MultiScale Data-Driven Modeling of Foundational Combustion Reaction Systems

Carly LaGrotta

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2023

Abstract

MultiScale Data-Driven Modeling of Foundational Combustion Reaction Systems

Carly LaGrotta

As the world becomes increasingly interconnected, modernized, and populated, the demand for energy across the globe is growing at an unprecedented rate. This growth in energy demand has an undeniable impact on increasingly pressing social issues including, climate change, energy security, energy economy, atmospheric chemistry, and air quality. Finding a way to address these issues on a rapid timescale is more important than ever. A common thread running through all of these challenges is that they can be partially or fully addressed with the development of new chemical energy conversion technologies which, in turn, rely on a comprehensive understanding of gas phase kinetics.

Examples of promising technologies include renewable fuels (i.e. methanol and hydrogen) and/or reliable, efficient, and clean engines that can accommodate renewable fuels. The development of such technology would enable the use of renewable fuels, thereby reducing emissions and cutting down on harmful byproducts released into the atmosphere. Computational simulations have become a powerful approach for developing and advancing energy technology in a safe, efficient, and effective manner. These computational approaches model reacting flows and are generally known as computational fluid dynamics (CFD). However, in order for these CFD simulations to work effectively and make meaningful predictions, the sub-models used to describe the underlying chemistry (gas phase kinetics) must be accurate; information about underlying chemistry is provided to computational simulations via a chemical kinetic model/mechanism, which describes the chemical reactions that drive the fuel oxidation within the system being simulated. Regarding combustion specifically, the reliability of predictive simulations depends on the availability of accurate data and models not only for chemical kinetics, but also

thermochemistry and transport.

Further complicating the problem, combustion and chemical kinetics provide a unique challenge in regard to obtaining accurate predictive models; underlying chemical kinetics mechanisms may require unprecedented accuracy to obtain truly predictive combustion modeling. For example, it has been shown in computational simulations that uncertainties in any of several kinetic parameters can yield uncertainties large enough in the physical system being modeled to cause system failure, thereby reducing the effectiveness of computational design approaches that could accelerate technology development. Hence, a strong need exists to develop a method that significantly reduces uncertainties in chemical kinetics parameters to meet the accuracy demands of advanced computational design tools. To this end, it is useful to draw on inspiration from existing methods in the field of combustion and chemical kinetics as well as tangential fields; the most compelling inspiration can be found in the field of thermochemistry in the form of the Active Thermochemical Tables (ATcT).

This work presents a novel, analogous approach for chemical kinetics called MultiScale Informatics, or MSI for short. The MSI approach identifies optimized values and quantified uncertainties for a set of molecular parameters (within theoretical kinetics calculations), rate parameters, and physical model parameters (within simulations of experimental observables) as informed by data from various sources and scales. The overarching objectives of this work are to demonstrate how the MSI approach can be used to determine physically meaningful optimized kinetics parameters and quantified uncertainties, unravel webs of interconnected rate constants in complex reaction systems, resolve discrepancies among data sets, and touch on key elements of MSI's implementation.

To demonstrate how these objectives are met, the MSI approach is used to explore the kinetics of three reaction sub-systems. The studies of these sub-systems will demonstrate some key elements of this approach including: the importance of raw data for quantifying the information content of experimental data, the utility of theoretical kinetics calculations for constraining experimental interpretations and providing an independent data source, and the subtleties of target data

selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties.

For the first sub-system explored ($CH_3 + HO_2$), the MSI approach is applied to carefully selected (mostly raw) experimental data and yields an opposite temperature dependence for the channel-specific $CH_3 + HO_2$ rate constants as compared to a previous rate-parameter optimization. While both optimization studies use the same theoretical calculations to constrain model parameters, only the present optimization, which incorporates theory directly into the model structure, yields results that are consistent with theoretical calculations.

