behaviour of the SOHR in somewhat greater detail for this simple problem. We note that even though the pseudo-first-order rate coefficients are time-independent, the pathway probabilities still possess a strong time-dependence. In Fig. 3.4, we show the pathway probabilities vs. time for the 100 most important paths connecting S_1 to S_{15} . The paths are color-coded according to their length. At very early times, the one-step path,

Figure 3.3: The concentration of several chemical intermediates vs. time for the linear kinetic model consisting of 20 species. The solid lines are the result of converged numerical integration of the conventional kinetics equations, eqn. 3.2, for species S_5 , S_{10} , and S_{15} . The symbols are the result of the SOHR using eqn. 3.1 for the pathway probabilities. The pathways were identified using a small stochastic simulation.



i.e. $S_1 \rightarrow S_{15}$, dominates the probability. At somewhat later times, the two-step paths, i.e. $S_1 \rightarrow S_j \rightarrow S_{15}$ begin to play a role. At still later times, the three- and four-step paths begin to contribute. Heuristically, we can understand that a molecule which arrives at species S_{15} at a later time will tend to take a more circuitous route than one that arrives at an earlier time. In Fig. 3.5, we show the growth of the mean path length as a function of time. Obviously the concentration of a kinetic intermediate is closely connected to the notion of arrival time distributions or first passage time distributions that are often studied in stochastic models.[96] If we consider the full passage through the network to the attracting final state, S_{20} , the pathway probabilities take a simple form. As $t \rightarrow \infty$, the sum of exponentials in eqn. 3.1 contains only the last term with $k_{20} = 0$. In this case, the pathway probability becomes the product of the branching ratios along the path and is time-independent. It is interesting to point out that even for a general non-linear mechanism, if the kinetics reaches a steady state condition then the same linear analysis may be invoked to implement the SOHR.

Figure 3.4: Pathway probabilities vs. time for the 100 most probable paths connecting S_1 with S_{15} in the linear kinetic model. At each time, the sum of probabilities is normalised to unity. The pathway probabilities are colour coded according to their path lengths.

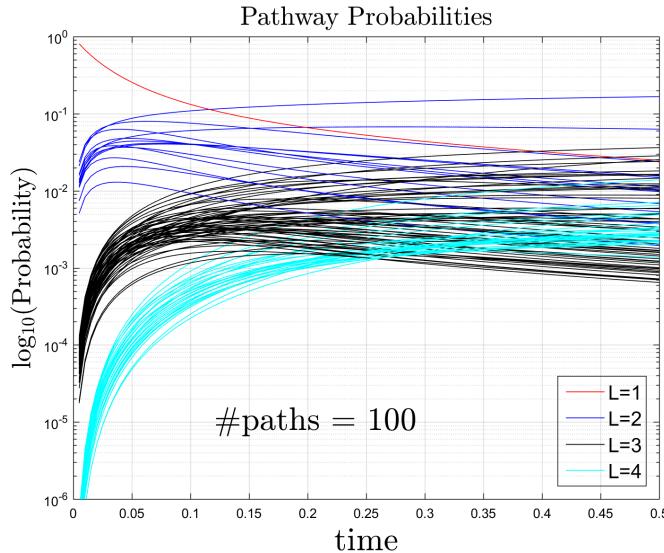
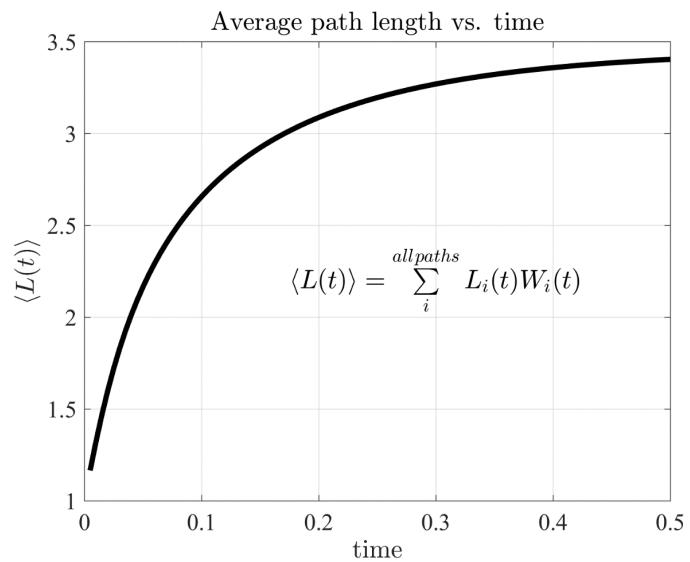


Figure 3.5: The average path length vs. time for chemical paths going from S1 to S15 in the linear kinetic model. The weight of each chemical path is computed using Equation (3.1) where the weights are normalised to unity at each time.



Chapter 4

Kinetic Sensitivity Analysis Using Sum Over Histories Representation

Most of the contents of this chapter are reprinted, with permission, from

- [7] Bai, S., Zhou, D., Davis, M. J., & Skodje, R. T. Sum over histories representation for chemical kinetics. *The journal of physical chemistry letters*, 6(1), 183-188. **2014**
- [4] Bai, S., Davis, M. J., & Skodje, R. T. Sum over Histories Representation for Kinetic Sensitivity Analysis: How Chemical Pathways Change When Reaction Rate Coefficients Are Varied. *The Journal of Physical Chemistry A*, 119(45), 11039-11052. **2015**
- [5] Bai, S., & Skodje, R. T. The Sum Over Histories Representation for Chemical Kinetics: A Quantitative Theory Based on Chemical Pathways. *International Reviews in Physical Chemistry*, 35(4), 539-567. **2016**

4.1 Introduction

Sensitivity analysis provides an important tool for the study of chemical kinetic networks.[54, 79, 95] In chemical kinetic sensitivity analysis the response of a target observable, such as a species concentration at time t , $[\mathbf{X}_i(t)]$, is probed as a function of variable factors, which are commonly chosen to be the rate coefficients, k_j . The use of sensitivity measures such as local or global sensitivity indices allows the chemistry described by complex mechanisms to be better understood physically, and in some cases permits the mechanisms to be systematically improved. For example, a large sensitivity index may indicate a key reaction step whose rate coefficient may then be im-

proved upon using better experimental or theoretical methods. Conversely, reactions with very low sensitivity indices may be targeted for removal from the mechanism which will enhance the computational efficiency. Furthermore, if the rate coefficients can be assigned uncertainty ranges, the overall uncertainty of the model predictions can be quantified using the function representation of the target.[74] Although the utility of sensitivity analysis in chemical kinetics is widely appreciated, it is not always possible to identify a specific chemical explanation that rationalizes the observed sensitivities. Thus, it is not uncommon for reaction to be found highly sensitive although it is unclear by what chemical mechanism this sensitivity arises. When the origin of observed kinetic sensitivity is not well understood, the practical utility of the results for model improvement or simplification is a less well motivated endeavor. In this work, we propose the use of our recently developed sum over histories representation as the basis to interpret kinetic sensitivity.[53, 7]

The sum over histories approach to chemical kinetics is based on the idea of expressing the quantitative kinetics of a chemical network using separate numerical contributions (via probabilities) from distinct chemical pathways. Our method to simulate the time dependence of a chemical network draws its motivation from Feynman's "sum over histories" approach to quantum mechanics[25] and is closely related to treatments of time-ordered Markov chains[96] discussed by Wiener[28] and Ito.[44] In our method, the paths lie in chemical space unlike Feynman's dynamical paths. The notion of a chemical pathway plays a central role in our understanding of chemical systems and is a quite familiar concept. In synthetic chemistry, for example, one follows a substrate molecule as it undergoes a sequence of alterations en route to a desired product. In general, a pathway is defined as a sequence of species, S_0, S_1, \dots, S_N connected by elementary reaction R_1, R_2, \dots, R_L such that the product of one reaction is a reagent for the next reaction along the pathway. Reaction pathways and graphical methods are also well-known in kinetic studies where they are used effectively to interpret and optimize chemical mechanisms.[93, 14, 65, 59, 41, 24, 47] The sum over histories approach is unique, however, in that it provides a fully quantitative and convergent procedure in which the concentration of any target species at time t , $[\mathbf{X}(t)]$ may be computed by

a linear combination of chemical pathways that lead from initial reactants to the target

$$[\mathbf{X}(t)] = \sum_i c_i p_i(t) \quad (4.1)$$

where $p_i(t)$ is the probability for path i and c_i are trivial time-dependent coefficients obtained from the stoichiometry. The pathway probabilities are given explicitly by an integral of a time-ordered product that is analogous to the Feynman path integral or the Wiener integral. The advantage of the sum over histories formulation is that the kinetics, and hence the behavior of sensitivity targets, is now represented by complete mechanisms that incorporate complete chemical pathways rather than by individual reaction steps that may participate in many distinct mechanisms. The origin of kinetic sensitivity may then be directly studied in terms of how a kinetic parameter affects the probability for a particular mechanism, i.e., through the pathway probabilities $p_i(t)$.

In this work we investigate sensitivity in the kinetics for ignition in the H₂-O₂ system at intermediate temperatures, i.e., T = 1000 K. We select for sensitivity targets the species concentrations at times before ignition and, thus, we focus on the chemistry underlying the growth of the radical pool, which determines the ignition induction period. It is well-known that speciation studies provide a particularly effective approach for the improvement and validation of chemical mechanisms.[46] The chemistry of the H₂ combustion problem is now fairly well understood and the rates of most of the elementary steps are known to good accuracy.[52, 40] Our purpose here, however, is not to improve the combustion model, but rather to investigate a new method to carry out sensitivity analysis. The H₂ combustion problem is ideal for such a test because it exhibits the generic characteristics of combustion kinetics whereas its mechanism is sufficiently simple to permit a full numerical investigation. We find that although the sum over histories approach is not a panacea, it provides useful insight into the origin of the kinetic sensitivity of speciation targets and it appears to be a useful tool to apply to more complicated problems.

This chapter is organized as follows. In section 4.2, the sum over histories approach to kinetics is reviewed and the use of the high dimension model representation (HDMR) for global sensitivity analysis (GSA) is summarized. In section 4.3, the details of the combustion simulation are dis-

cussed and the results of the GSA for speciation in the H₂ combustion problem are presented. In section 4.4, the chemical pathways that contribute to the various species targets are found and the associated probabilities are computed. It is shown that the species concentrations can be converged by summing a small number of chemical pathways. In section 4.5, the sensitivity of the various target species is discussed using the chemical pathways uncovered in section 4.4. Finally, section 4.6 is a short conclusion.

4.2 Methodology

We begin with a brief review of the concepts of sensitivity analysis and of the sum over histories representation that will also serve to define the required mathematical notation. We shall explicitly consider the case of spatially homogeneous kinetics described by the continuous concentrations (number of molecules per unit volume) of the N distinct species S_i , $X_i(t) \geq 0$. The total reaction mechanism is described by L elementary reaction steps, of the form (2.1), where the rate laws $R_j(\mathbf{X})$ are parametrized by the rate coefficients k_i , $i = 1, \dots, L$. In a typical application of sensitivity analysis, one imagines that these rate coefficients are variable over a range Δk_i . This range of variation may be due to an actual uncertainty in our knowledge of k_i or may simply be introduced as a formal tool in the analysis of the chemical network.

4.2.1 Sum over Histories Representation

In the sum over histories representation introduced in Chapter 2, the solution to the kinetics problem is expressed in terms of a sum over chemical pathways that relate an initial condition at $t = 0$ to some final target observable at time t. A pathway is specified through a sequence of chemical species S_0, S_1, \dots, S_n where the product of one reaction is the reagent for the next by means of an elementary reaction, R_i , i.e., $S_{i-1} \xrightarrow{R_i} S_i$. We have adopted a followed-atom approach in which a given atom (or, more generally, a stable chemical moiety, or even a "hypothesized" atom) is used to specify the paths. Thus, we take a single molecule perspective in which the chemical moiety hops discretely from one species to the next due to the action of the elementary reactions.

The sum over histories method consists of the following practical steps:

- (1) Find the most important chemical pathways for the problem at hand.
- (2) Find the probabilities for each path.
- (3) Combine the probabilities for the pathways to form the observable.

The method is quantitative and the chemical kinetics (i.e., concentrations versus time) can be converged by including the contributions from more and more paths. The details of these steps are now briefly reviewed.

The important chemical pathways may be enumerated in one of several ways. In a brute force approach, the chemical pathways can be identified and roughly ranked in importance using a small stochastic simulation using an ensemble of Monte Carlo trajectories in which a sample of molecules are followed through the network. Unlike the more common applications of stochastic simulation methods, the present approach follows a single tagged atom from reactants to the target species.[69, 31, 29, 32] The pathways that are followed by the most molecules are then used for the representation. Alternatively, we may generate the paths using a graph theoretic enumeration procedure. In the language of graph theory, the species are vertices (or nodes) and the reactions connecting them are edges and the chemical pathways are n-step paths on the graph. Because the reactions have a direction the graph is directed and if more than one edge connects two vertices it is a multigraph. The edges of the graph are weighted by the kinetic flux for the product formation due to the edge reaction. Search algorithms may be employed to identify the most probable pathways on the weighted graph.

The next task is to compute the probability for a general pathway consisting of n-reactions and $n + 1$ species, i.e., $S_0 \xrightarrow[R_1]{t_1} S_1 \xrightarrow[R_2]{t_2} S_2 \cdots S_{n-1} \xrightarrow[R_n]{t_n} S_n$, where the reactions obey the time ordering $t_f > t_n > \cdots > t_1 > t_0$. Because the occurrence of each reaction constitutes a random event, we must integrate the probability density over all allowed times consistent with the time ordering. To find the probability for a given pathway, We adopted the importance sampling based technique mentioned in Section 2.3 using a modest number of MC string.

The final step in the method is to combine the pathway probabilities to obtain the desired observable. As discussed in Section 2.4, most observables of interest can be expressed in terms of the chemical concentrations. Thus, we describe how to compute the concentration of some species S_i at a time t_f , $X_i(t_f)$, given the initial conditions $\mathbf{X}(t)$ at time t_0 . We follow a specific atom found in $X_i(t_f)$ that may potentially originate from one of several initial species, S_k , each of which may contain more than one chemically equivalent followed atom, viz. μ_k . The pathways are labeled by j_k where k specifies the species of origin for the path of the followed atom. The probability of the atom following pathway j_k from initial species S_k arriving at the target species S_i at time t_f is given by $P_{j_k}(t_f)$. Then the concentration of S_i is given by eqn. 2.18. Where we define X_i to be the concentration of species S_i . Each chemically equivalent followed atom in a reagent species is sampled equally and the sum of all probabilities for a given followed atom is 1, and thus, $\sum_{j_k} P_{j_k}(t) = 1$, at any time $t > t_0$.

4.2.2 Sensitivity Analysis

In sensitivity analysis, we define a target observable, τ , that is computed using solutions to the kinetics equation eqn. 2.2 and is a function of the variable factors. Examples of τ might be the concentration of a species S_i at a specific time or the ignition delay time in a combustion problem. To be explicit, we shall consider the variable factors to be the L rate coefficients, k_i , and we assume the initial conditions of the reaction are fixed. Thus, we have a well defined target function

$$\tau = \tau(k_1, \dots, k_L) \quad (4.2)$$

The local variation of τ with respect to k_i in the vicinity of the nominal mechanism, \mathbf{k}_0 , is measured by the derivative $\partial\tau/\partial k_i$. A dimensionless local sensitivity index may be defined using the form

$$s_i = \left. \frac{\partial \ln(\tau)}{\partial \ln(k_i)} \right|_{\mathbf{k}=\mathbf{k}_0} \quad (4.3)$$

or through the normalized expression

$$w_i = \frac{\left| \frac{\partial \tau}{\partial k_i} \right|}{\sqrt{\sum_j \left(\frac{\partial \tau}{\partial k_j} \right)^2}} \Bigg|_{\mathbf{k}=\mathbf{k}_0} \quad (4.4)$$

If the factors k_i enter eqn. 4.2 in a highly nonlinear fashion, or if ranges δk_i are large, then it is advisable to analyze the kinetics using Global Sensitivity Analysis.[74, 17, 18, 70, 103, 103, 98, 81, 84] The global sensitivity analysis is carried out using the analysis of variance (ANOVA)[88] strategy with the aid of the high dimensional model representation (HDMR).[77, 61, 62, 108, 107, 106, 87, 51, 20, 104] The global sensitivity index for rate coefficient k_i is then given by the ratio of the partial variance to the total variance of the target function, i.e., the main effect

$$W_i = \frac{\sigma_i^2}{\sigma_T^2} \quad (4.5)$$

The total variance is given by $\sigma_T^2 = \langle \tau^2 \rangle - \langle \tau \rangle^2$ where the bracket indicates an average over the uncertainty hypercube in the L-dimension rate coefficient space. The partial variance is obtained using the HDMR expansion

$$\tau(k_1, k_2, \dots, k_N) = \tau_0 + \sum_{i=1}^L F_i(k_i) + \sum_{i>j} G_{i,j}(k_i, k_j) + \dots \quad (4.6)$$

where $\tau_0 = \langle \tau \rangle$. Because it can be shown that $\langle F_i \rangle = \langle F_i F_j \rangle_{i \neq j} = \langle G_{i,j} \rangle = \langle F_i G_{n,m} \rangle = \dots = 0$ if the component functions are expanded using orthogonal polynomials, then the partial variance at first-order is given by

$$\sigma_i^2 = \langle F_i^2 \rangle \quad (4.7)$$

The L component functions, $F_i(k_i)$, are obtained using a least-squares fitting of a Legendre expansion, $F_i(k_i) = \sum_{j=1}^n c_j P_j(k_i)$. Similarly, the higher order component functions are obtained using the product form, e.g., $G_{i,j}(k_i, k_j) = \sum_{i',j'}^n c_{i',j'} P_{i'}(k_i) P_{j'}(k_j)$. Thus, the uncertainty hypercube is randomly sampled using a large number M of Monte Carlo values, $(k_1, \dots, k_L)_l$, $l = 1, \dots, M$, for which τ is computed, and this sample is used to determine all of the component functions. The HDMR expansion is carried out to higher and higher order until the predicted variance is sufficiently close to the exact variance obtained from the data set. Thus, if the expansion in first-order

component functions F_i proves inadequate, then some or all the second-order terms, $G_{i,j}$, are included, which then incorporates part of the interaction between the factors k_i . An advantage of global sensitivity analysis over local sensitivity analysis is that the global index not only reflects the response of the target to variation of the factors but also accounts for the size of the uncertainty range of the factors. Thus, a large global sensitivity coefficient suggest that an improved measurement of k_i could lead to a large reduction in the overall uncertainty of the model.