For the second sub-system explored ($HO_2 + HO_2$), the MSI approach is applied to carefully selected experimental data, leveraging the hydrogen reaction system from the first study with the addition of high level theory calculations for the reaction of $HO_2 + HO_2$. Recent high-level theoretical calculations predict a mild temperature dependence for $HO_2 + HO_2$, which is inconsistent with state-of-the-art experimental determinations that upheld the stronger temperature dependence observed in early experiments. Via MSI analysis of the theoretical and experimental data, alternative interpretations of the raw experimental data that uses $HO_2 + HO_2$ rate constants nearly identical to theoretical predictions are identified – implying that the theoretical and experimental data are actually consistent, at least when considering the raw data from experiments. Similar analyses of typical signals from low-temperature experiments indicate that an HOOOOH intermediate – identified by recent theory but absent from earlier interpretations – yields modest effects that are smaller than, but may have contributed to, the scatter in data among different experiments. More generally, the findings demonstrate that modern chemical theories and experiments have progressed to a point where meaningful comparison requires joint consideration of their data simultaneously.

The third sub-system explored builds a larger web of interconnected reaction systems in an attempt to achieve data redundancy and demonstrate how interpreting coupled reaction systems is necessary to accurately determine many key rate constants. The ability of the MSI method to interpret raw experimental data and untangle rate constant reaction systems is demonstrated. The

study also reinforces how implementing theory into the model structure is imperative to yield results that are consistent with experimental data as well as theoretical calculations and achieve physically realistic branching ratios.

Finally, this work will present how results from all the studied reaction systems culminate into a complex hydrogen/syngas combustion model validated against data from various combustion experiments.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost I would like to thank my friends and family for their continued love and support throughout this process. Mom and Dad, words can not express how grateful I am for your support and every opportunity you have given me. You have both been an inspiration personally and professionally. Thank you for allowing me to chase my dreams and helping me grow into the person I have become. Most of all thank you for reminding me in my lowest moments when I wanted to give up that I am a "big brave dog"!

My brothers, Anthony and Ricky, who could never understand why I would sign up for six additional years of school, thank you for supporting me anyway. Thank you for always making me laugh and helping me put my troubles in perspective. Most of all, thank you for reminding me that I am never too busy to take a break and enjoy life.

Aunt Elisa and Uncle Mike, thank you for all the help you have given me along the way. Aunta, you are always there when I need you and I know you would do anything for me, even read through a technical document you don't understand to tell me where I should have put commas.

Eric, thank you for supporting me throughout this very long journey. Your understanding and compassion through this process has meant so much to me. Thank you for always reminding me to be strong in moments I felt small. You always push me to be the best version of myself and fight for what I deserve.

The Volpe family, life truly would not be the same without you. Thank you for always reminding me of who I am and where I came from. You make every Sunday dinner more than just a dinner.

Christy and Reid, thank you for always being my second phone call (after my mom of course) when I get good or bad news. Thank you for always knowing the difference between when I want advice and when I just want to complain, an underrated skill I will always be thankful for.

Robert, Andrew, and Roth (three more people who could never understand why I signed up for six more years of school) thank you for keeping me humble and never letting me forget who I am. Your years of support and friendship mean so much to me.

Thank you to my fellow Burke Lab members, both past and present: Rodger Cornell, Mark Barbet, Lei Lei, Joe Lee, Qinghui Meng, Jonathan Pankauski, Patrick Singal, and Justin Gomes. I would not have been able to finish my degree without your help. You were always there to answer questions and lend a hand when I needed help. You have helped me maintain my sanity through the most stressful of moments, something that I will always be thankful for. Most importantly, you were all always down for a distracting conversation or a coffee trip to Mudd when I needed to step away from work. You started off as coworkers but have since become dear friends. The only thing I will not miss about our time together is Mark's space heater.

Thank you to all of my committee members: Professor Arvind Narayanaswamy, Professor Vijay Vedula, Professor Daniel Esposito, and Professor Lauren Marbella. I cannot thank you all enough for taking the time to review my thesis and provide invaluable feedback that I have no doubt will improve the quality of my work.

Last but not least, I would like to thank my advisor, Professor Michael P. Burke, for all the time and energy you have given me over the last six years. Your dedication and guidance throughout this process has been invaluable.

Dedication

To my family and friends,

your unwavering support and encouragement has meant the world to me.

Chapter 1: Background and Introduction

1.1 Background

As society becomes increasingly globalized, the demand for energy across the world is growing at an unprecedented rate [1]. The increasing global energy demand is undeniably tied to pressing social issues including climate change, energy security, energy economy, atmospheric chemistry, and air quality [2, 3]. A common thread running through all of these challenges is that they can be partially or fully addressed with the development of new chemical energy conversion technologies, which in turn require a comprehensive understanding of gas phase kinetics (the study of the motion of gasses and the chemical reactions / interactions occurring between them).

Climate change, which is the most ubiquitous of these challenges, is being exacerbated by the the extensive use of fossil fuels (i.e. combustibles). Begging the question, why as a society do we continue to use fuels? The energy density and efficiency of fuels are unparalleled compared to that of most other energy sources. Fig. 1 and Fig. 2 give clear visual representations of the advantage that fuels provide for energy storage as compared to renewables and current battery technology.

These significant differences in energy density exemplify why fuels will continue to remain as the standard power source for applications where minimizing weight and maximizing continuous travel time are important factors (i.e. cargo shipping, freight trains, long haul trucking, air travel, or space travel). Given that fuels will remain the standard power source in these scenarios, the improvement and development of alternative fuel sources that minimize greenhouse gas emissions and curb climate change is critical. Two promising alternative fuels for minimizing harmful emissions (and most relevant to the research herein) are methanol (CH_3OH) and hydrogen (H_2).

Methanol is a versatile fuel that can be used either on its own or blended with other fuels to improve engine efficiency and reduce emissions; methanol also has a comparable energy density to

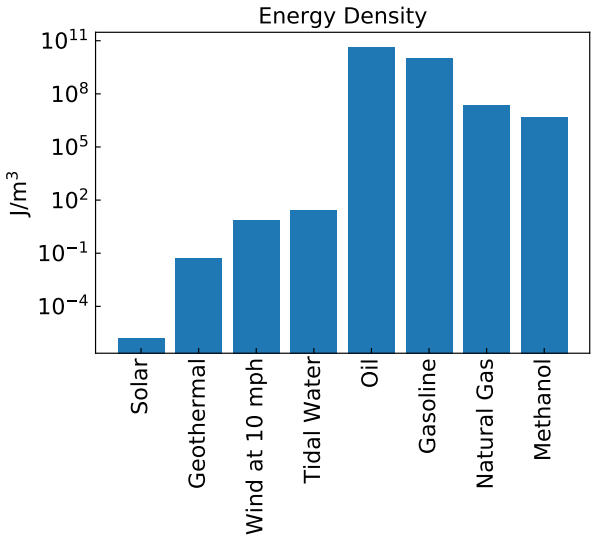


Figure 1: Energy density for various fuels verses various renewable energy sources. [4]

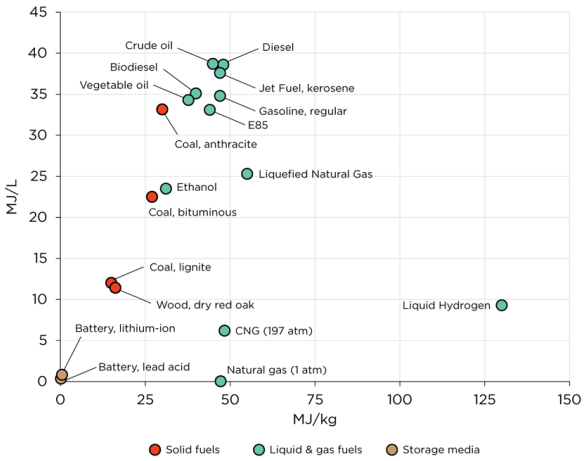


Figure 2: Energy density for various fuels verses batteries. [5]

that of traditional combustibles such as gasoline or oil [6, 7]. The formulation process for methanol has traditionally been a steam reforming process powered by fossil fuels, which produce harmful emissions, making methanol an unattractive option as an alternative fuel source. Although steam reforming is still the most common method for producing methanol, there is ongoing research to determine more environmentally sound production methods, [8] namely using more sustainably sourced reactants (CO_2 and H_2). For example, in one method of production, excess energy produced by renewable resources powers an electrolysis process, yielding hydrogen gas; carbon

dioxide is either extracted from industrial waste streams or collected with carbon capture techniques. The hydrogen and carbon dioxide gases are heated, pressurized, and reacted in a synthesis reactor (also powered by excess energy from renewable resources) to produce methanol [9].