In the present work, we shall focus on the most basic set of target functions of the rate coefficients, viz. the S_i species concentration at time t , i.e., $\tau = X_i(t)$. The sensitivity indices are obtained by simply solving the differential eqn. 2.2 from fixed initial conditions, $\mathbf{X}(t = 0)$ while varying k_j 's. These quantities are a probe of the speciation of the reaction network which may, under some circumstances, be directly measured in experiments.

The target functions $X_i(t)$ are among many examples of kinetic observables that can be written as linear combinations of the contributions of separate chemical pathways via

$$\tau = \sum_i c_i p_i \quad (4.8)$$

In eqn. 4.8, p_i are the pathway probabilities and the constants c_i do not depend explicitly on p_i . From eqn. 2.18 we see that for speciation the c_i 's are given by the initial concentrations and the stoichiometric coefficients and do not depend on the rate coefficients. Hence the sensitivity of this target to the k 's is carried totally by the pathway probabilities themselves.

4.3 Sensitivity Analysis of Hydrogen Combustion Chemistry

The combustion of hydrogen is often chosen as a model system to investigate the performance of new chemical kinetic techniques. Though the chemistry of this problem is fairly simple by the standards of combustion kinetics, it exhibits most of the generic features seen in the burning of more complicated fuels while at the same time being describable by a manageable small mechanism. The H₂-O₂ system is described here by a model consisting of 19 reversible reactions (plus two duplicate reactions) and 8 distinct chemical species. The reactions and species are labeled using a numbering

scheme defined in Table 4.1 and Table 4.2. The corresponding rate coefficients and uncertainty ranges are provided in the Appendix A. The simulations were carried out using a stoichiometric mixture of pure H₂-O₂ at an initial temperature of T = 1000 K and an initial pressure of p = 5.62 bar. The calculations were carried out while the volume and total energy were kept fixed. Under these conditions, the mixture exhibits ignition at $\tau_{ign} = 3.4$ ms. To illustrate the general behavior

Table 4.1: Species Labeling of H₂-O₂ system.

Species Index	Species Name
0	O ₂
1	H ₂ O
2	H ₂
3	H ₂ O ₂
4	H
5	OH
6	HO ₂
7	O

[†]Indices start at 0.

of the kinetics, the concentrations versus time are plotted on a logarithmic scale in Fig. 4.1. It is seen that the radical pool shows a steady increase until it spikes upward at the ignition time, τ_{ign} . For the present work, we shall consider the sensitivity exhibited by the chemistry for times before τ_{ign} . The highly reactive radicals O, H, and OH are seen to be roughly in steady state for a significant fraction of the time shown. The concentration of the hydroperoxy radical, HO₂, is seen to exhibit substantially more time variation than the other radicals. In this preignition phase of the reaction, the contributions from some of the more important elementary reaction steps are shown in Fig. 4.2 where we plot the logarithm of the instantaneous rates versus time. We note in particular the behavior of reactions R13 (HO₂ + HO₂ → H₂O₂ + O₂) and R17* (HO₂ + H₂ → H₂O₂ + H), which are both sink reactions for HO₂ and source reactions for H₂O₂. (The asterisk indicates the backward reaction in Table 4.2.) There is clearly a crossover behavior where the abstraction path (R17*) is largest at short times and the radicalradical disproportionation reaction (R13) becomes larger at longer times. We also notice that the rate for R15 (H₂O₂ + M → 2OH + M), which is

Table 4.2: Reaction Labeling of H₂-O₂ system. The backward reactions, where the right-hand side is the reagent, are denoted with an * in the text; i.e., R17* denotes HO₂ + H₂ → H₂O₂ + H and R17 denotes H₂O₂ + H → HO₂ + H₂.

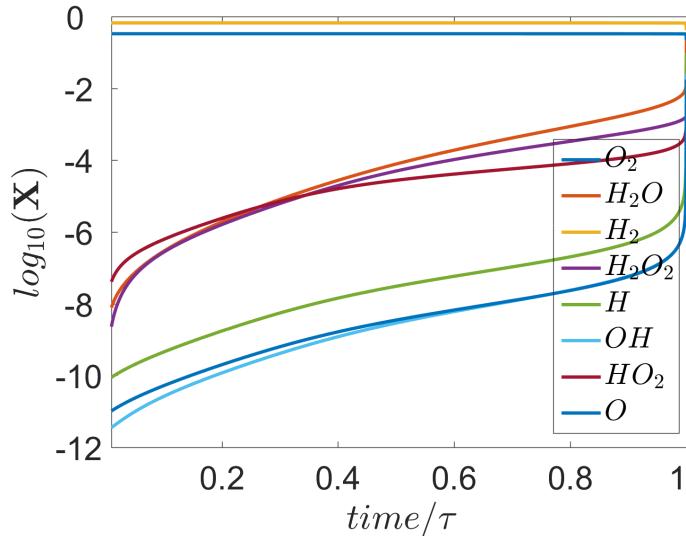
Species Index	Species Name
0	H + O ₂ = O + OH
1	O + H ₂ = H + OH
2	H ₂ + OH = H ₂ O + H
3	O + H ₂ O = OH + OH
4	H ₂ + M = H + H + M
5	O + O + M = O ₂ + M
6	O + H + M = OH + M
7	H + OH + M = H ₂ O + M
8	H + O ₂ (+M) = HO ₂ (+M)
9	HO ₂ + H = H ₂ + O ₂
10	HO ₂ + H = OH + OH
11	HO ₂ + O = O ₂ + OH
12	HO ₂ + OH = H ₂ O + O ₂
13	HO ₂ + HO ₂ = H ₂ O ₂ + O ₂
14	duplicated reaction
15	H ₂ O ₂ (+M) = OH + OH (+M)
16	H ₂ O ₂ + H = H ₂ O + OH
17	H ₂ O ₂ + H = HO ₂ + H ₂
18	H ₂ O ₂ + O = OH + HO ₂
19	H ₂ O ₂ + OH = HO ₂ + H ₂ O
20	duplicated reaction

[†]Indices start at 0.

an important sink for H_2O_2 , crosses that for both R13 and R17*. We also show, in Fig. 4.3, the behavior of the species decay rate (defined as $A_i(t)$ in eqn. 2.11) for all eight species versus time. It is seen that the chemical lifetimes of the key species H_2O_2 and HO_2 are on the order of 10^{-3} s, which is comparable to the ignition time itself.

The global sensitivity indices for the speciation targets $\tau = X_i(t)$ for $i = 18$ were computed at

Figure 4.1: Concentrations versus time obtained from the combustion simulation. The simulation was carried out at $T=1000$ K, $p=5.62$ bar, and $\Phi=1$. No buffer was present and the reaction occurred in a constant volume and constant energy. We set τ to be the ignition delay time.



a grid of times between $t = 0$ and the ignition time, τ_{ign} . These computations employed numerical solutions of the kinetic rate equations, eqn. 2.2, with rate coefficients randomly selected from within their uncertainty ranges. For rate coefficients of the form $k(T) = AT^\delta \exp(-E_a/(RT))$, this sampling is accomplished by varying the value of the pre-exponential factor A. The rate coefficients reverse reactions are set by detailed balance and thus are not independent from the forward rates. We note that the sensitivity index for a forward reaction, such as R17, is thus exactly equal to that for the backward reaction, R17*, regardless of the value of the reactive flux in the forward and backward directions. Using 5000 sets of random rate coefficients, the HDMR component functions

Figure 4.2: Rates of the most important reactions versus time in units of the ignition time. The reaction index is defined in Table 4.1 and the * notation indicates a backward reaction.

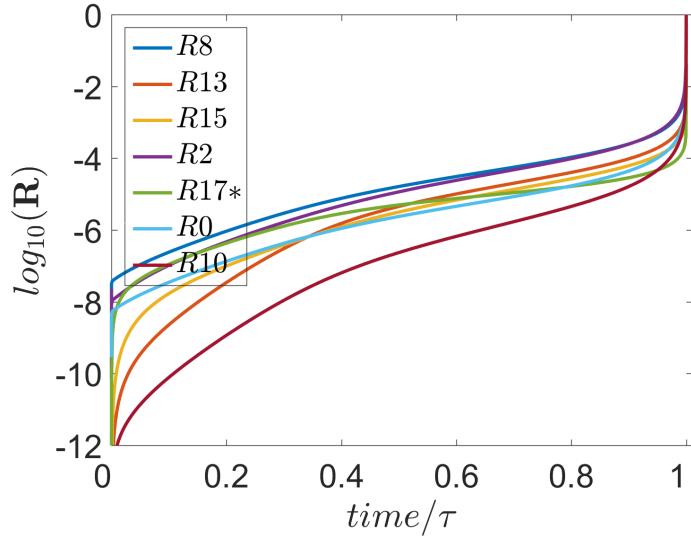
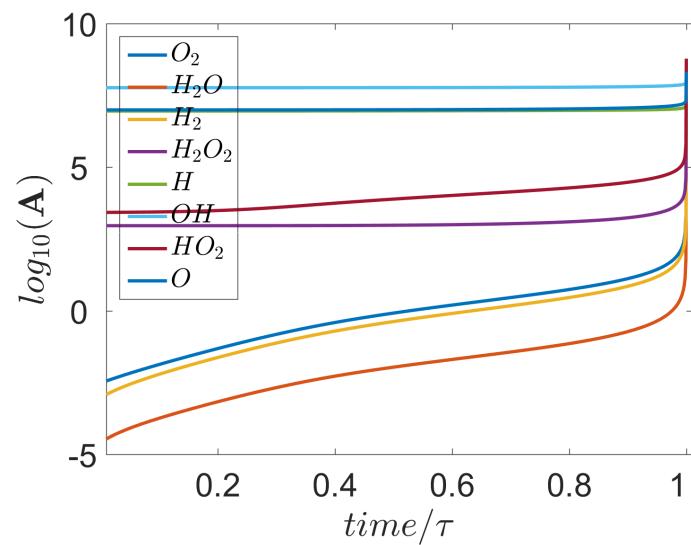


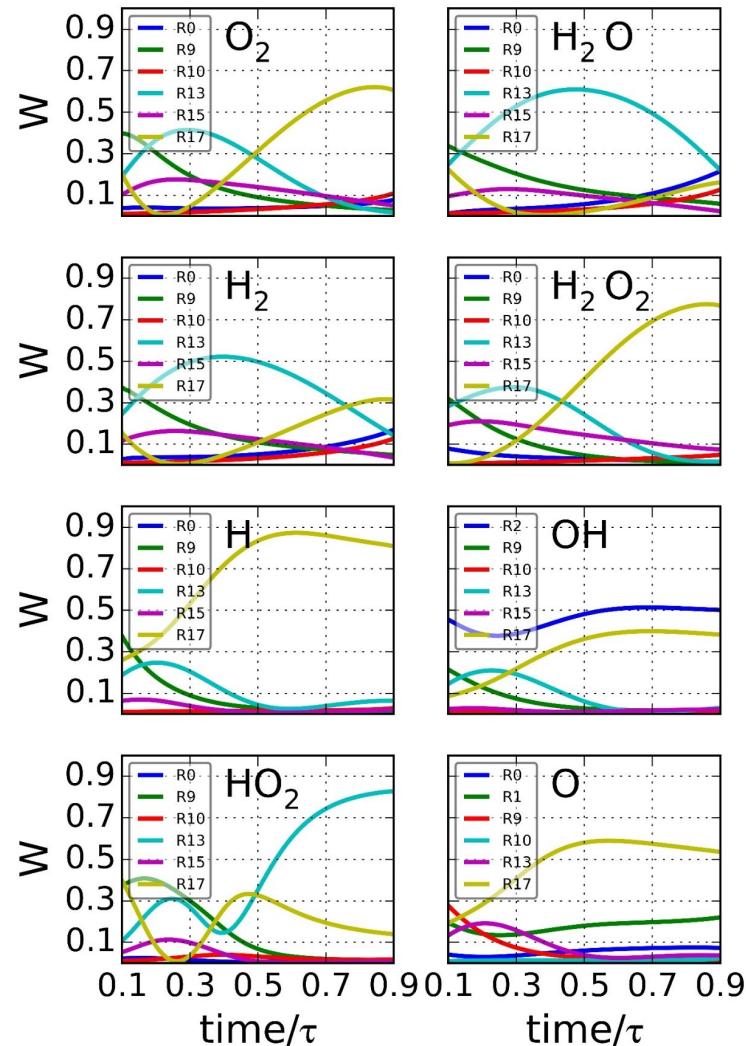
Figure 4.3: Species decay rate, i.e., the sum of the rates of all the sink reactions, plotted versus time for all eight species.



$F_i(k_i)$ from eqn. 4.6 were fit by regression to fifth-order expansions in Legendre polynomials in a single fitting procedure consisting of $19 \times 6 = 114$ parameters. As in our previous work, the sensitivity indices were then obtained by integration using eqn. 2.18.[87, 51, 20, 104] In Fig. 4.4 we show the sensitivity indices W_i versus time for all 8 species. For clarity, we have included in the plot only the most important reactions. It is seen that reactions R17, R15, R13, R9, and R0 have consistently the highest values of W_i . We note that reactions R17, R13, and R9 involve the production and destruction of the HO₂ radical, which is known to play an important role in determining the ignition properties of H₂ and other fuels.[101, 49, 48] Reactions R17, R15, and R13 also involve the H₂O₂ species, which is also important in the generation of the radical OH.[58, 105] Reaction R0, H + O₂ = OH + O, is well-known to be a critical chain branching step in all combustion reactions. Here, R0 is seen to be particularly important for the formation of OH radicals. The species O is unique in being the only target that is significantly sensitive to the reaction R1, O + H₂ = OH + H.

Despite differences in the shape of the sensitivity curves for the various species, there is actually a significant amount of similarity in overall behavior of the results. All the species exhibit large values at early times for the initiation reaction H₂ + O₂ = HO₂ + H (R9), which then falls off sharply at later times. This is understandable because reaction R9 is the main initiator for autoignition but is not important for chain propagation and chain branching. During the later stages of the reaction, $t \geq 0.4\tau_{ign}$, reactions R17 (HO₂ + H₂ = H₂O₂ + H) and R13 (HO₂ + HO₂ = H₂O₂ + O₂) usually carry the largest sensitivities, as occasionally does R15 (H₂O₂ + M = M + 2OH). It is interesting to note for many targets we see a trend where W_{17} is more sensitive at earlier times but then is surpassed by W_{13} , which grows larger at later times. For all species except H and O, there is a crossover where $W_{17} > W_{13}$ at short times but then $W_{13} > W_{17}$ at longer times. That the sensitivity coefficients of the different species in a large model behave similarly has been noted previously by Skodje and Davis[51] and Turanyi et al;[109] this behavior is related to the fact that the kinetics tends to rapidly approach a low-dimensional manifold in which the species concentrations are slaved to one another. The equivalence between the different sensitivity plots is only qualitative

Figure 4.4: Global sensitivity index W for the speciation targets $[X_i(t)]$ versus time. The results shown are restricted to the most sensitive reactions for each case.



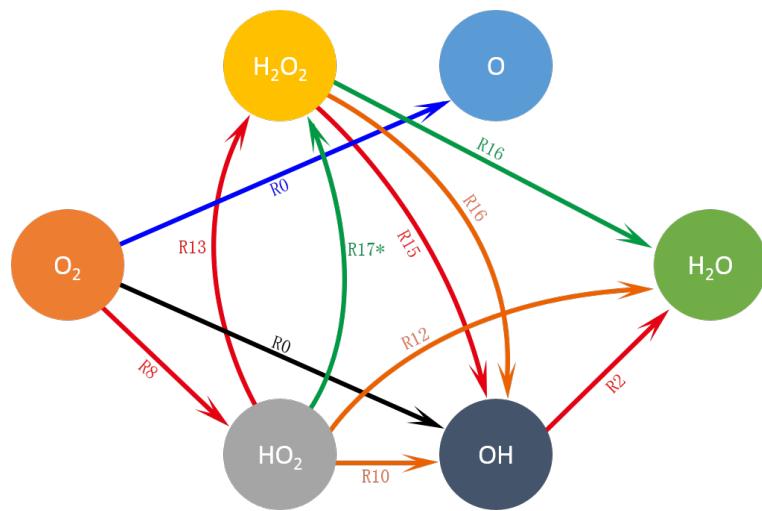
in the present case because for global analysis the different values of the uncertainty ranges, Δk_i , introduces quantitative differences between the various W_i . This can be appreciated by noting that global sensitivity index can be crudely approximated by $W_i \approx \left| \frac{d[X(t)]}{dk_i} \times \Delta k_i \right|$, where $\mathbf{X}(t)$ is the species concentration computed using the nominal (unperturbed) rate coefficients. Thus, clearly, the relative values of W_i can be adjusted using the externally provided uncertainty ranges.

4.4 Description of Chemical Pathways for H₂-O₂ Combustion

When the species target functions are expressed as linear combinations $\tau = \sum_i c_i p_i$, the sensitivity of τ to \mathbf{k} can be understood in terms of the behavior of the pathway probabilities. Hence, we must first determine a sufficient number of chemical pathways and their probabilities to converge the speciation targets of interest, $X_i(t)$. Then, we investigate the \mathbf{k} dependence of the probabilities. In our previous work, we studied the early stage combustion of hydrogen using an H atom following scheme. It was found that convergence of the concentrations to within several percent was obtained using about a dozen chemical pathways. Here, we reexamine the pathway representation using, instead, O atom following pathways. Although convergence to the exact result may be obtained by following any moiety present in the target species, the choice of followed atom may affect the efficiency of the algorithm. In the present case, we find that O atom following converges significantly faster than H atom following. For reference, in Fig. 4.5 we show the full graph that shows all species (nodes) and the most important reactions (lines) that contribute to the present mechanism. The O atom following pathways lie on the subgraph containing only the oxygen possessing species.