Methanol's viability as a mainstream fuel source increases as it becomes more environmentally friendly to produce [10]. Studies have shown that the addition of methanol to gasoline significantly improves the brake thermal efficiency of traditional combustion engines and lowers CO and NO_x emissions [11]. Additionally, the use of an 85% / 15% methanol to gasoline ratio can reduce CO and NO_x emissions by about 25% and 80%, respectively, as compared to pure gasoline [11], while engines operating using pure methanol could reduce CO emissions by an additional 70% and would completely eliminate SO_x and particulate matter emissions [8]. Considering the transport sector alone contributed 20% of total greenhouse gas emissions worldwide in 2021, the widespread adoption of vehicles powered by renewable methanol would dramatically curb CO_2 and other greenhouse gas emissions [12, 13].

However, adopting methanol as an alternative fuel comes with a variety of challenges, three of which are notable. First, methanol poses the major operating risk of cold engine start failure due to both the lower volatility and burning temperature as compared to that of traditional gasoline [14]. Second, the acidic byproducts of methanol combustion cause the engine valves, valve seats, and cylinders to wear out at a higher rate compared to traditional gasoline combustion [15]. Third, methanol is hygroscopic and absorbs water vapor directly from the atmosphere, causing engine failure once a certain concentration of water is introduced. Through the addition of different fuel additives or alternative engine designs, these challenges among others can certainly be addressed, but addressing these challenges and determining ideal and safe operating conditions for the system requires complex modeling and simulation. Simulations of these reacting flows require a comprehensive understanding of the gas kinetics at work.

Hydrogen (H_2) is another promising alternative fuel source as it produces minimal NO_x emissions and no CO_x, SO_x, or particulate matter emissions [16]. In addition to the clean burning properties of hydrogen, studies have shown that hydrogen combustion engines can reach up to

60% brake thermal efficiency, while on average, traditional commercial engines achieve 30% brake thermal efficiency [17]. Similar to methanol, hydrogen was not considered a viable alternative fuel source until recently because its production traditionally required fossil fuels. As discussed, hydrogen gas can now be generated through water electrolysis or stream-methane reforming using excess energy from renewable resources [18, 19].

The physical properties of hydrogen introduce unique design challenges when it is used as a fuel source. Hydrogen gas is the smallest molecule in the world, meaning it easily leaks through traditional engine components. Its small molecule size and the gaseous state of hydrogen require redesigns for valves, valve seats, the fuel injector, and various other engine components [20]. Another major design challenge with hydrogen combustion engines is determining an air/fuel ratio and operating conditions that optimize power output while minimizing NO_x emissions. When hydrogen is burned at a stoichiometric ratio, the combustion temperature is very high and a large amount of nitrogen oxides (NO_x) are produced; for this reason hydrogen engines are not normally designed to operate at these conditions. Typically, hydrogen engines are designed to use about twice as much air as theoretically required for complete combustion to almost entirely eliminate NO_x emissions. Unfortunately, this reduces the power output to about half that of a similarly sized gasoline engine. It is possible to address all of these challenges with continued studies, design, and simulations, and in fact there are numerous studies [21, 22, 23, 24] (and references therein) working to do just this. These simulations and studies of the reacting flows require a comprehensive understanding of the gas kinetics at work.

Alternative fuels are not only useful for the transportation sector but also have promising applications for energy distribution / power generation for everyday consumer use. The power grid is the infrastructure used to distribute energy (in the form of electricity) from producers to consumers, and in order to maintain grid stability, the energy generated must be roughly equal to the energy consumed. Grid stability relies on the grid's ability to respond to volatility in voltage and frequency disturbances within an acceptable time frame to maintain equilibrium. The stochastic nature of traditional renewable energy sources (wind, solar, geothermal etc.) makes it impossible

for them to reliably and consistently provide energy on demand. This means the power grid can not exclusively rely on them [25, 26] without risking instability. In fact, studies have shown once a certain percentage of power is provided to the electrical grid by renewable sources (anywhere between 20% and 77%), the grid begins to become unstable. Instability of the power grid can lead to power failures and sweeping blackouts. Studies suggest that by adding large energy storage capability (which could supply synchronous power) to the grid, it may be possible to operate the grid with up to 96% of power generated by renewable sources [27]. Batteries are a possible storage solution, but their lower energy density as compared to that of fuels means that the batteries would require much more space to store the same amount of energy. Thus, using renewable fuels as an energy storage solution is an extremely viable, long-term, space-saving, and environmentally friendly solution [28, 29]. Specifically, when considering hydrogen as an alternative fuel, its long term energy storage capabilities are superior to those of batteries [30].

Grid stability must be maintained whether energy production exceeds demand or demand exceeds production. If energy production is exceeding demand, excess energy produced by renewables is used to power production of hydrogen or methanol. The hydrogen or methanol produced can be stored as either liquids or compressed gasses. Alternatively, if peak power demand exceeds power provided by renewable resources, the hydrogen or methanol can be burned to power generators (providing a more environmentally friendly alternative to natural gas, which is currently used). If the stored methanol or hydrogen are burned in industrial turbines as energy sources for power production, the same types of design challenges discussed above need to be addressed to safely, efficiently, and cleanly combust them on an industrial scale. Moreover, this would again require the intimate understanding of the gas phase kinetics in the reacting flows at play.

Developing renewable fuels and advancing the state of chemical energy conversion technology is a natural next step towards a better future; the development of such technology would help reduce emissions and thus cut down on harmful byproducts released into the atmosphere [31, 32]. Computational simulations have become a powerful approach for developing and advancing energy technology in a safe, efficient, and effective manner [33, 34, 35, 36, 37, 38, 39]. These computa-

tional approaches model reacting flows and are generally known as computational fluid dynamics (CFD) [40]. However, in order for these CFD simulations to work effectively and make meaningful predictions, the sub-models used to describe the underlying chemistry (gas kinetics) must be accurate. Information about underlying chemistry is provided to computational simulations (CFD simulations) via a chemical kinetic model/mechanism, which describes the chemical reactions that drive the fuel oxidation within the combustion system being simulated. In fact, accelerating energy technology development with predictive simulations that use chemical kinetics models has already proven to be successful, including the development of a new engine concept with about 50% better fuel economy [36]. However, it has been shown that uncertainties in any of several kinetic parameters can yield uncertainties large enough in the physical system being modeled to cause system failure. Given the proven success of this workflow, which hinges on the accuracy of chemical kinetics parameters, the next natural step is to focus on improving the chemical kinetics sub-models.

1.2 Introduction

Defining an accurate chemical kinetics model depends on having an intimate understanding of a complex reaction network; the challenge of understanding complex reaction networks is that they do not proceed through a single step. Rather, these reaction networks proceed through many smaller, intermediate steps that can occur simultaneously and sequentially, depending on the system. In fact, complex reaction networks often involve thousands of individual chemical reactions. Meaning, to understand a full system of reactions, one needs to fully understand the rate constants governing hundreds to tens of thousands of individual reactions.

In the most general terms, a rate constant is defined as a coefficient of proportionality, relating the rate of a chemical reaction at a given temperature, and/or pressure, and/or mixture composition to the concentration of a reactant (in a unimolecular reaction) or to the product of the concentrations of reactants.

$$\sum_{i=1}^{N} \nu_i' \mathrm{M}_i \xrightarrow{k_f} \sum_{i=1}^{N} \nu_i'' \mathrm{M}_i \tag{1.1}$$

One way these chemical reactions can be represented is in the form of Eq. 1.1, where M_i is the chemical symbol for the i^{th} species, ν_i is the corresponding molar concentration coefficient in the reactants (') and the products("). Associated with every forward reaction Eq. 1.1, is the corresponding reverse reaction 1.2.