To identify a set of candidate pathways for study, we first use a small size stochastic simulation, i.e., $10^5 - 10^6$ simulated paths, computed at the nominal value of \mathbf{k} . Thus, during a time interval Δt , a given molecule of S_i has a decay probability of $A_i(t)\Delta t$. Using one random number to sample this decay, and a second random number to select a product branch j , $\Gamma_j(t)$, a given O atom is then followed through the network one node (species) at a time. Recording each pathway that occurs, an approximately ordered set of candidates is produced that describes the passage of a followed atom

Figure 4.5: Graph representing all species and the most important reactions for the H₂-O₂ system by following Oxygen atom.



to a given target. For the present problem, typically about 100 pathways were identified in this way, although many fewer were needed to converge the targets. Although this calculation is not efficient enough to compute accurate pathway probabilities, it does yield the important pathways for high level computation using eqn. 2.15 or 2.16. Because the ordering of the pathway probabilities will be a function of \mathbf{k} , we retain some of the smaller probability pathways because they may become larger in certain regions of k -space.

We begin by showing the most probable pathways for the creation of several important target species. Consider first the formation of the transient intermediate H_2O_2 which was used as a test case in our previous work. Although we see from Table 4.2 that there are six elementary reactions that generate H_2O_2 , we find that just two chemical pathways are sufficient to converge $[\text{H}_2\text{O}_2]$ to within 1% at times in the range $0 < t < 0.9\tau_{ign}$. In Fig. 4.6 we show the pathway probabilities for these paths, along with the prediction for $[\text{H}_2\text{O}_2(t)]$. Many more paths were required to converge the results using the H atom following algorithm. The two dominate chemical pathways for the formation of this species are path a3 (where a means the most probable pathway and 3 indicates that H_2O_2 is species 3 in Table 4.1)



and path b3 (b for the second most probable pathway, etc.)



In the chemical pathway formulas, we indicated the followed atom in red, the nonfollowed reagents are in parentheses, and the nonfollowed products are dropped from the equation. We can see by the figure that the pathway probabilities for H_2O_2 formation show a crossover behavior where the path involving R_{17}^* is largest at short times whereas the path containing R_{13} is larger at long times. The remaining pathways we computed were lower in probability by factors of 100 or more.

The next three pathways are

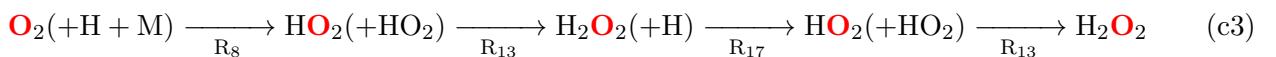
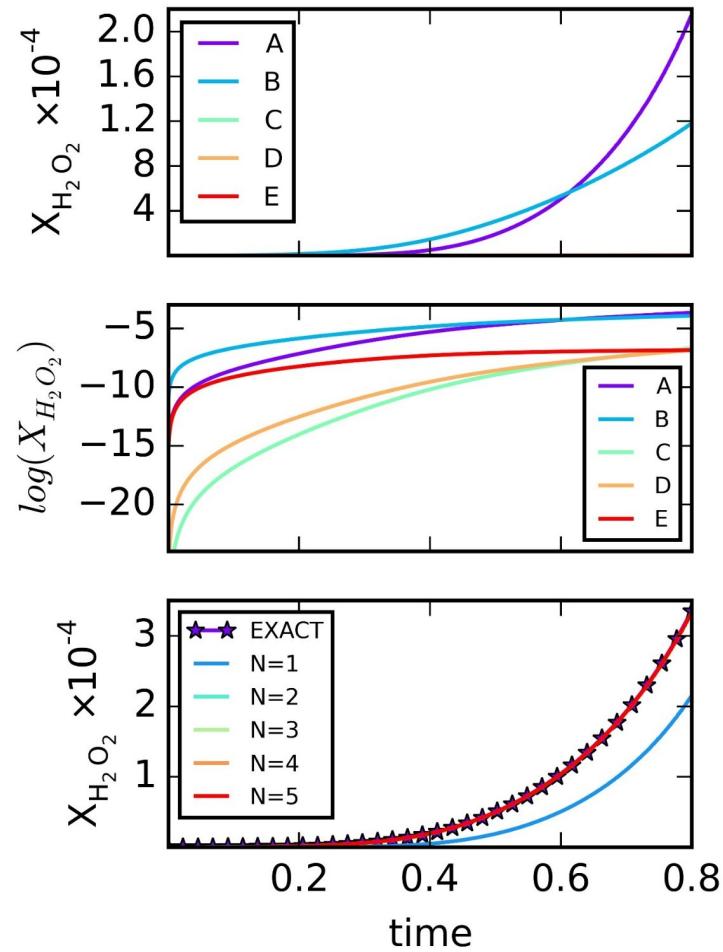
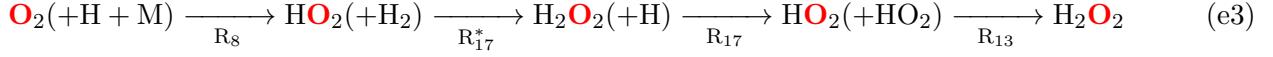


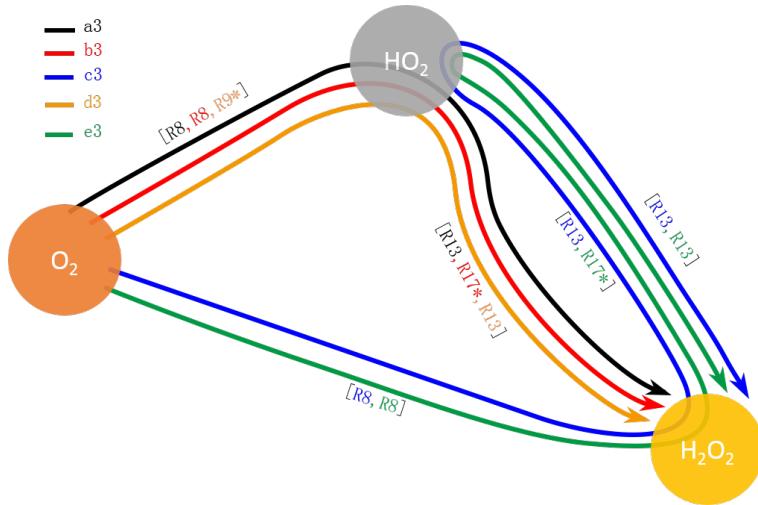
Figure 4.6: Pathway probabilities for H_2O_2 . In the upper panel, the probabilities for the five most important pathways, ranked at time $t = 0.8\tau_{ign}$, are shown versus time in units of τ . In the middle panel, the same probabilities are reproduced on a logarithmic scale. In the lower panel, the convergence of the concentration to the exact simulation result is shown; it is seen that probability converges to good accuracy with just two paths.





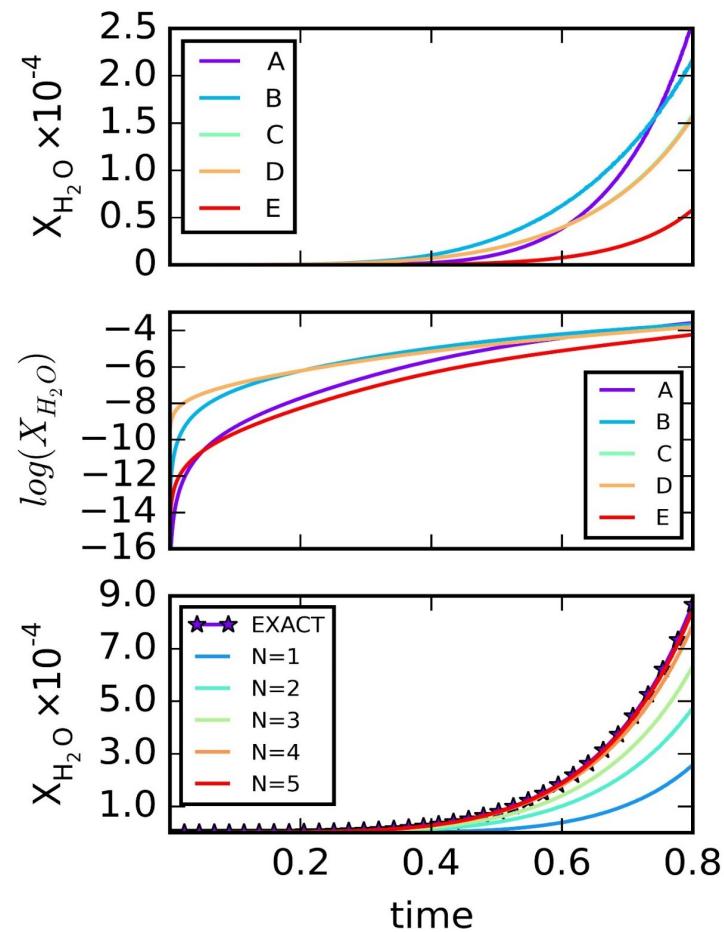
It is interesting to note that pathways c3 and e3 are quite closely related to the main pathways a3 and b3. We see that in mechanism c3, the H_2O_2 is formed along pathway a3 but then is recycled back to HO_2 by H atom attack, , only then to regenerate H_2O_2 by radicalradical recombination, R13. In mechanism e3, we see the same scenario except the initial H_2O_2 molecule is generated by path b3. In mechanism d3, the H_2O_2 is formed by radical recombination R13 but the followed species HO_2 is generated by the initiation reaction R9, i.e., $\text{H}_2 + \text{O}_2 \rightarrow \text{HO}_2 + \text{H}$ rather the higher flux reaction R6. For reference, we indicate the contributing pathways on the O atom subgraph in Fig. 4.7.

Figure 4.7: Graph for the oxygen following paths, a3 - e3, leading to the H_2O_2 product.

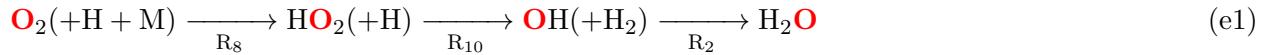
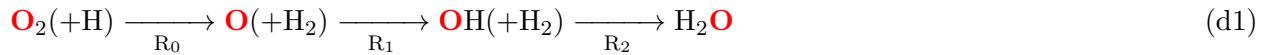
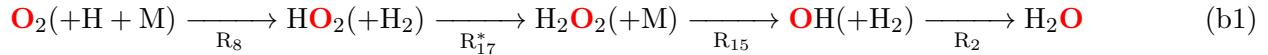
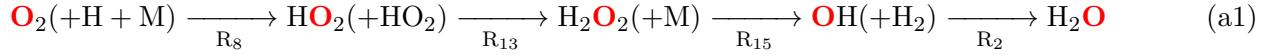


For the H_2O product species (species 1 in Table 4.1), we find that there are more contributing pathways than for H_2O_2 . We find that five paths account for about 98formation under the present conditions. The convergence of the target $[\text{H}_2\text{O}(t)]$ versus time is shown in Fig. 4.8. The most

Figure 4.8: Same as Fig. 4.6 except for the target function $[H_2O(t)]$.



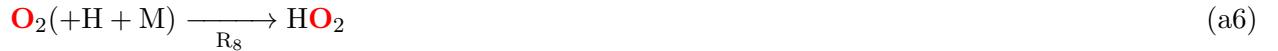
important O atom following pathways that connect the O_2 reagent to the H_2O product are



The two most probable pathways, a1 and b1, again generate an H_2O_2 intermediate from HO_2 radicals, which indicates that a large fraction of the H_2O product passes through this intermediate species. Indeed, the first three steps for a1 and b1 are seen to be identical to the dominate pathways for H_2O_2 formation, i.e., a3 and b3. In both a1 and b1, the H_2O_2 breaks down via the unimolecular dissociation reaction R15 with the resulting OH radical converted to water through the reaction R2, i.e., $\text{OH} + \text{H}_2 \rightarrow \text{H}_2\text{O} + \text{H}$. From the figure, we again note a crossover behavior where path a1 (containing R13) is more probable at long times whereas path b1 (contain R17*) is more probable at shorter times. The three next most probable pathways are mechanistically distinct from the H_2O_2 forming pathways a1 and b1. It is seen in Fig. 4.8 that the contributions from c1 and d1 are nearly identical because it is clear that these paths are closely related. In these two pathways, we are tracking the separate progress of the two oxygen atoms of O_2 following the initiation reaction R0, i.e., $\text{O}_2 + \text{H} \rightarrow \text{OH} + \text{O}$. In the first case, c1, the OH radical from R0 reacts quickly with H_2 to produce the water product in a single step, R2. Along the other path, d1, the O atom product of R0 first abstracts an H atom via R1 to form OH and then the OH reacts with H_2 to form water. Because nearly all the OH is consumed to form H_2O and nearly all the O reacts with H_2 to form

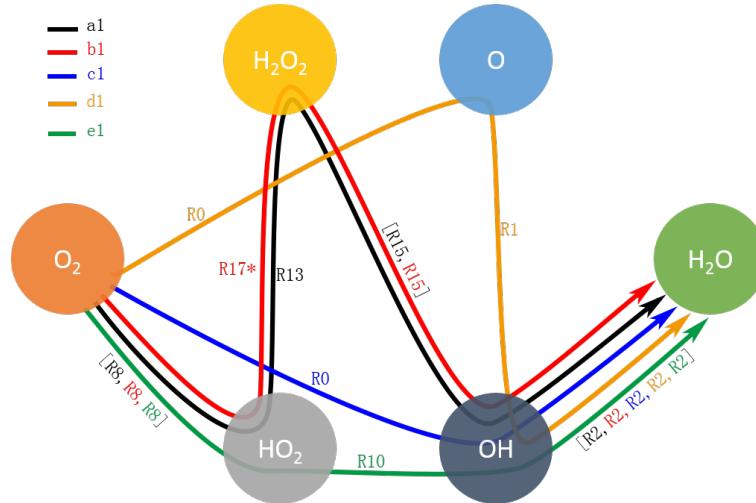
OH, it is not surprising that the probabilities for c3 and d3 are very close. In both cases, the species containing the followed O atom quickly forms water upon reaction with H₂. Finally, pathway e5 again is initiated by formation of HO₂ through the recombination reaction R8. However, in e5 the HO₂ radical reacts with H via R10 to form the OH radical, which then quickly converts to H₂O by R10. The five pathways are depicted graphically in Fig. 4.9.

Finally, we consider the speciation of the critical HO₂ radical. In Fig. 4.10, we show the relevant pathway probabilities for of the [HO₂(t)]. It is very interesting to note that one pathway carries the vast majority of the probability. Specifically, almost all of the HO₂ species is generated by the recombination mechanism



As seen in the figure, the probabilities for other production pathways are lower by at least a factor

Figure 4.9: Oxygen atom following pathways leading to the formation of the H₂O product.



of 1000. For completeness, however, we note that the algorithm does quantify the probabilities for the other minor pathways

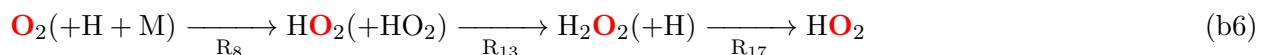
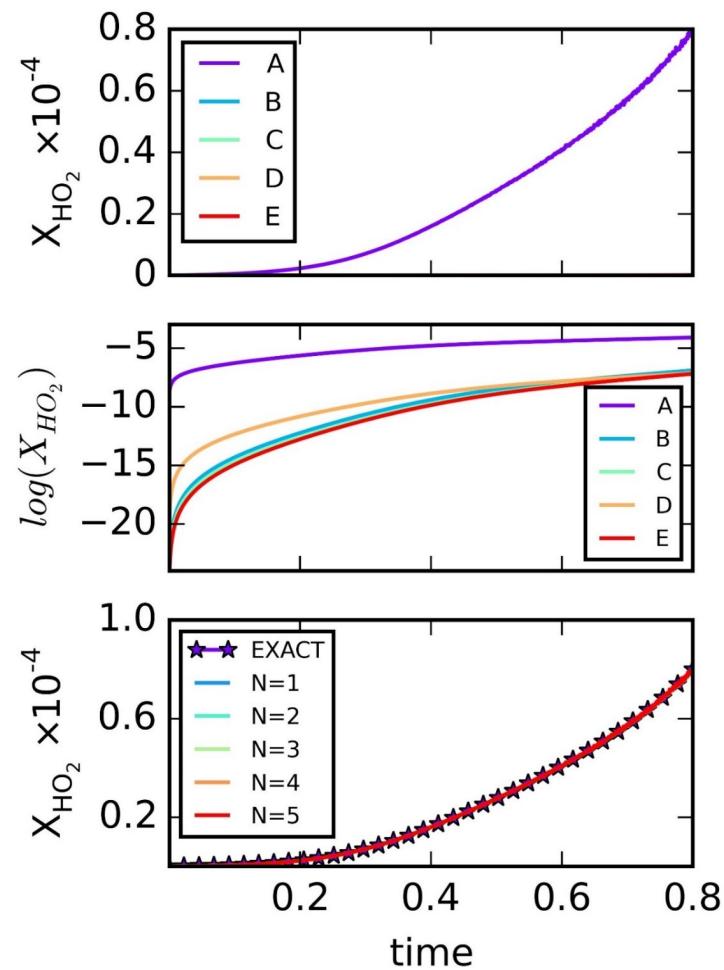


Figure 4.10: Same as Fig. 4.6 except for the HO₂ product.



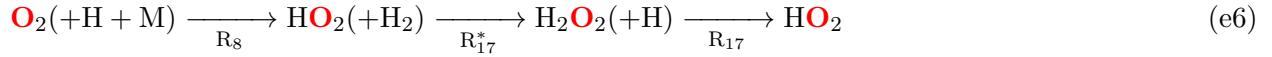
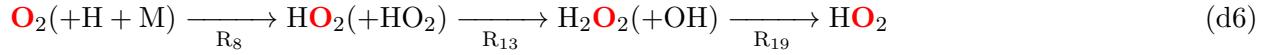
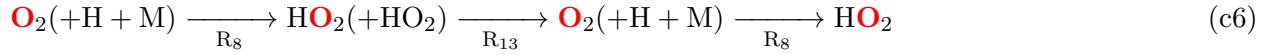
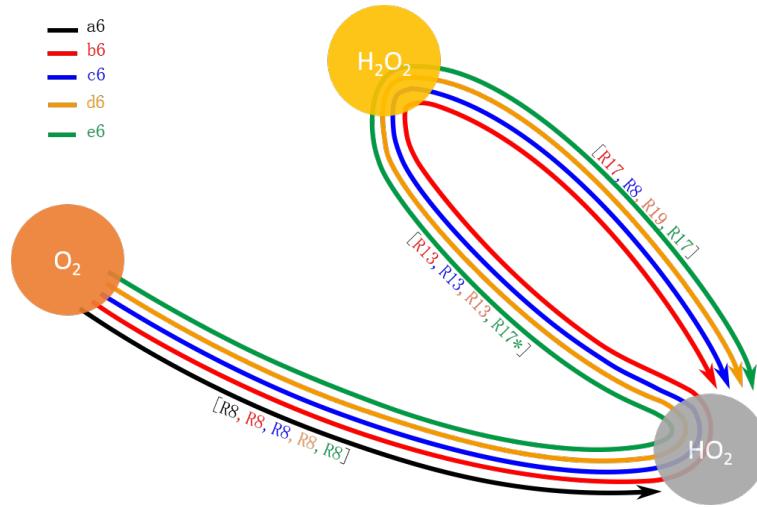


Figure 4.11: Oxygen atom following pathways leading to the formation of the HO_2 product.



It is seen that b6–e6 are cyclic paths. For all these pathways, the HO_2 radical is first generated by recombination but then reacts to form other species, such as H_2O_2 , only then to regenerate the HO_2 radical. These pathways are depicted in graphical form in Fig. 4.11. Because we are using O atom following, the origin of the H atom in these pathways is not specified. Obviously, the formation of H radicals does involve important rate limiting behavior for the formation of HO_2 .

4.5 Pathway Interpretation of Sensitivity

Now that the dominant pathways that lead to various species targets have been uncovered, we can interpret the sensitivity results presented in Fig. 4.4 using the sum over histories approach

via $\tau = \sum_i c_i p_i$. Thus, each probability p_i is a function of \mathbf{k} and has a total variance over the uncertainty range of $\sigma_T^{i,2}$. To quantify the global sensitivity of the pathway probabilities with respect to the individual rate coefficient k_j , we express each p_i as a first-order HDMR expansion, i.e.

$$p_i(k_1, k_2, \dots, k_L) = p_0^i + \sum_{j=1}^L F_j^i(k_j) \quad (4.9)$$

The component functions for pathway i , $F_j^i(k_j)$, were expressed as fifth-order expansions in Legendre polynomials fit using linear regression from 2000 sets of rate coefficients that were randomly selected from the uncertainty range. We have verified by direct computation that the totality of the second order terms of the HDMR were small, typically on the scale of $\sim 5\%$ of the first-order terms, although they are occasionally larger. We have also verified that there was not over-fitting of the first-order component functions. We determine the first order global sensitive coefficient for the pathway probability p_i with respect to rate coefficient k_j using the normalized partial variance

$$\begin{aligned} \widehat{W}_j^i &\equiv \frac{\sigma_j^{i,2}}{\sigma_T^{i,2}} \\ &= \frac{\langle F_j^{i,2} \rangle}{\sigma_T^{i,2}} \end{aligned} \quad (4.10)$$

which is completely analogous to eqs 4.5 and 4.7. Correlations between pathways may be gauged using the coefficients

$$Corr(i, i') = \frac{\langle p_i p_{i'} \rangle - \langle p_i \rangle \langle p_{i'} \rangle}{\sqrt{\sigma_T^{i,2} \sigma_T^{i',2}}} \quad (4.11)$$

If the first-order HDRM expansion, eqn. 4.9, is inserted into eqn. 4.11, it is seen that expressions of the following form appears.

$$W_j^{i,i'} = \frac{\langle F_j^i F_j^{i'} \rangle}{\sigma_T^{i,2} \sigma_T^{i',2}} \quad (4.12)$$

The total correlation between paths i and i' may thus be decomposed into contributions from the individual rate coefficients as

$$Corr(i, i') = \sum_{j=1}^L \widehat{W}_j^{i,i'} \quad (4.13)$$

We should note that this pathway correlation, $Corr(i, i')$, reflects how the change in flux along path i affects the flux along path i' ; it is not a correlation between individual reaction steps which

would require a higher order HDMR expansion to model.

Although the general chemical flux in the reaction network may depend on the rate coefficients in a complicated manner, there are two scenarios that are particularly simple for interpretation. In the first, a single chemical pathway dominates the flux to the target function $X_n(t)$. Under these circumstances, the rate limiting step (i.e., the slowest step) to the target, and the fastest step away from the target may prove the most sensitive. For simplicity of illustration, assume that reactions along the path are described by pseudo-first-order kinetics with constant decay rates and branching ratios. The rate limiting step to the target $X_n(t)$ is modeled by a pseudo-first-order rate law $K_J X_J$ and the fastest decay channel for S_n is described by $K_n X_n$. We therefore expect the speciation target function to behave as

$$X_n(t) \approx X_0 \cdot \frac{K_J}{K_n - K_J} [e^{-K_n t} - e^{-K_J t}] \quad (4.14)$$

where X_0 is a constant. This approximate expression applies at times that are long compared to the transient rates along the pathway. Clearly, the target will show strong sensitivity to the rate limiting formation step and/or the dominate decay step and only weak sensitivity to the other steps. Although this result is not surprising, it does serve to emphasize the connection between pathways and sensitivity. A different scenario occurs for pathway switching in which a sensitive rate coefficient causes the reactive flux to change paths. Consider the case where the largest flux pathway splits into two branches (labeled a and b) that subsequently flow to the target $X_n(t)$. If we again assume pseudo-first-order kinetics, we have the branching ratios

$$\begin{aligned} \Gamma_a &= \frac{k_a}{k_a + k_b} \\ \Gamma_b &= \frac{k_b}{k_a + k_b} \end{aligned} \quad (4.15)$$

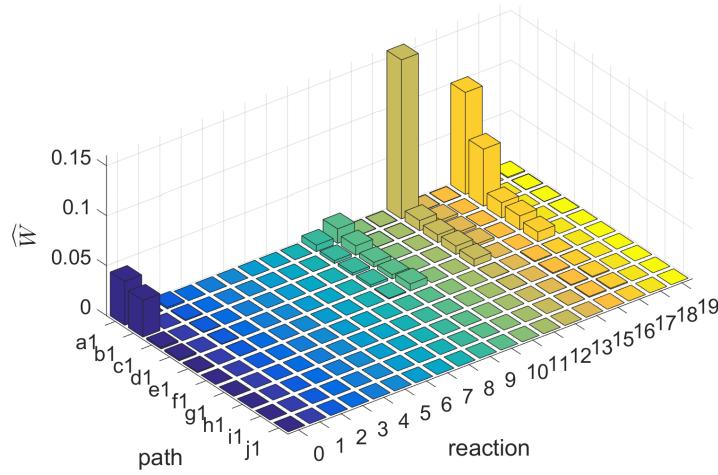
where k_a and k_b are the conventional rate coefficients at the branching step. Then the contribution to the target will be roughly given by

$$X_n(t) \approx c (\Gamma_a p_a + \Gamma_b p_b) \quad (4.16)$$

where c is a constant and p_a and p_b are the probabilities for paths a and b. Thus, we expect there to be a large sensitivity of the target to k_a and k_b provided that p_a and p_b are significantly different

numerically. We also predict a high level of correlation between the factors k_a and k_b .

Figure 4.12: Sensitivity indices of the pathways leading to the H_2O target concentration evaluated at $t = 0.8\tau_{ign}$. In this figure we plot $F_j^{i,2}/\sigma_T^{i,2}$, where each path has a common normalizing factor so that less probable pathway show smaller indices.



As an application, consider the $[\text{H}_2\text{O}]$ target, which possesses the most interesting and complicated pathway representation. We focus first on a single time, $t = 0.8\tau_{ign}$. From the global sensitivity results of Fig. 4.4, we see that the most important reactions for the H_2O formation at this time are, in descending order, R13, R0, R17, R9, and R10. We can understand this pattern from the behavior of the paths. At $t = 0.8\tau_{ign}$, the pathway probabilities are 0.264 (a1), 0.229 (b1), 0.187 (c1), 0.184 (d1), and 0.074 (e1). In Fig. 4.12 we show the spectrum of sensitivity indices, i.e., \widehat{W}_j^1 versus j , obtained for the largest pathways ranked by order of probability. In this figure, we have plotted sensitivity indices using a common normalization factor, i.e., $\langle F_j^{i,2} \rangle / \sigma_T^2$ where σ_T^2 is the total variance for all paths, so that the less important pathways show lower indices. In Fig. 4.13 we show the sensitivity spectrum for the five most probable paths, ale1, using the conventional normalization, eqn. 4.10. First, we note that the reactions that have the largest value

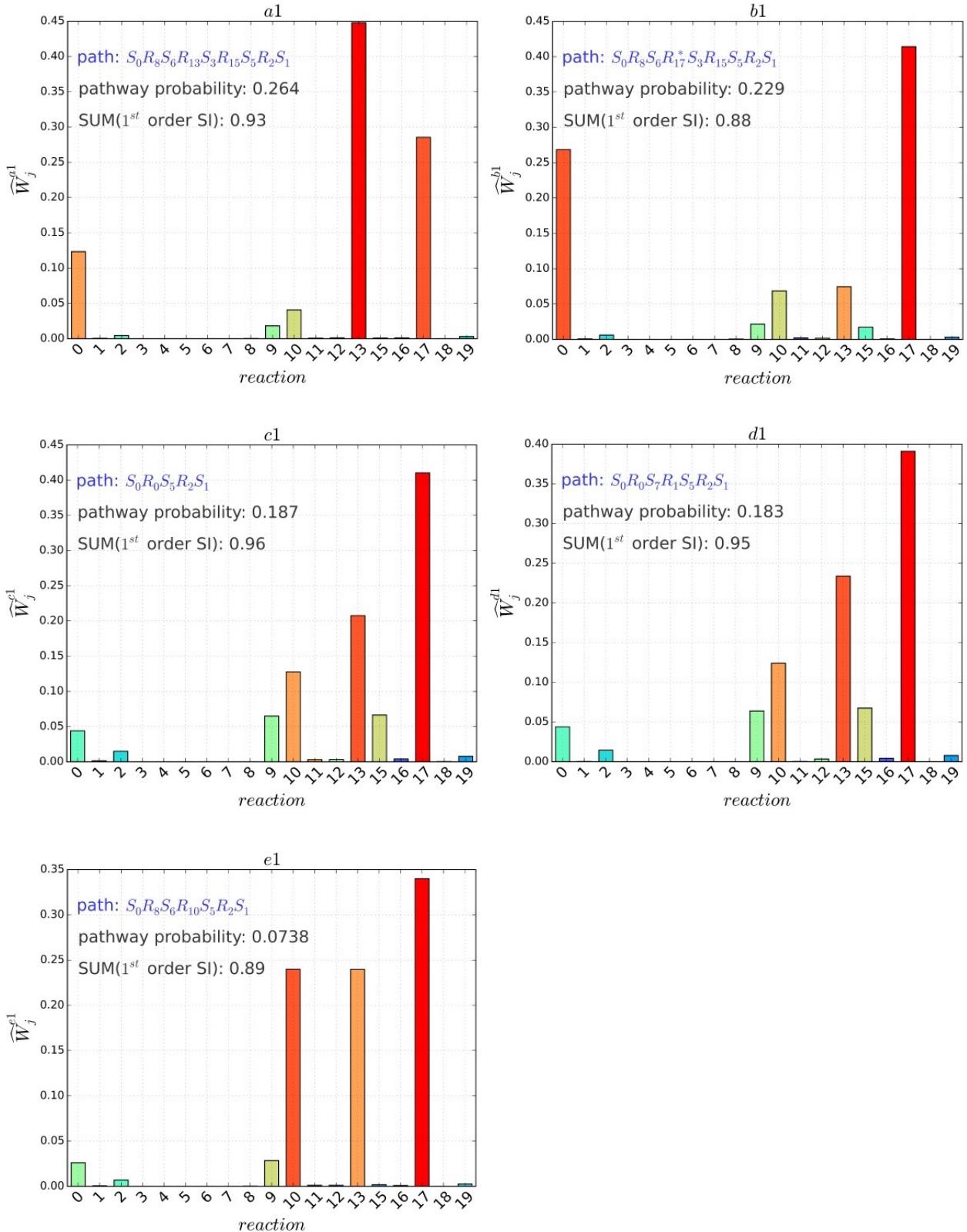


Figure 4.13: Sensitivity indices versus reaction index for the five most important paths leading to the H₂O target evaluated at time $0.8\tau_{ign}$. The relevant reactions for each path are reproduced on the figure panels. Also shown in each panel is the pathway probability and the sum of first-order indices obtained by the HDMR expansion of p_i .

of H_2O target sensitivity, W_j , in Fig. 4.4 are also, on average, the most sensitive reactions for the paths leading to the target. However, we notice that the detailed distributions of \widehat{W}_j^1 values differ significantly from pathway to pathway. For brevity we use the pathway superscript i going from 1–5 for pathways a1–e1. The most probable path, a1, shows strongest sensitivity to R13 and R17 with a modest contribution from R0, i.e., $\widehat{W}_{13}^1 > \widehat{W}_{17}^1 > \widehat{W}_0^1$. For the second most important path, b1, the contribution from R17 is now clearly the largest, with R2 second and R13 and R10 also present, i.e., $\widehat{W}_{17}^2 > \widehat{W}_2^2 > \widehat{W}_{13}^2 \approx \widehat{W}_{10}^2$. Paths c1 and d1 give a nearly identical spectrum because, as discussed in section 4.4, they are closely related physically. For those two paths, we see that $\widehat{W}_{17}^3 > \widehat{W}_{13}^3 > \widehat{W}_{10}^3$. For e1, we have $\widehat{W}_{17}^5 > \widehat{W}_{13}^5 \approx \widehat{W}_{10}^5$. We can understand the results of Fig. 4.13 from the detailed character of the chemical pathways. The pathway a1 is most sensitive to R13, which is rate limiting along this pathway. For b1, on the contrary, the reaction R17* is most sensitive because it is rate limiting for this pathway. We note that both \widehat{W}_{17}^3 and \widehat{W}_{13}^3 are appreciable in both the a1 and b1 spectra due to the pathway switching mechanism. Clearly, by changing the ratio k_{13}/k_{17}^* , we will alter whether the O atom follows the a1 path or the b1 path. Indeed, inspection of the graph in Fig. 4.8 shows that a1 and b1 are identical except for the step where the HO₂ radical decays via $\text{HO}_2(+\text{HO}_2) \xrightarrow{\text{R}_{13}} \text{H}_2\text{O}_2$ for a1 and via $\text{HO}_2(+\text{H}_2) \xrightarrow{\text{R}_{17}^*} \text{H}_2\text{O}_2$ for b1. It is interesting to notice that R0, $\text{H} + \text{O}_2 \rightarrow \text{OH} + \text{O}$, is quite important for a1 and b1 even though this step does not enter into either path. More understandably, this reaction is also sensitive in c1 and d1, where it is the initial step. For a1 and b1, the graphs in Figs. 4.5 and 4.9 make it is easy to see that R0 is in competition with R8, $\text{H} + \text{O}_2 + \text{M} \rightarrow \text{HO}_2 + \text{M}$, which is the initial step for a1 and b1. Thus, again the sensitivity to R0 is a case of pathway switching between (c1, d1) and (a1, b1). The pathway e1 is interesting in that the reaction R10, $\text{HO}_2 + \text{H} \rightarrow \text{OH} + \text{OH}$, emerged as the key step. This is consistent with the observation that the sensitivity index \widehat{W}_{10}^5 for this reaction is seen to be quite large. That R13 and R17* are also sensitive reactions again reflects the competition from those reactions for the HO₂ radical and thus pathway switching.

In addition to the sensitivity indices of the pathway probabilities, we have also computed the correlations between the pathways that lead to the H₂O target. In the interest of brevity, we focus

on the correlation that develops within just two pairs of pathways, (a1, b1) and (a1, c1). Using eqn. (4.12), we see in Fig. 4.14 that the largest interactions for the (a1, b1) pair, $\widehat{W}_j^{a1,b1}$, are due to reaction R17*, a negative correlation, and reactions R0 and R13, positive correlations. For the (a1, c1) pair, $\widehat{W}_{17}^{a1,c1}$, is again the largest, with a positive $\widehat{W}_{13}^{a1,c1}$, the next largest but $\widehat{W}_0^{a1,c1}$ now yielding a negative correlation. The explanation for these results can be provided by examining the first-order component functions, $F_j^i(k_j)$, for paths $i = a1, b1, c1$. In Fig. 4.15 we plot $F_j^i(k_j)$ for a1, b1, and c1 over the uncertainty range for each rate coefficient. Note that $F_j^i(k_j) = 0$ because the mean probability is already taken out. The negative correlation for reaction R17* arises because $F_{17}^{a1}(k_{17}^*)$ decreases with increasing k_{17}^* whereas $F_{17}^{b1}(k_{17}^*)$ and $F_{17}^{c1}(k_{17}^*)$ both increase. This is quite understandable because for path a1 reactions R13 and R17* are in competition to consume the HO₂ radical; thus an increase in k_{17}^* decreases the flux along path a1 for which the R13 step is rate limiting. On the contrary, pathway b1 includes the rate limiting R17* step and so increasing k_{17}^* promotes this pathway. The pathway probability for c1 ($i = 2$) increases with k_{17}^* because reaction R17* produces an H-radical which then promotes reaction R0, H + O₂ → OH + O. Because reaction R13 leads to the formation of H₂O₂, which almost always decays to OH + OH, increasing k₁₃ also increases the probability for c1 by promoting reaction R2, OH + H₂ → H₂O + H. Finally, $F_0^2(k_0)$ increases with increasing k₀ because R0 is the initiation step of this pathway; on the contrary, $F_0^0(k_0)$ and $F_0^1(k_0)$ both decrease with increasing k₀ because R0 is in competition with R8, which initiates pathways a1 and b1. In a similar fashion we could also understand the correlations that develop between other pathways in the model.

One of the most interesting aspects of the H₂–O₂ system is the crossover behavior where the R17* is the main source of H₂O₂ at early times whereas R13 becomes largest at long times. This mechanistic crossover is apparent in the overall reaction rates, Fig. 4.2, which shows that the rate for R13 crosses and then exceeds that for R17* at about $t = 0.5\tau_{ign}$. In terms of pathways, H₂O₂ production is dominated by just two, a3 and b3, which account for over 99% of the total. These two paths differ only in the step that produces H₂O₂ from HO₂, reactions R13 and R17*, respectively. Thus, in Fig. 4.6, we saw that the pathway probability for a3 likewise crosses and exceeds that for

Figure 4.14: Correlation between pathways (a1, b1) in the upper panel and (a1, c1) in the lower panel for the H₂O target evaluated at $t = 0.8\tau_{ign}$. The correlation has been decomposed into the contributions from specific reactions via eqn. 4.12.