$$\sum_{i=1}^{N} \nu_i'' \mathrm{M}_i \xrightarrow{k_b} \sum_{i=1}^{N} \nu_i' \mathrm{M}_i \tag{1.2}$$

The rates of each reaction, ω_k, depend on the local and instantaneous temperature, pressure, and composition via the law of mass action (Eq. 1.3).

$$\omega = k_f(T,P,M) \prod_{i=1}^{N} c_i^{\nu_i'} - k_b(T,P,M) \prod_{i=1}^{N} c_i^{\nu_i''} \tag{1.3}$$

Only k_f or k_b must actually be determined because at equilibrium, $\omega \equiv 0$, and the rate of forward reaction is balanced by that of backward reaction. Therefore, Eq. 1.3 yields Eq. 1.4.

$$k_f(T,P,M)/k_b(T,P,M) = \prod_{i=1}^{N} c_i^{(\nu_i''-\nu_i')} = K_c \tag{1.4}$$

Substituting Eq. 1.4 into Eq. 1.3 yields Eq. 1.5

$$\omega = k_f(T,P,M)(\prod_{i=1}^{N} c_i^{\nu_i'} - K_c(T,P,M)^{-1} \prod_{i=1}^{N} c_i^{\nu_i''}) \tag{1.5}$$

Since K_c can usually be determined to a much greater accuracy than k_b by using the ATcT [41], Eq. 1.5 is preferred over Eq. 1.3. The proportionality factors, $k_f(T,P,M)$ and $k_b(T,P,M)$ (in the law of mass action), are the forward and reverse reaction rate constants and are functions of temperature (T), pressure (P), and mixture composition (M).

Given the inherent complexity of chemical reaction networks with regard to the shear number of rate constants and their thermodynamic dependencies, it is not surprising that deriving com-

putational models to represent these systems may prove challenging. However, combustion and chemical kinetics provide an additional, unique challenge with regard to obtaining accurate predictive models; underlying chemical kinetics mechanisms require unprecedented accuracy to obtain truly predictive combustion modeling. For example, engine simulations show that uncertainties in any of several rate constants yield uncertainties in ignition crank angle of 1-2°, enough of a deviation to cause misfire [37]. Additionally, even knowing all rate constants as accurately as we know the best characterized rate constants (10-15%) will not yield a truly predictive simulation of flame propagation [42] or extinction [43]. These results exemplify how achieving the full potential of predictive simulations poses a significant challenge given that the uncertainties for even the most studied reactions are still non-negligible, and the uncertainties for other reactions are much larger (and sometimes go undefined) even after they are extensively investigated [44].

A nearly universal theme in experimental rate constant determinations is that the uncertainties generally exceed that which is dictated by measurement noise alone. Completely isolating a single reaction (especially at combustion temperatures) is nearly impossible, therefore uncertainties in rate constants for other reactions are often among the largest sources of uncertainty. This interdependence results in a complex web of implicit connections in rate constant determinations– with two extremely significant implications:

1. The systematic uncertainties associated with ignoring these conditions lead to inconsistencies among determinations that often exceed experimental precision [44].

2. There is more information in the existing data than has been possible to extract previously using traditional methods.

A strong need exists to develop a method to obtain the most accurate kinetics parameters with well-defined uncertainties. As a result of the rapid decrease in the cost of computing power, several research groups have attempted various optimization approaches to determine more accurate kinetics parameters. However, with regard to addressing the aforementioned challenges, the field of thermochemistry (which has many parallels to the chemical kinetics field) offers the most in-

spirational precedent– the Active Thermochemical Tables (ATcT) [41]. The ATcT use informatics based strategies to achieve thermochemical data of unprecedented accuracy by quantitatively accounting for the implicit connections in networks of interdependent data. Additionally, the ATcT have enabled major advancements in the *ab initio* electronic structure and other theoretical methods.