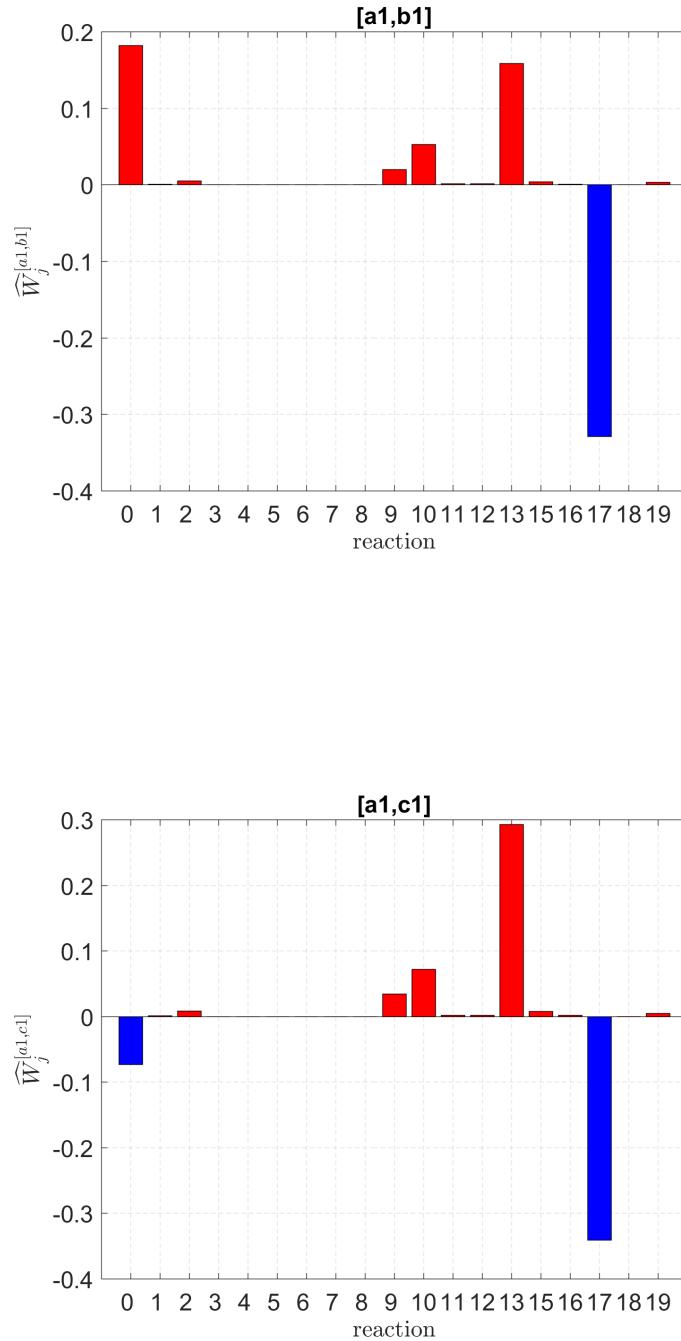
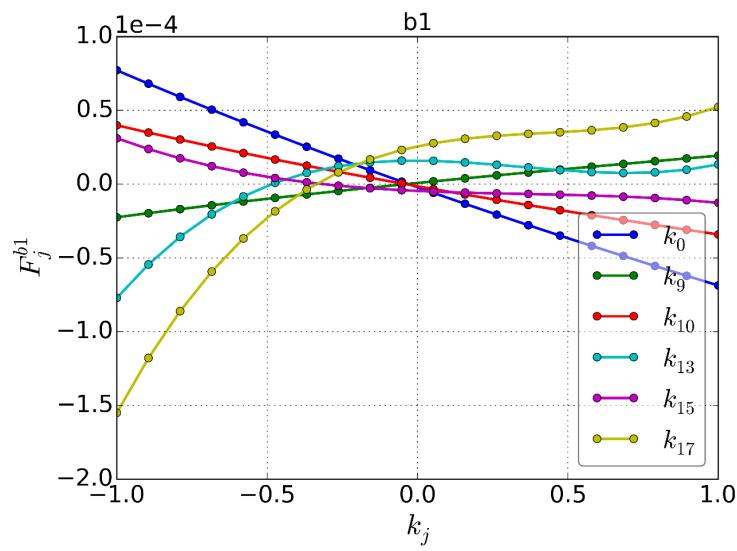
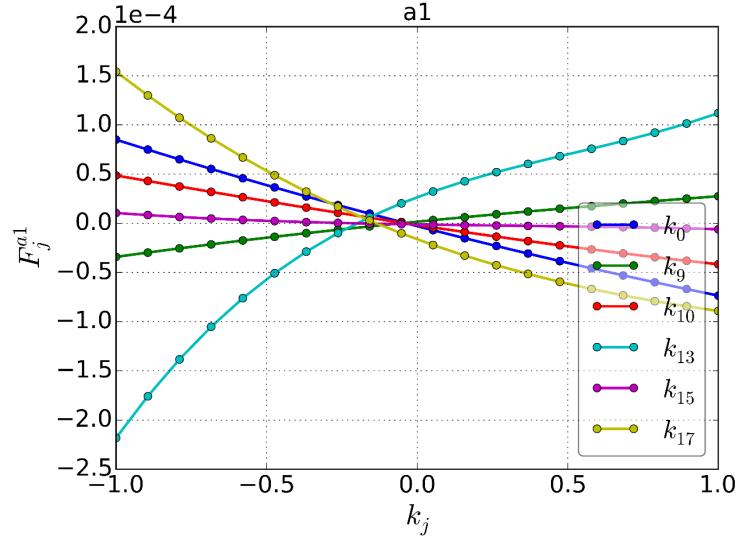
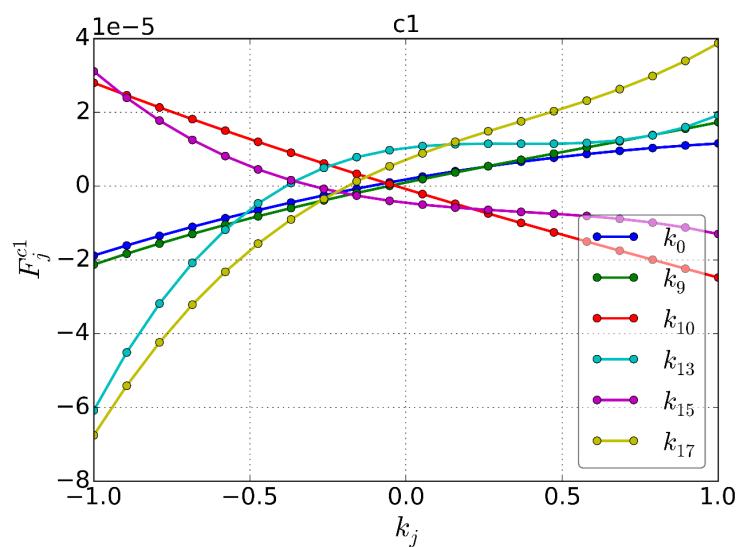


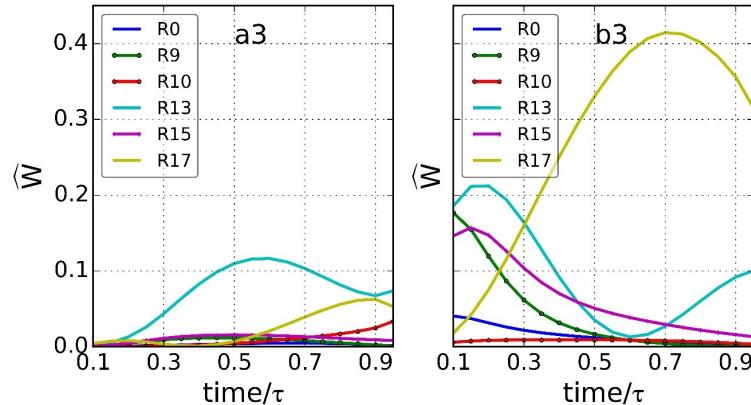
Figure 4.15: First-order component function for the HDMR expansion of the pathway probabilities a_1 , b_1 , and c_1 for the H_2O target evaluated at $t = 0.8\tau_{ign}$.





b3 at roughly the same point in time as the rates. It is instructive to study the pathway sensitivities versus time for this crossover regime. In Fig. 4.16 we see the sensitivities for paths a3 and b3 are dominated by R13 and R17*, respectively, but in a quite asymmetrical way. As we might expect, pathway a3 is more sensitive to R13 than to R17* and b3 is more sensitive to R17* than R13. However, the degree of sensitivity and the shapes of the profiles versus time are dissimilar, which reflects the differing functional dependence of p_{a3} and p_{b3} on the rate coefficients. The a3 pathway shows a growth in both the R13 and R17* sensitivity near the $t = 0.5\tau_{ign}$ crossover point. The b3 pathway shows a significant growth in R17* sensitivity early in the ignition process and a more complicated oscillating sensitivity to R13 versus time.

Figure 4.16: Sensitivity indices for the two dominant paths leading to H_2O_2 as a function of time.



Finally, we note the unexpected results obtained for the HO_2 target. The sensitivity indices for this target shown in Fig. 4.4 are generally similar to that for other species, viz. strong contributions from the R13, R17*, R15, and R9. Unlike the other species, however, HO_2 formation is dominated by a single one-step chemical pathway, $\text{O}_2 + \text{H} + \text{M} \rightarrow \text{HO}_2 + \text{M}$, described by R8. Hence, it is at first surprising that R8 does not significantly contribute to the sensitivity spectrum at any time. Instead, the W_j are largest for reactions involved in the decay of the HO_2 species, R13 and R17*, and the secondary decay of H_2O_2 , i.e., R15. The lack of R8 dependence is partly due to use of O atom following. Clearly, HO_2 formation is limited not by oxygen but instead by the formation of

H-radicals. Indeed, R8 is found to be much faster than R17* ($\text{HO}_2 + \text{H}_2 \rightarrow \text{H} + \text{H}_2\text{O}_2$) and R9 ($\text{H}_2 + \text{O}_2 \rightarrow \text{HO}_2 + \text{H}$), which are both sources of H atoms and both appear in the sensitivity spectrum.

4.6 Conclusions

The sum over histories approach offers a uniquely quantitative method for investigating the contribution of distinct and competing chemical pathways to the behavior of large mechanisms. The sum over histories method allows the concentration of species to be represented in terms of a sum over the probabilities for chemical pathways that generate that species. This method differs from other techniques for reaction path analysis that represent the pathways at an instant of time using a flux snapshot. Instead, our method permits the time evolving chemistry of the system to be quantitatively incorporated into the probability for the chemical pathways. Hence, the dominant chemical pathways followed by chemical moieties are allowed to change as time progresses. In the present work we have used the sum over histories approach to analyze the sensitivity of speciation in the H_2-O_2 combustion system. Thus, the formation of any target species is described in terms of a sequence of reactions that carry a tagged atom from reactants to the target. Using O atom following, it was found that a small number of chemical pathways could account for the formation of the products and intermediates prior to the ignition time. The observed sensitivities for species targets to the values of the rate coefficients could, for the most part, be traced to key steps along the chemical pathways. In particular, it was observed that the chemistry of HO_2 radicals and the subsequent formation of H_2O_2 were essential steps along many dominate chemical pathways. It was found that high sensitivity was often associated with rate limiting steps or pathway switching phenomena involving HO_2 self-reaction, R13, and $\text{HO}_2 + \text{H}_2$, R17*.

As a final point, we note that there remains an element of choice when applying the sum over histories representation, specifically in the selection of the chemical moiety to be followed. In principle, we may select any atom in the target molecule as our followed atom to implement the algorithm. However, the efficiency of the method, and the insight it affords, does depend on this

choice. For the present H₂–O₂ combustion problem, e.g., we found that O atom following procedure was more rapidly convergent and more useful for interpretation than the H atom following approach. However, the sensitivity of HO₂ species formation to the rate coefficients could not be well explained in this representation. Instead, the formation of HO₂ is largely controlled by H atom production required for the association reaction H + O₂ + M → HO₂ + M. Hence, for this species it is advisable to switch to the H atom following scheme to understand the results of the sensitivity analysis.

Chapter 5

Simulating Chemical Kinetics Without Differential Equations

Most of the contents of this chapter are reprinted, with permission, from

- [5] Bai, S., & Skodje, R. T. The Sum Over Histories Representation for Chemical Kinetics: A Quantitative Theory Based on Chemical Pathways. *International Reviews in Physical Chemistry*, 35(4), 539-567. **2016**
- [6] Bai, S., & Skodje, R. T. Simulating Chemical Kinetics Without Differential Equations: A Quantitative Theory Based on Chemical Pathways. *The Journal of Physical Chemistry Letters*, 8(16), 3826-3833. **2017**

5.1 Introduction

Simulation of chemical reaction networks using mass-action chemical kinetics provides an invaluable tool for uncovering chemical mechanisms and predicting time-dependent chemical behavior as functions of controllable parameters such as temperature, pressure, and initial concentrations. As chemical mechanisms grow large, such as for hydrocarbon combustion, the resulting kinetic behavior can become quite difficult to physically decipher and, in some cases, the simulations themselves can bog down due to the proliferation of species and the need for small time steps owing to the existence of multiple time scales. Thus, considerable effort has been devoted in the disciplines of chemistry, biochemistry, physics and engineering to interpreting and simplifying kinetic behavior in complex networks.

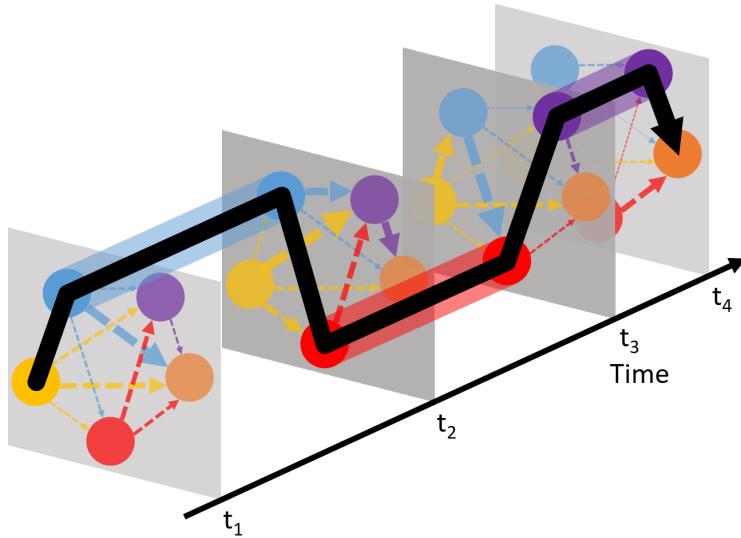
The usual computational approach to mass-action kinetics involves constructing and solving a set of differential equations in the continuous concentration variables $X_i(t)$ that model the local rates of formation and destruction of each of the N chemical species, S_i . (Underpinning kinetics equations are defined in eqns. 2.2 and 2.3.) For the M elementary reaction steps, defined in eqn. 2.1, the explicit rate laws $R_j(\mathbf{X})$ are provided by an external experimental or theoretical determination and $v_{i,j}$ are known constants.[54, 91, 76] This representation is properly regarded as a mean field theory that becomes valid in the thermodynamic limit. For systems involving very low concentrations or strong local environment effects, more intensive simulations may be used that employ stochastic or Monte Carlo algorithms that may account for the influence of fluctuations.[69, 31, 45]

In this chapter, we propose a new method to simulate the kinetics of general nonlinear mechanisms that builds on, and fundamentally extends, our earlier work.[53, 7, 4, 5] In contrast to the local rate equation approach to kinetics, we propose a self-contained global method based on chemical pathways that tracks molecules and fragments as they move throughout the network. A pathway follows a chemical moiety, at a molecular level, as it hops from species to species under the influence of elementary chemical steps. A convenient approach to the unique specification of a chemical pathway utilizes an atom-following algorithm. If, e.g., the carbon-atom in a CH_4 reagent is initially tagged, $\text{CH}_4(+\text{OH}) \xrightarrow{\text{R}_1} \text{CH}_3(+\text{OH}) \xrightarrow{\text{R}_2} \text{CH}_2(+\text{O}) \xrightarrow{\text{R}_3} \text{CH}_2\text{O}$ defines a three-step pathway leading from CH_4 to the product CH_2O . Each reaction is modeled as a random event governed by transition probabilities obtained from the elementary rate laws. By constructing paths for a complete set of atoms in the initial state at a time t_0 , it is possible to predict the subsequent values of any kinetic observable given the pathway probabilities (see below). Hence, the method is global in the sense that it replaces local rate equations based on instantaneous rates for individual species with expressions that span the entire network over finite times using multistep chemical pathways.

The use of chemical paths to analyze the behavior of chemical networks has a long history. A most convenient way to visualize and enumerate chemical paths involves the use of graph theory.[71, 8, 27] Graph theoretic analysis of chemical mechanisms was initiated by the early work of Christiansen[15]

and also King and Altman[50] who exploited the analogy between chemical and electrical circuit theory. Since then, there have been numerous treatments of chemical networks, especially for biochemical systems, that have demonstrated the utility of graph theory.[75, 43, 26, 41, 59, 24, 66] Closely related work on the solution of master equations via network analysis has also been explored using graph theory.[82, 2] Much of work on kinetic graphs has focused on the developing expressions for the steady state rates for first-order or pseudo-first-order chemical mechanisms. Typically, the vertices of those graphs comprise a subset of the species, such as biochemical intermediates, and the edges are reaction steps. Quantitative treatment of the kinetics are possible using edge weights obtained from the steady state fluxes. In our work, we have developed a quantitative pathway

Figure 5.1: A schematic representation of a time-dependent pathway followed by a tagged atom through a chemical network of five species. The reactions occur instantaneously at times t_1, t_2, t_3 , and t_4 , which are randomly selected and averaged to obtain the pathway probability.



analysis, called the Sum Over Histories Representation or SOHR, which goes beyond steady state conditions and fully embraces the time-dependent and nonlinear behavior of general mechanisms. As emphasized in the schematic diagram of Fig. 5.1, the SOHR method employs time-resolved chemical pathways in which reaction occur at specific randomly selected times. This treatment employs dynamical graph theory[38] in which the edge weights are time-dependent. This is quite

important for quantitative modeling since static snapshots of the reactive flow can be very deceptive for interpreting and modeling time-dependent chemical flux. Our approach draws inspiration from Feynmans path integral representation of quantum mechanics.[25] We have previously used the SOHR method to interpret the workings of several kinetic networks, including those involved with surface catalysis and combustion. While we could readily identify and quantify the important chemical paths that guided the kinetics, the results were interpretative rather than predictive. That is, for general nonlinear problems, the determination of the pathway probabilities requires a knowledge of the time-dependence of the pseudo-first order rate coefficients which, in turn, requires an a priori knowledge of the concentrations versus time. For interpretive analysis, the nonlinear kinetics was first solved using conventional kinetics to obtain a reference trajectory, $\mathbf{X}(t)$, and then the kinetics was decomposed into contributing pathways. In this work, we overcome this limitation and formulate the SOHR as a predictive theory that does not require a solution to eqn. 2.2. In this way, the need to solve the usual differential equations of mass action kinetics is avoided, and the pathway probabilities can be solved in a self-consistent manner.

The key to our formulation is the use of an iterative solution technique to obtain the pathway probabilities. Here The iterative SOHR method is presented. Given a set of chemical pathways, the pathway probabilities and the species concentrations are obtained simultaneously by the iteration. A strategy to identify the most important chemical pathways is given. The new predictive SOHR method is numerically tested on the Lotka Volterra model and H₂–O₂ combustion system. The results presented demonstrate that the method is accurate and quickly convergent. We conclude with a discussion of the possible advantages of the new method.

5.2 The Predictive SOHR Method

For a linear kinetic network, the SOHR formulation is automatically predictive since the pathway probabilities are known analytically via eqn. 3.1. Thus, if a sufficient number of pathways are enumerated, the sums in eqn. 2.18 are guaranteed to converge to the correct concentration functions. However, we can immediately see a difficulty in applying this formalism to general

nonlinear kinetic networks. Namely, the quantities $A_{i,m}(t)$ are no longer constants and pathway probabilities seen to require the solution to the conventional kinetics eqn. 2.2 as input into eqn. 2.12. In a departure from our previous work, we now show how this requirement can be avoided using a new methodology.

We shall assume that a finite number of distinct pathways, J , is adequate to accurately describe the evolution of the N species concentrations $X_i(t)$ over a time interval $t \in (0, T)$. In the next paragraph, we discuss how those pathways may be identified. If the concentrations $X(t) = (X_1(t), \dots, X_N(t))$, which solve eqn. 2.2 with initial conditions $\mathbf{X}(t = t_0) = \mathbf{X}_0$ are known, by some method, then the J probability functions $\bar{\mathbf{P}}(t_0, t) = (P_1(t_0, t), \dots, P_J(t_0, t))$ can be calculated using eqn. 2.15 or 2.16. We can express this fact using the functional relation

$$\bar{\mathbf{P}}(t_0, t) = \bar{\mathbf{F}}[\mathbf{X}] \quad (5.1)$$

The concentrations are likewise represented as linear combinations of pathway probabilities (eqn. 2.18) that can be symbolically represented as

$$\mathbf{X}(t) = \bar{\mathbf{G}} \cdot \bar{\mathbf{P}}(t_0, t) \quad (5.2)$$

Here $\bar{\mathbf{P}}$ and $\bar{\mathbf{F}}$ are J -dimensional vectors, \mathbf{X} is a N -dimensional vector, and $\bar{\mathbf{G}}$ is a $N \times J$ matrix. In order to solve the kinetics problem, we need to simultaneously solve eqns. 5.1 and 5.2 to obtain the self-consistent solution $[\mathbf{X}(t), \bar{\mathbf{P}}(t_0, t)]$. A standard approach to a problem such as this is to employ functional iteration.[78] Thus, one makes an initial guess for the concentrations that is then used to obtain the pathway probabilities. These probabilities are then used to recompute the concentrations. If m is the iteration number, we have

$$\bar{\mathbf{P}}^{m+1}(t_0, t) = \bar{\mathbf{F}}[\mathbf{X}^m] \quad (5.3)$$

$$\mathbf{X}^{m+1}(t) = \bar{\mathbf{G}} \cdot \bar{\mathbf{P}}^m(t_0, t) \quad (5.4)$$

The iteration is continued until convergence is obtained. We view the convergence properties and stability of this functional iteration as an empirical matter. In our numerical applications, we have

found that the iteration is stable, given a sensible initial guess, and the iteration converges within 10 steps to a level consistent with the numerical error of the MC integration of $\bar{\mathbf{P}}(t_0, t)$.

5.3 Pathway Generation

An important ingredient in the SOHR method is the choice of the "important" chemical pathways needed to treat a given kinetic network. Since we have discussed this procedure elsewhere[5], and since this issue is not the primary focus of this work, we shall only address this point briefly. A general chemical network will typically have an infinite number of pathways on a kinetic graph that can deliver a tagged atom from an initial species S_r to a final species S_i ; however, the probabilistic weights fall off for long path lengths, and the pathway expansion can be truncated. For the simple special case of acyclic graphs, which do not satisfy micro-reversibility and form trees, there will actually be only a finite number of possible paths. We have previously outlined three strategies to generate the paths for the general problem:

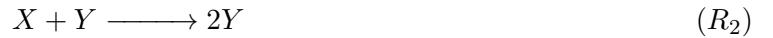
- (1) Matrix enumeration of the paths using the adjacency matrix, as previously introduced in Section 2.5.
- (2) Monte Carlo path sampling using a small stochastic simulation of the kinetics, as previously discussed in Section 2.6.
- (3) Search algorithms to find optimal paths on weighted graphs.[60, 21, 10, 39, 22, 38]

To illustrate and explore the performance of the iterative SOHR method, we have chosen two test problems: the model Lotka-Volterra (LV) system and the high-temperature combustion problem H₂-O₂. The LV is a simple three-reaction model that has often been employed in the literature to explore new ideas in nonlinear kinetics and dynamical systems. This application allows us to investigate the workings of the SOHR method without the complications encountered in "real" problems. In particular, there are only a finite number of chemical pathways in the LV mechanism due to the absence of reverse reactions. The H₂-O₂ combustion system is, in contrast, a

true physical problem of some complexity with 21 reversible reactions and empirically determined temperature/pressure dependent rate coefficients. (Detailed mechanism is described in Appendix A) There are an infinite number of potential chemical pathways that must be efficiently pruned to yield a manageable numerical problem. Success with this system demonstrates that the predictive iterative SOHR method has the potential of practical utility to physical systems.

5.4 The Lotka-Volterra System

The LV system is a nonlinear dynamical system that has been used to study chemical kinetics[64] as well other problems such as predator-prey population ecology[97] and macroeconomics.[23] In kinetics, the autocatalytic LV system corresponds to the three reactions



This yields the simple kinetics equations,

$$\frac{dA}{dt} = -k_1 A \cdot X \quad (5.5)$$

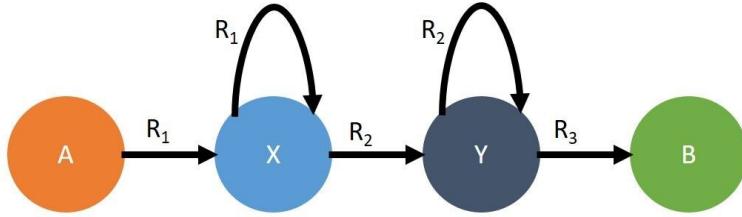
$$\frac{dX}{dt} = k_1 A \cdot X - k_2 X \cdot Y \quad (5.6)$$

$$\frac{dY}{dt} = k_2 X \cdot Y - k_3 Y \quad (5.7)$$

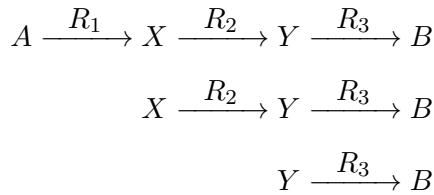
$$\frac{dB}{dt} = k_3 Y \quad (5.8)$$

where A , B , X , and Y stand for the time-dependent concentrations, and k_1 , k_2 , and k_3 are the rate coefficients in reactions R_1 , R_2 and R_3 . When we hold fixed the value of A , the intermediates X and Y can undergo a variety of interesting nonlinear effects such as persistent chemical oscillation. For the autonomous LV problem, where A varies in time, the kinetics approaches a simple stable equilibrium state. Since reactions R_1 , R_2 and R_3 conserve the number of molecules, we can hypothesize a fictitious atom at the core of each species that we choose to follow. The directed

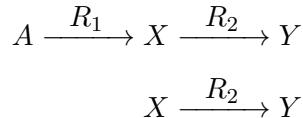
Figure 5.2: Kinetic graph for the Lotka-Volterra model.



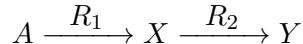
chemical graph, shown in Fig. 5.2, illustrates the simple connectivity of the chemical mechanism. We list in Table 5.1 the parameters used in the calculation and the possible reaction routes followed by a fictitious atom. In the pathway representation, the kinetics is represented by 10 independent chemical paths: three terminating on B



two additional paths terminating on Y



and one path terminating on X



If we include the trivial zero-step paths, where A, X, Y, and B do not react, then there are only 10 pathways that describe any kinetic observable.

Given the complete set of 10 chemical pathways, we can numerically test the predictive SOHR

method for the LV system. The concentration for any species Z_i is given by

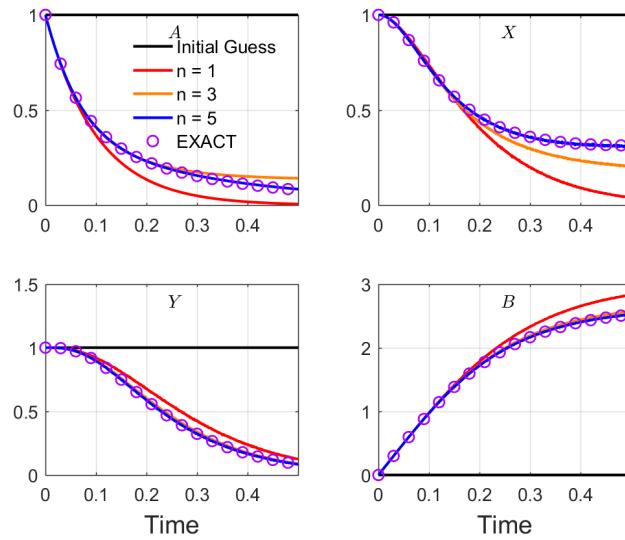
$$Z_i(t) = \sum_{j=1}^{10} c_j P_j(0, t) W_j(t=0)$$

Here, j labels the paths that terminate on species Z_j , $P_i(0, t)$ is the probability from eqn. 2.16, $W_j(t=0)$ is the initial concentration of the first species along the pathway containing the "fictitious atom", and the stoichiometric coefficients c_j are all equal to 1. Using eqns. 5.3 and 5.4, we can obtain a set of iterates, $Z_i^n(t)$ based on an initial guess $Z_i^0(t)$. To illustrate the robustness of the algorithm, we choose the trivial initial guess of constant concentrations, $Z_i^0(t) = Z_i(t=0)$. The MC integration of the pathway probabilities (eqn. 2.16) is carried out using $L = 10000$ random strings, which is expected to yield the concentrations to about 1% accuracy. The concentration profiles are evaluated using eqns. 2.18 and 2.16 on a uniform grid of times between $t = 0$ and $t = 0.5$. In Fig. 5.3, we show the convergence of the iteration for $n = 0, 1, 3$, and 5 along with the "exact" result obtained by integrating eqns. 5.5, 5.6, 5.7 and 5.8 using an ODE solver. We see that the iterative SOHR method converges quite quickly to the exact result. Obviously, the nonlinear effects are strong in the LV system, and so the rapid convergence of the results is a positive omen for the iterative computational strategy.

5.5 H₂-O₂ Combustion

Hydrogen combustion kinetics presents a more computationally challenging problem for the SOHR method since it exhibits strong nonlinearity, stiffness, and thermokinetic behavior. The kinetic behavior of the H₂-O₂ system is nevertheless well understood.[100] To emphasize the "traditional" characteristics of this combustion application, we carried out the SOHR simulations using standard CHEMKIN subroutines for the mechanistic and thermokinetic aspects of the problem. The mechanism used here consists of 8 species and 21 reversible reactions (see Table 4.1, 4.2 and A.1) and is carried out at high temperature T = 2000 K and pressure P = 10 bar using an initial stoichiometric mixture of H₂ and O₂. Under these conditions, the system approaches equilibrium within about 1μs. There are an infinite number of paths; thus, an important aspect of the method

Figure 5.3: Convergence of the iteration for the predictive SOHR method in the LV system. The initial guess for the concentrations are the constants $A(t_0) = X(t_0) = Y(t_0) = 1$ and $B(t_0) = 0$. The $n = 1, 3$, and 5 iterations are shown along with the exact results. It is seen that by the fifth iteration, the SOHR method has effectively converged to the true result.

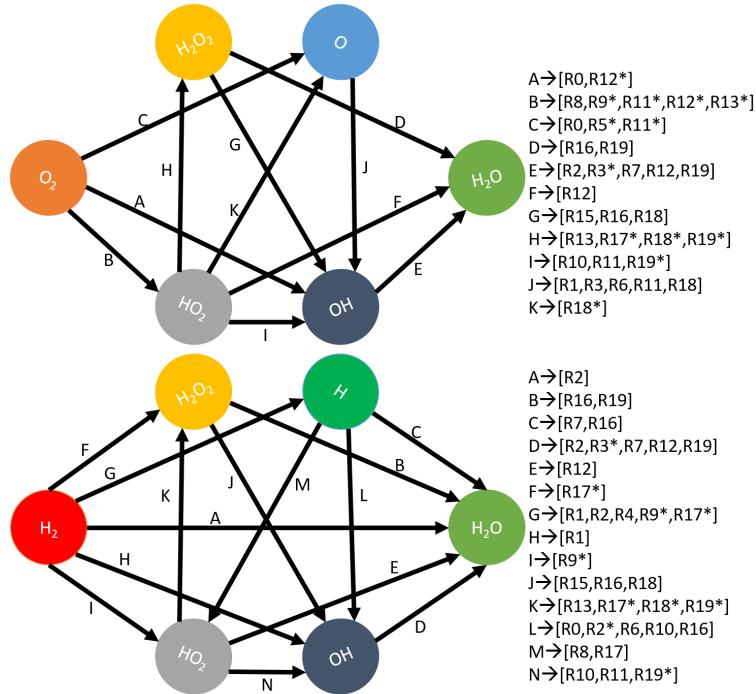


is the selection and ordering of these reaction routes. If, using chemical intuition, we could guess the important paths, they could then be directly inserted into the SOHR algorithm. Here, to demonstrate the generality of the method, we apply an automated pathway selection procedure that assumes nothing physical about the H₂-O₂ combustion problem. In previous work, we have applied the interpretive SOHR method to identify the key pathways for the H₂-O₂ system. Under the conditions studied there, T = 1000 K and P = 5 bar, it was found empirically that many of the key reaction paths passed through the H₂O₂ and HO₂ intermediates. For the higher temperature conditions considered here, we have identified a different set of paths.

We employ a pathway generation method that is a hybrid of the matrix and Monte Carlo procedures, i.e., a set of "short" pathways found by the matrix method are augmented with "long" pathways found by the MC-method. First, using the matrix method, all potential pathways of up to length n_{max} are identified using powers of the sparse **SR**-matrix, eqn. 2.23. The kinetic graphs for the O atom and H atom following pathways are shown in Fig. 5.4. These are multigraphs since vertices generally have more than one edge connecting them. For the initial concentration guess $\mathbf{X}_0(t)$ in the iteration scheme, the probabilities for those pathways are quickly estimated using a low-level MC-evaluation of eqn. 2.16, i.e., using a small value of L . The top J paths obtained by this screening process are used for a higher-level estimation of the probabilities, i.e., a larger value of L , which are then used in eqns. 5.3 and 5.4 to obtain the first iterative concentrations, $\mathbf{X}_1(t)$. The $\mathbf{X}_1(t)$ is then used in an updated screening process to obtain a new (possibly different) set of J paths for the second iteration. A small number of MC pathway sampling stochastic trajectories are propagated to test whether paths longer than n_{max} contribute. If so, they are added to the list of contributing pathways used in the iteration. The process continues until convergence is achieved. Since path lengths tend to grow longer with propagation time, we have found it efficient to break up the total propagation time into a set of shorter sub-intervals (or sectors) for which the iteration process is successively converged. In this way, the chemical pathways during each phase the chemical evolution are allowed to separately optimize.

In Fig. 5.5, we illustrate the convergence of the sector-by-sector iteration scheme for several of

Figure 5.4: Directed kinetic graphs for the forward reactions in the hydrogen combustion system. In the upper panel, we show the graph for O atoms. In the lower panel, we show the graph for H atoms. The reactions $R_0 - R_{20}$ correspond to the elementary reactions presented in Table 4.2. For brevity, single arrows labeled $A - N$ are used to represent the multiple edges listed in the caption. The * notation is used to indicate a reverse reaction in the nominal left-to-right direction of the reactions in Table 4.2. All forward and backward reactions are included in the calculations.



the concentration functions. The initial guesses again are taken to be the trivial constant functions, $\mathbf{X}^0(t) = \mathbf{X}^0(t_0)$ where $\mathbf{X}^0(t_0)$ is the initial concentration at the beginning of each sector. For a total propagation time of $T = 0.2 \mu\text{s}$, 20 sectors of $0.01 \mu\text{s}$ were chosen. The maximum path length was taken to be $n_{max} = 5$ and the number of paths in the iteration was selected as $J = 20$. It is seen that within seven iterations, the concentration profiles of all species were converged and stable. The selection of convergence parameters requires some experimentation, but no strenuous efforts were required to obtain convergence. The full concentration profiles obtained by combining the sector-bysector results are shown in Fig. 5.6 along with a reference result obtained by conventional integration of the kinetics equations using an implicit integration algorithm. It is seen that predictive SOHR results converge globally to exact results to within the tolerance set by the MC-integration scheme. We note that the selection and ordering of chemical pathways changes considerably over the propagation time. This reflects, e.g., the differences in the chemistry of the initiation and steady state regimes of the kinetics. Since our objective in this work is the development of the iteration scheme, and not the interpretation of the chemical pathways, we shall not delve further into an analysis of these interesting changes.

5.6 Utility of the Method

The success of the iterative algorithm demonstrates that it is possible to model the chemical kinetics of a realistic mechanism without the need to solve the conventional system of coupled differential equations. This establishes that the SOHR method provides a consistent alternative formalism for modeling the evolution of chemical systems. The question then arises, under what circumstances will the SOHR method be an advantageous approach for solving a kinetic problem? In our previous work, we demonstrated that SOHR methodology was valuable to interpret the behavior of complex mechanisms. Indeed, the chemical pathways identified by SOHR provide a physical picture of how the chemistry occurs. These dynamical pathways can be quite different from the static snapshots provided by earlier steady state treatments. Under certain conditions, we suggest that the predictive SOHR method may be useful approach to carry out the numerical

Figure 5.5: Convergence of the SOHR iteration for the H₂-O₂ system. The calculation is carried out using the sector-by-sector method, so the iteration is converged for each sector of $0.01 \mu\text{s}$ before moving on to the next sector to repeat the process. The initial guess for the iteration in each new sector is the constant values of the concentrations at the end of the previous sector. The conditions chosen for the simulation are T = 2000 K, P = 10 bar, and an initially stoichiometric mixture, [H₂(t₀)]/[O₂(t₀)] = 2. The results show that the iteration converges accurately to the exact solution to eqn. 2.2 within about five iterations.

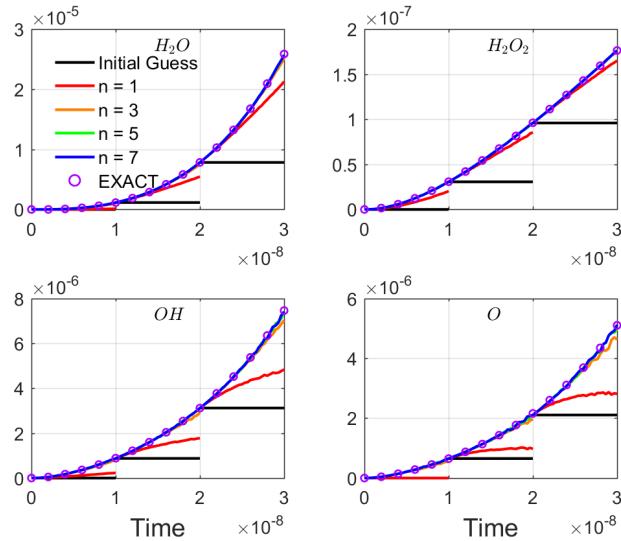
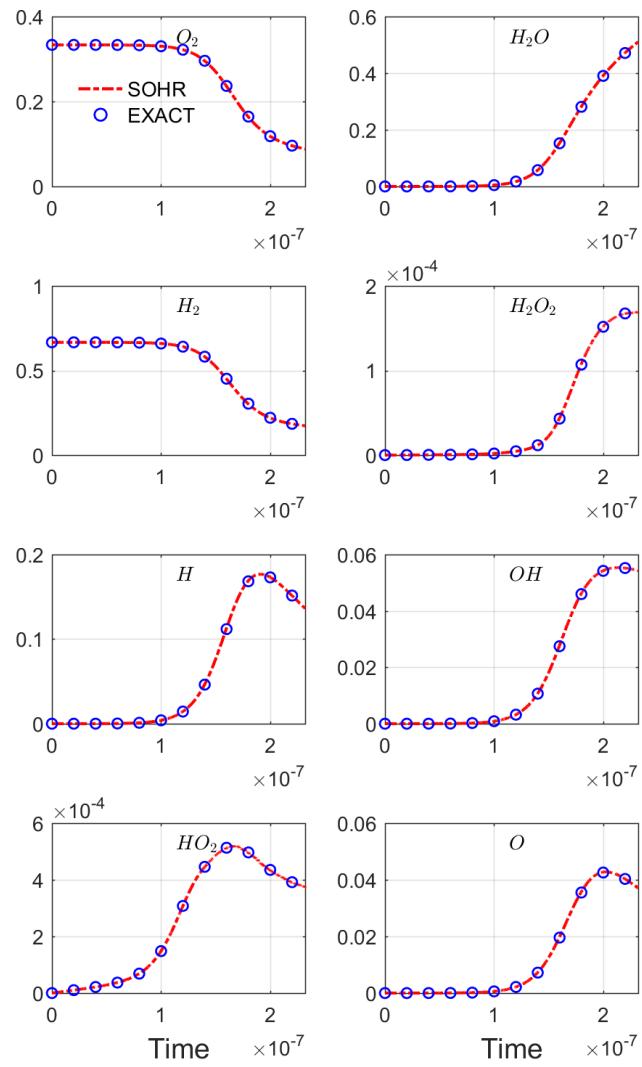


Figure 5.6: A comparison of the predictive SOHR method with the conventional ODE kinetic modeling for the H₂-O₂ combustion problem. Using approximately 20-25 chemical pathways automatically selected by the algorithm, and combining the sector-by-sector trajectory fragments obtained with five iterations, the SOHR results are seen to converge to the exact kinetics.



simulations. For example, there are cases where the number of species variables becomes extremely large such as for the conformation kinetics of large polymers. The solution of eqn. 2.2 can then become difficult or impossible. However, should it be feasible to identify a much smaller number of pathways, by chemical intuition or other means, the SOHR representation can then become a more efficient approach. Another potential application is to systems exhibiting multiple time scales that lead to extremely stiff ODE's. While conventional implicit integration methods may require very small time-steps to stabilize extremely stiff systems, thus greatly lengthening computation time, the pathway representation has different convergence properties. The importance sampling used the MC evaluation of the pathway probabilities uniformly samples the survival probabilities, \mathcal{P}_i in, roughly, the interval $(0, 1)$ for each reaction step along the pathway. Thus, even though the species lifetimes may be vastly different, each step is integrated on its own appropriate scale. A related issue is use of parallelization in the integration of single trajectories. In fact, we used parallel coding to carry out the H₂-O₂ SOHR simulation where individual pathways were distributed to different processors. Under some circumstances, which we are investigating, there can be parallelism across the system. By contrast, conventional ODE solvers can be difficult to parallelize for individual trajectories since they involve sequential time stepping.[85] Finally, the SOHR method provides new avenue for the quantification of rare events occurring in kinetic simulations. A rare event may be identified with a very low probability chemical pathway that leads to a given outcome, such as the production of a trace species. This rare event may exist simultaneously with other pathways that have much higher probabilities that dominate the chemistry and obscure the rare event.

Table 5.1: Lotka-Volterra System: The System Parameters and Paths Used in the Iterative Solution Procedure.

	reactions	rate coefficients	initial conditions
R_1	$A + X \rightarrow 2X$	$k_1 = 10$	$A(0) = 1.0$
R_2	$X + Y \rightarrow 2Y$	$k_2 = 10$	$B(0) = 1.0$
R_3	$Y \rightarrow B$	$k_3 = 10$	$X(0) = 1.0$ $Y(0) = 1.0$
	path		length
	A		0
	B		0
	X		0
	Y		0
	$A \rightarrow X$		1
	$X \rightarrow Y$		1
	$Y \rightarrow B$		1
	$A \rightarrow X \rightarrow Y$		2
	$X \rightarrow Y \rightarrow B$		2
	$A \rightarrow X \rightarrow Y \rightarrow B$		3

[†]Concentration of A can vary with time.

Chapter 6

Conclusion

The sum over histories representation provides a quantitative approach to investigate the contributions from distinct chemical pathways in determining the chemistry of large reaction networks. The basic concept motivating the SOHR is that the value of an observable target in a kinetic system may be computed from a knowledge of all the chemical pathways leading to that target. Thus, if we know the pathway probabilities P_j , then the observable can be composed from a linear combination $\sum c_j P_j$ where c_j are trivial coefficients obtained from the rate equations. The implementation of the SOHR thus involves two main tasks: (1) enumerating the most important chemical pathways associated with a given observable, and (2) efficiently computing the probabilities of those pathways. The chemical pathways were defined using a atom-following approach. We have shown that for sufficiently simple problems an enumeration of paths based on graph theoretical methods can converge the expansion of observables by increasing path length. For complicated mechanisms, a more efficient approach utilizes random walks through the reaction network to identify the most important chemical paths. Once the pathways have been located, the associated probabilities are computed using the time-ordered product, eqn. 2.15. This formula was derived assuming that the reactions along the pathway form a Markov chain of random events. The integral of the time-ordered product can be evaluated analytically for linear kinetics or for nonlinear systems in steady state. For more general nonlinear systems, the expression is efficiently evaluated using a MC-integration scheme, eqn. 2.16. The SOHR method differs from other techniques for reaction path analysis which represent the pathways at an instant of time using a flux snapshot.[47] Instead,

our method permits the time evolving chemistry of the system to be quantitatively incorporated into the probabilities for the chemical pathways. Hence, the ranking of the chemical pathways followed by chemical moieties are allowed to change as time progresses.

The SOHR provides a new tool to interpret the quantitative kinetics of large networks of reactions. Since chemical pathways are, in essence, mechanisms for the creation of specified species, this method provides an attractive physical picture to understand the behaviour of large models. In the sum over histories representation, the concentrations of the chemical species are decomposed into the sum of probabilities for chemical pathways that follow molecules from reactants to products or intermediates. Unlike static flux methods for reaction path analysis, the sum over histories approach includes the explicit time dependence of the pathway probabilities. Using the sum over histories representation, the sensitivity of an observable with respect to a kinetic parameter such as a rate coefficient is then analyzed in terms of how that parameter affects the chemical pathway probabilities. The method is illustrated for species concentration target functions in H₂ combustion where the rate coefficients are allowed to vary over their associated uncertainty ranges. As we have shown, e.g. the sensitivity of target functions to rate coefficients in kinetic simulations can have a physically appealing interpretation in terms of the behaviour of chemical pathways. The reaction can prove to be sensitive if it is a rate limiting step on a dominant pathway to the target. Kinetic simplification[86, 57, 67] of large mechanisms may likewise be guided by a knowledge of important (or unimportant) paths leading to products.

In addition to interpreting chemical kinetics, we demonstrated SOHR can be developed into an independent computational technique that is not dependent upon other representations of the kinetics. This newly developed method provides an alternative to conventional modeling of mass-action kinetics that involves solving differential equations for the species concentrations. The method presented in this thesis avoids the need to solve the rate equations by switching to a representation based on chemical pathways. In the SOHR method, any time-dependent kinetic observable, such as concentration, is written as a linear combination of probabilities for chemical pathways leading to a desired outcome. In this work, an iterative method is introduced that allows the time-dependent

pathway probabilities to be generated from a knowledge of the elementary rate coefficients thus avoiding the pitfalls involved in solving the differential equations of kinetics. The method was successfully applied to the model Lotka-Volterra system and to a realistic H₂-O₂ combustion model.

Bibliography

- [1] E Abel. Über einen besonders einfachen fall von stufenreaktionen. *Zeitschrift für Physikalische Chemie*, 56(1):558–564, 1906.
- [2] David Andrieux and Pierre Gaspard. Fluctuation theorem for currents and schnakenberg network theory. *Journal of statistical physics*, 127(1):107–131, 2007.
- [3] Abbin Antony, Aravind Asthagiri, and Jason F Weaver. Pathways and kinetics of methane and ethane c–h bond cleavage on pdo (101). *The Journal of chemical physics*, 139(10):104702, 2013.
- [4] Shirong Bai, Michael J Davis, and Rex T Skodje. Sum over histories representation for kinetic sensitivity analysis: How chemical pathways change when reaction rate coefficients are varied. *The Journal of Physical Chemistry A*, 119(45):11039–11052, 2015.
- [5] Shirong Bai and Rex T Skodje. The sum over histories representation for chemical kinetics: a quantitative theory based on chemical pathways. *International Reviews in Physical Chemistry*, 35(4):539–567, 2016.
- [6] Shirong Bai and Rex T Skodje. Simulating chemical kinetics without differential equations: A quantitative theory based on chemical pathways. *The Journal of Physical Chemistry Letters*, 8(16):3826–3833, 2017.
- [7] Shirong Bai, Dingyu Zhou, Michael J Davis, and Rex T Skodje. Sum over histories representation for chemical kinetics. *The journal of physical chemistry letters*, 6(1):183–188, 2014.
- [8] Alexandru T Balaban. *Chemical applications of graph theory*. Academic Press, 1976.
- [9] Frédérique Battin-Leclerc. Detailed chemical kinetic models for the low-temperature combustion of hydrocarbons with application to gasoline and diesel fuel surrogates. *Progress in Energy and Combustion Science*, 34(4):440–498, 2008.
- [10] Richard Bellman. On a routing problem. *Quarterly of applied mathematics*, 16(1):87–90, 1958.
- [11] Bruce J Berne and D Thirumalai. On the simulation of quantum systems: path integral methods. *Annual Review of Physical Chemistry*, 37(1):401–424, 1986.
- [12] Norman Biggs. *Algebraic graph theory*. Cambridge university press, 1993.

- [13] Jianshu Cao and Gregory A Voth. The formulation of quantum statistical mechanics based on the feynman path centroid density. i. equilibrium properties. *The Journal of chemical physics*, 100(7):5093–5105, 1994.
- [14] Jia-Ming Chern and Friedrich G Helfferich. Effective kinetic modeling of multistep homogeneous reactions. *AIChE Journal*, 36(8):1200–1208, 1990.
- [15] JA Christiansen. The elucidation of reaction mechanisms by the method of intermediates in quasi-stationary concentrations. *Advances in Catalysis*, 5:311–353, 1953.
- [16] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. Introduction to algorithms second edition, 2001.
- [17] RI Cukier, CM Fortuin, Kurt E Shuler, AG Petschek, and JH Schaibly. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i theory. *The Journal of chemical physics*, 59(8):3873–3878, 1973.
- [18] RI Cukier, JH Schaibly, and Kurt E Shuler. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. iii. analysis of the approximations. *The Journal of Chemical Physics*, 63(3):1140–1149, 1975.
- [19] Henry J Curran, Paolo Gaffuri, William J Pitz, and Charles K Westbrook. A comprehensive modeling study of n-heptane oxidation. *Combustion and flame*, 114(1):149–177, 1998.
- [20] Michael J Davis, Rex T Skodje, and Alison S Tomlin. Global sensitivity analysis of chemical-kinetic reaction mechanisms: construction and deconstruction of the probability density function. *The Journal of Physical Chemistry A*, 115(9):1556–1578, 2011.
- [21] Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.
- [22] David Eppstein. Finding the k shortest paths. *SIAM Journal on computing*, 28(2):652–673, 1998.
- [23] Charles Hilliard Feinstein. Socialism, capitalism and economic growth. Technical report, Cambridge University Press, 1975.
- [24] Haidong Feng, Bo Han, and Jin Wang. Dominant kinetic paths of complex systems: gene networks. *The Journal of Physical Chemistry Letters*, 1(12):1836–1840, 2010.
- [25] Richard P Feynman, Albert R Hibbs, and Daniel F Styer. *Quantum mechanics and path integrals*. Courier Corporation, 2010.
- [26] Ilie Fishtik, Caitlin A Callaghan, and Ravindra Datta. Reaction route graphs. i. theory and algorithm. *The Journal of Physical Chemistry B*, 108(18):5671–5682, 2004.
- [27] Ramón García-Domenech, Jorge Gálvez, Jesus V de Julián-Ortiz, and Lionello Pogliani. Some new trends in chemical graph theory. *Chemical Reviews*, 108(3):1127–1169, 2008.
- [28] Crispin W Gardiner. Handbook of stochastic methods for physics, chemistry and the natural sciences, + berlin. Springer-Verlag Gibson, JJ (1986). *The ecological approach to visual perception*. Hillsdale, Nj: Lawrence Erlbaum Associates, Inc.(Original work published 1979) Hosking, JRM (1981), Fractional differencing. *Biometrika*, 68:165–176, 1985.

- [29] Michael A Gibson and Jehoshua Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *The journal of physical chemistry A*, 104(9):1876–1889, 2000.
- [30] RG Gilbert, K_ Luther, and J Troe. Theory of thermal unimolecular reactions in the fall-off range. ii. weak collision rate constants. *Berichte der Bunsengesellschaft für physikalische Chemie*, 87(2):169–177, 1983.
- [31] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.
- [32] Daniel T Gillespie, Andreas Hellander, and Linda R Petzold. Perspective: Stochastic algorithms for chemical kinetics. *The Journal of chemical physics*, 138(17):05B201_1, 2013.
- [33] Jeff Greeley and Manos Mavrikakis. Competitive paths for methanol decomposition on pt (111). *Journal of the American Chemical Society*, 126(12):3910–3919, 2004.
- [34] Xiang-Kui Gu and Wei-Xue Li. First-principles study on the origin of the different selectivities for methanol steam reforming on cu (111) and pd (111). *The Journal of Physical Chemistry C*, 114(49):21539–21547, 2010.
- [35] Bjørk Hammer and Jens Kehlet Nørskov. Theoretical surface science and catalysiscalculations and concepts. *Advances in catalysis*, 45:71–129, 2000.
- [36] Peter J Hansen and Peter C Jurs. Chemical applications of graph theory. part i. fundamentals and topological indices. *J. Chem. Educ.*, 65(7):574, 1988.
- [37] Frank Harary. The determinant of the adjacency matrix of a graph. *Siam Review*, 4(3):202–210, 1962.
- [38] Frank Harary and Gopal Gupta. Dynamic graph models. *Mathematical and Computer Modelling*, 25(7):79–87, 1997.
- [39] Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.
- [40] Hamid Hashemi, Jakob M Christensen, Sander Gersen, and Peter Glarborg. Hydrogen oxidation at high pressure and intermediate temperatures: Experiments and kinetic modeling. *Proceedings of the Combustion Institute*, 35(1):553–560, 2015.
- [41] Kaiyuan He, Marianthi G Ierapetritou, and Ioannis P Androulakis. A graph-based approach to developing adaptive representations of complex reaction mechanisms. *Combustion and Flame*, 155(4):585–604, 2008.
- [42] Graeme Henkelman, Blas P Uberuaga, and Hannes Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of chemical physics*, 113(22):9901–9904, 2000.
- [43] Terrell L Hill. Studies in irreversible thermodynamics iv. diagrammatic representation of steady state fluxes for unimolecular systems. *Journal of theoretical biology*, 10(3):442–459, 1966.

- [44] Kiyosi Itô. Wiener integral and feynman integral. In Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, volume 2, pages 227–238, 1961.
- [45] NV Kampen. The langevin approach. Stochastic Processes in Physics and Chemistry (Third Edition), Elsevier, Amsterdam, page 219, 2007.
- [46] Darshan MA Karwat, Scott W Wagnon, Paul D Teini, and Margaret S Wooldridge. On the chemical kinetics of n-butanol: ignition and speciation studies. The Journal of Physical Chemistry A, 115(19):4909–4921, 2011.
- [47] RJ Kee, FM Ruply, and JA Miller. Chemkin-pro, reaction design. Inc., San Diego, CA, 2008.
- [48] Alan Kéromnès, Wayne K Metcalfe, Karl A Heufer, Nicola Donohoe, Apurba K Das, Chih-Jen Sung, Jürgen Herzler, Clemens Naumann, Peter Griebel, Olivier Mathieu, et al. An experimental and detailed chemical kinetic modeling study of hydrogen and syngas mixture oxidation at elevated pressures. Combustion and Flame, 160(6):995–1011, 2013.
- [49] Nick J Killingsworth, Vi H Rapp, Daniel L Flowers, Salvador M Aceves, J-Y Chen, and Robert Dibble. Increased efficiency in si engine with air replaced by oxygen in argon mixture. Proceedings of the Combustion Institute, 33(2):3141–3149, 2011.
- [50] Edward L King and Carl Altman. A schematic method of deriving the rate laws for enzyme-catalyzed reactions. The Journal of physical chemistry, 60(10):1375–1378, 1956.
- [51] Stephen J Klippenstein, Lawrence B Harding, Michael J Davis, Alison S Tomlin, and Rex T Skodje. Uncertainty driven theoretical kinetics studies for ch 3 oh ignition: Ho 2+ ch 3 oh and o 2+ ch 3 oh. Proceedings of the Combustion Institute, 33(1):351–357, 2011.
- [52] Alexander A Konnov. Remaining uncertainties in the kinetic mechanism of hydrogen combustion. Combustion and flame, 152(4):507–528, 2008.
- [53] Zeb C Kramer, Xiang-Kui Gu, Dingyu DY Zhou, Wei-Xue Li, and Rex T Skodje. Following molecules through reactive networks: surface catalyzed decomposition of methanol on pd (111), pt (111), and ni (111). The Journal of Physical Chemistry C, 118(23):12364–12383, 2014.
- [54] Keith JKJ Laidler. Chemical kinetics. Harper Collins, New York, 3rd edition, 1987.
- [55] J-F Lamarque, Drew T Shindell, B Josse, PJ Young, I Cionni, Veronika Eyring, D Bergmann, Philip Cameron-Smith, William J Collins, R Doherty, et al. The atmospheric chemistry and climate model intercomparison project (accmip): overview and description of models, simulations and climate diagnostics. Geoscientific Model Development, 6:179–206, 2013.
- [56] JF Lamarque, LK Emmons, PG Hess, DE Kinnison, S Tilmes, F Vitt, CL Heald, EA Holland, PH Lauritzen, J Neu, et al. Cam-chem: description and evaluation of interactive atmospheric chemistry in the community earth system model, geosci. model dev., 5, 369–411, doi: 10.5194, 2012.
- [57] Chung K Law, Chih Jen Sung, Hai Wang, and TF Lu. Development of comprehensive detailed and reduced reaction mechanisms for combustion modeling. AIAA journal, 41(9):1629–1646, 2003.

- [58] Daeyup Lee and Simone Hochgreb. Hydrogen autoignition at pressures above the second explosion limit (0.6–4.0 mpa). *International journal of chemical kinetics*, 30(6):385–406, 1998.
- [59] Ralph Lehmann. An algorithm for the determination of all significant pathways in chemical reaction systems. *Journal of atmospheric chemistry*, 47(1):45–78, 2004.
- [60] Charles E Leiserson, Ronald L Rivest, Clifford Stein, and Thomas H Cormen. Introduction to algorithms, 2001.
- [61] Genyuan Li, Carey Rosenthal, and Herschel Rabitz. High dimensional model representations. *The Journal of Physical Chemistry A*, 105(33):7765–7777, 2001.
- [62] Genyuan Li, Sheng-Wei Wang, and Herschel Rabitz. Practical approaches to construct rs-hdmr component functions. *The Journal of Physical Chemistry A*, 106(37):8721–8733, 2002.
- [63] Juan Li, Zhenwei Zhao, Andrei Kazakov, Marcos Chaos, Frederick L Dryer, and James J Scire. A comprehensive kinetic mechanism for co, ch₂o, and ch₃oh combustion. *International Journal of Chemical Kinetics*, 39(3):109–136, 2007.
- [64] Alfred J Lotka. Contribution to the theory of periodic reactions. *The Journal of Physical Chemistry*, 14(3):271–274, 1910.
- [65] Tianfeng Lu and Chung K Law. A directed relation graph method for mechanism reduction. *Proceedings of the Combustion Institute*, 30(1):1333–1341, 2005.
- [66] Tianfeng Lu and Chung K Law. On the applicability of directed relation graphs to the reduction of reaction mechanisms. *Combustion and Flame*, 146(3):472–483, 2006.
- [67] Ulrich Maas and Stephen B Pope. Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space. *Combustion and flame*, 88(3):239–264, 1992.
- [68] Nancy Makri. Time-dependent quantum methods for large systems. *Annual review of physical chemistry*, 50(1):167–191, 1999.
- [69] Donald A McQuarrie. Stochastic approach to chemical kinetics. *Journal of applied probability*, 4(3):413–478, 1967.
- [70] Gregory J McRae, James W Tilden, and John H Seinfeld. Global sensitivity analysisa computational implementation of the fourier amplitude sensitivity test (fast). *Computers & Chemical Engineering*, 6(1):15–25, 1982.
- [71] O Mekenyany, SS Basak, D Bonchev, and O Mekenyany. Graph theoretical approaches to chemical reactivity, 1994.
- [72] James A Miller, Robert J Kee, and Charles K Westbrook. Chemical kinetics and combustion modeling. *Annual Review of Physical Chemistry*, 41(1):345–387, 1990.
- [73] Eric Mjolsness. Time-ordered product expansions for computational stochastic system biology. *Physical biology*, 10(3):035009, 2013.
- [74] Habib N Najm, Bert J Debusschere, Youssef M Marzouk, Steve Widmer, and OP Le Maître. Uncertainty quantification in chemical systems. *International journal for numerical methods in engineering*, 80(6-7):789–814, 2009.

- [75] George Oster, Alan Perelson, and Aharon Katchalsky. Network thermodynamics. *Nature*, 234(5329):393–399, 1971.
- [76] Michael J Pilling, Paul W Seakins, et al. *Reaction kinetics*. Oxford University Press, 1996.
- [77] Herschel Rabitz and Ömer F Aliş. General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry*, 25(2-3):197–233, 1999.
- [78] Marc R Roussel and Simon J Fraser. Geometry of the steady-state approximation: Perturbation and accelerated convergence methods. *The Journal of chemical physics*, 93(2):1072–1081, 1990.
- [79] A Saltelli, K Chan, and EM Scott. 2000, sensitivity analysis.
- [80] Andrea Saltelli, Karen Chan, E Marian Scott, et al. *Sensitivity analysis*, volume 1. Wiley New York, 2000.
- [81] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.
- [82] Jürgen Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern physics*, 48(4):571, 1976.
- [83] Lawrence S Schulmann. Techniques and applications of paths integrals, 1996.
- [84] James J Scire, Frederick L Dryer, and Richard A Yetter. Comparison of global and local sensitivity techniques for rate constants determined using complex reaction mechanisms. *International Journal of Chemical Kinetics*, 33(12):784–802, 2001.
- [85] Robert D Skeel and Hon-Wah Tam. Limits of parallelism in explicit ode methods. *Numerical Algorithms*, 2(3):337–349, 1992.
- [86] Rex T Skodje and Michael J Davis. Geometrical simplification of complex kinetic systems. *The Journal of Physical Chemistry A*, 105(45):10356–10365, 2001.
- [87] Rex T Skodje, Alison S Tomlin, Stephen J Klippenstein, Lawrence B Harding, and Michael J Davis. Theoretical validation of chemical kinetic mechanisms: combustion of methanol. *The Journal of Physical Chemistry A*, 114(32):8286–8301, 2010.
- [88] Ilya M Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, 55(1):271–280, 2001.
- [89] Sibendu Som, Wei Liu, Dingyu DY Zhou, Gina M Magnotti, Raghu Sivaramakrishnan, Douglas E Longman, Rex T Skodje, and Michael J Davis. Quantum tunneling affects engine performance. *The journal of physical chemistry letters*, 4(12):2021–2025, 2013.
- [90] Gabor A Somorjai and Yimin Li. *Introduction to surface chemistry and catalysis*. John Wiley & Sons, 2010.
- [91] Jeffrey I Steinfeld, Joseph Salvadore Francisco, and William L Hase. *Chemical kinetics and dynamics*. Prentice Hall Englewood Cliffs (New Jersey), 1989.

- [92] Oleg N Temkin and Danail G Bonchev. Application of graph theory to chemical kinetics: Part 1. kinetics of complex reactions. *J. Chem. Educ.*, 69(7):544, 1992.
- [93] Oleg N Temkin, Andrew V Zeigarnik, and DG Bonchev. *Chemical reaction networks: a graph-theoretical approach*. CRC Press, 1996.
- [94] Luca Tosatto and Youssef Marzouk. Simplifying chemical kinetic systems under uncertainty using markov chains. In *extended abstract, Workshop on Numerical Combustion*, 2013.
- [95] Tamás Turányi. Sensitivity analysis in chemical kinetics. *International Journal of Chemical Kinetics*, 40(11):685–686, 2008.
- [96] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.
- [97] Vito Volterra. Variations and fluctuations of the number of individuals in animal species living together. *ICES Journal of Marine Science*, 3(1):3–51, 1928.
- [98] Hai Wang and David A Sheen. Combustion kinetic model uncertainty quantification, propagation and minimization. *Progress in Energy and Combustion Science*, 47:1–31, 2015.
- [99] J Warnatz, U Maas, and RW Dibble. Combustion: physical and chemical fundamentals, modeling and simulation, experiments, pollutant formation. 2006. *Int. J. Comp. Meth. and Exp. Meas.*, 5(1), 2017.
- [100] Jürgen Warnatz, Ulrich Maas, and Robert W Dibble. Physical and chemical fundamentals, modeling and simulation, experiments, pollutant formation. *Springer, Berlin Frenklach M, Wang H (1990) Detailed kinetic modelling of soot particle nucleation and growth. Proc Comb Inst*, 23:1559–1566, 2006.
- [101] Charles K Westbrook. Chemical kinetics of hydrocarbon ignition in practical combustion systems. *Proceedings of the Combustion Institute*, 28(2):1563–1577, 2000.
- [102] Judit Zádor, Craig A Taatjes, and Ravi X Fernandes. Kinetics of elementary reactions in low-temperature autoignition chemistry. *Progress in energy and combustion science*, 37(4):371–421, 2011.
- [103] Judit Zador, Istvan Gy Zsely, Tamas Turanyi, Marco Ratto, Stefano Tarantola, and Andrea Saltelli. Local and global uncertainty analyses of a methane flame model. *The Journal of Physical Chemistry A*, 109(43):9795–9807, 2005.
- [104] Dingyu DY Zhou, Michael J Davis, and Rex T Skodje. Multitarget global sensitivity analysis of n-butanol combustion. *The Journal of Physical Chemistry A*, 117(17):3569–3584, 2013.
- [105] Dingyu DY Zhou, Keli Han, Peiyu Zhang, Lawrence B Harding, Michael J Davis, and Rex T Skodje. Theoretical determination of the rate coefficient for the $\text{ho}_2 + \text{ho}_2 \text{ h}_2\text{o}_2 + \text{o}_2$ reaction: Adiabatic treatment of anharmonic torsional effects. *The Journal of Physical Chemistry A*, 116(9):2089–2100, 2012.
- [106] T Ziehn, KJ Hughes, JF Griffiths, R Porter, and AS Tomlin. A global sensitivity study of cyclohexane oxidation under low temperature fuel-rich conditions using hdmr methods. *Combustion Theory and Modelling*, 13(4):589–605, 2009.

- [107] T Ziehn and Alison S Tomlin. Gui-hdmr—a software tool for global sensitivity analysis of complex models. *Environmental Modelling & Software*, 24(7):775–785, 2009.
- [108] Tilo Ziehn and Alison S Tomlin. A global sensitivity study of sulfur chemistry in a premixed methane flame model using hdmr. *International Journal of Chemical Kinetics*, 40(11):742–753, 2008.
- [109] I Gy Zsély, J Zádor, and T Turányi. On the similarity of the sensitivity functions of methane combustion models. *Combustion Theory and Modelling*, 9(4):721–738, 2005.

Appendix A

The Hydrogen Combustion Model

The combustion of H₂ is modeled using a detailed mechanism that consists of 21 individual steps and 8 distinct chemical species. The specific mechanism employed here is a sub-mechanism of a larger model for methanol combustion used by Li et al.[63] The reactions and species used are listed in the tables below. The temperature dependence of the rate coefficients is given by the three parameter fit

$$k(T) = AT^\delta E^{-E_a/k_B T} \quad (\text{A.1})$$

where the values of A , δ , and E_a for each reaction are listed in Table A.1. The pressure-dependent reactions rate are calculated using the Troe centering expression

$$k = k_\infty \left(\frac{P_r}{1+P_r} F \right) \quad (\text{A.2})$$

where

$$P_r = \frac{k_0 [M]}{k_\infty} \quad (\text{A.3})$$

and the high pressure limit (k_∞) and low pressure limit (k_0) rate coefficients are parameterized by the three parameter expression in eqn. A.1. The function F ¹ is given by Troe and coworkers.[30] The reverse reaction rates are defined consistently with micro-reversibility using the equilibrium constant

$$k_r = \frac{k_f}{k_{eq}^c} \quad (\text{A.4})$$

¹ The function F is defined as $\log F = \left[1 + \left[\frac{\log P_r + c}{n - d(\log P_r + c)} \right] \right]^{-1} \log F_{cent}$ where $F_{cent} = (1 - \alpha) \exp(-T/T^{***}) + \alpha \exp(T/T^*) + \exp(-T^{**}/T)$ and centering parameters T^* , T^{**} , and T^{***} are provided in the table.

In the mechanism, there are two pairs of duplicated reactions; i.e., they have the same reactants and products but different rate coefficients. This allows a much more flexible representation of the overall temperature dependence of the rate coefficients. The combustion kinetics is simulated using the constraints of constant volume and constant energy. The time development of each species is modeled by the following set of equations:

$$\begin{aligned}\frac{dY_k}{dt} &= \frac{W_k \dot{\omega}_k}{\rho}, \quad k = 1, \dots, 8 \\ \frac{dT}{dt} &= -\frac{1}{\rho C_v} \sum_k W_k \dot{\omega}_k e_k\end{aligned}\tag{A.5}$$

where Y_k represents the mass fraction of species k (which is proportional to its concentration $[X_k]$), W_k is its molecular weight, $\dot{\omega}_k$ refers to the production rate of the k^{th} species obtained from the mechanism, ρ is the bulk density, C_v is the constant volume heat capacity and e_k is the internal energy of the k^{th} species.

Table A.1: Rate coefficients and uncertainty ranges of H₂-O₂ system.

Index	Reaction	A(cm ³ /mol-s) or s ⁻¹	δ	E _a (kcal/mol)	Uncertainty
0	H+O ₂ =O+OH	3.55E+15	-0.4	16599.0	1.26
1	O+H ₂ =H+OH	5.08E+04	2.7	6290.0	1.58
2	H ₂ +OH=H ₂ O+H	2.16E+08	1.5	3430.0	2.00
3	O+H ₂ O=OH+OH	2.97E+06	2.0	13400.0	2.50
4	H ₂ +M=H+H+M	4.58E+19	-1.4	104380.0	3.00
	H ₂			Enhanced by 2.500E+00	
	H ₂ O			Enhanced by 1.200E+01	
5	O+O+M=O ₂ +M	6.16E+15	-0.5	0.0	2.00
	H ₂			Enhanced by 2.500E+00	
	H ₂ O			Enhanced by 1.200E+01	
6	O+H+M=OH+M	4.71E+18	-1.0	0.0	5.00
	H ₂			Enhanced by 2.500E+00	
	H ₂ O			Enhanced by 1.200E+01	
7	H+OH+M=H ₂ O+M	3.80E+22	-2.0	0.0	2.00
	H ₂			Enhanced by 2.500E+00	
	H ₂ O			Enhanced by 1.200E+01	
8	H+O ₂ (+M)=HO ₂ (+M)	1.48E+12	0.6	0.0	3.16
	Low pressure limit		0.63660E+21	-0.17200E+01	0.52480E+03
	TROE centering		0.80000E+00	0.10000E-29	0.10000E+31
	H ₂			Enhanced by 2.000E+00	
	H ₂ O			Enhanced by 1.100E+01	
	O ₂			Enhanced by 7.800E-01	
9	HO ₂ +H=H ₂ +O ₂	1.66E+13	0.0	823.0	2.00
10	HO ₂ +H=OH+OH	7.08E+13	0.0	295.0	2.00
11	HO ₂ +O=O ₂ +OH	3.25E+13	0.0	0.0	3.16
12	HO ₂ +OH=H ₂ O+O ₂	2.89E+13	0.0	-497.0	3.16
13	HO ₂ +HO ₂ =H ₂ O ₂ +O ₂	4.20E+14	0.0	11982.0	5.00
	Declared duplicate reaction...				
14	HO ₂ +HO ₂ =H ₂ O ₂ +O ₂	1.30E+11	0.0	-1629.3	5.00
	Declared duplicate reaction...				
15	H ₂ O ₂ (+M)=OH+OH(+M)	2.95E+14	0.0	48430.0	3.16
	Low pressure limit		0.12020E+18	0.00000E+00	0.45500E+05
	TROE centering		0.50000E+00	0.10000E-29	0.10000E+31
	H ₂			Enhanced by 2.500E+00	
	H ₂ O			Enhanced by 1.200E+01	
16	H ₂ O ₂ +H=H ₂ O+OH	2.41E+13	0.0	3970.0	5.00
17	H ₂ O ₂ +H=HO ₂ +H ₂	4.82E+13	0.0	7950.0	5.00
18	H ₂ O ₂ +O=OH+HO ₂	9.55E+06	2.0	3970.0	3.00
19	H ₂ O ₂ +OH=HO ₂ +H ₂ O	1.00E+12	0.0	0.0	5.00
	Declared duplicate reaction...				
20	H ₂ O ₂ +OH=HO ₂ +H ₂ O	5.80E+14	0.0	9557.0	5.00
	Declared duplicate reaction...				

[†]Indices start at 0.