Achieving an analog to the ATcT in chemical kinetics presents additional complexities and challenges, which dictate strict requirements on the elements of any universally applicable approach. Such an approach would require not only the interpretation of raw experimental data (to unravel the web of implicit connections among rate constant determinations), but also a theoretical kinetic backbone (to enable reliable extrapolation of rate constants whose temperature/pressure/mixture composition (T/P/M) dependence are ill-constrained by data at limited conditions or to handle reactions defying typical rate constant descriptions/fits) [45, 46, 47]. Given these additional complexities encountered in chemical kinetics, many current studies that would be considered the closest analogs to the ATcT are limited in their scope. Some of these studies include, but are not limited to, inverse uncertainty quantification (UQ) of theoretical kinetics parameters based on experimentally determined rate constants, for abstraction and multi well reactions [48, 49, 50], as well as inverse UQ determination of rate constant parameters based on experimentally determined rate constants and/or multi-well reaction observables [51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64]. Additionally, studies employing modeling techniques such as hierarchical comprehensive modeling [65, 66, 67] and static evaluation modeling [44, 68] have been performed. Generally, these studies can be considered less mathematically formal optimizations of rate parameters and use rate constant determinations and/or multi-reaction observables as optimization targets. These studies do not account for uncertainty quantification or cross-correlations between parameters.

From these previous works, it can be concluded that all rate parameter-based approaches rely heavily on rate parameter fitting expressions (i.e. Arrhenius or Troa) which are either inappropriate, ill-constrained, or often times both for many important reactions. Thus, the vast majority

of kinetics informatics studies either infer molecular parameters (within theoretical calculations) based on rate constant determinations or infer rate constant parameters (within semi-empirical rate constant formulas) based on rate constant determinations and/or multi-reaction observables. The MultiScale Informatics (MSI) [69, 70, 71, 72] approach proposed here is unique in that it incorporates both aforementioned aspects, combining theoretical and experimental data in a manner that both unravels data from multi-reaction systems and copes with sparse experimental data sets and/or unconventional kinetic behavior through the incorporation of theory. MSI is a data-driven optimization approach that combines theoretical and carefully selected experimental data across multiple scales, under one mathematical framework, to make highly accurate predictions with quantified uncertainties. Regarding experimental data selection, primarily raw data will be used to capture experimental information content most accurately and effectively instead of exclusively relying on derived data. Additionally, experimental data will be selected from experiments involving small sub-systems with at most a few reactions (to avoid introducing too many weak links [41]), as well as homogeneous environments (to minimize uncertainties due to transport) and dilute conditions (to minimize structural uncertainties [43, 70] in treating kinetics of multi-component, reactive mixtures [45, 46, 73, 74, 75, 76, 77, 78, 79].) Moreover, specific subsystems with overlapping reactions will be selected to grow a network of strongly linked, redundant data (to leverage information from multiple sources and to test the validity of values and their uncertainties). All of these are key elements of the ATcT [41] that when implemented in MSI will allow for the development of a method which can be used to make novel contributions within the combustion community.

1.3 Objectives of this Work

The overarching objectives of this work are as follows:

1. Demonstrate how the MultiScale Informatics (MSI) method can be used to determine physically meaningful optimized kinetics parameters and quantified uncertainties for chosen reaction systems.

2. Demonstrate how MSI can be used to help unravel a web of interconnected rate constants in a complex reaction system.

3. Demonstrate how MSI can resolve discrepancies among data sets as well as theoretical and experimental determinations.

4. Touch on key elements of the MSI approach by implementing it on three kinetics systems. In doing so, demonstrate some key elements for developing analogs to the ATcT for kinetics, including the importance of raw data for quantifying the information content of experimental data, the utility of theoretical kinetics calculations for constraining experimental interpretations and providing an independent data source, and the subtleties of target data selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties. This will also showcase the unique features of MSI as compared to other approaches, as well as demonstrate its validity as a method to be used in future studies.

5. Demonstrate how the results from all studied reaction systems presented can culminate into a complex hydrogen model validated against various experiment types.

The three subsystems that have been explored in detail will be presented in the following sections; a brief description of the motivation behind exploring each system and a summary of the study is also provided.

Chapter 2: Technical Approach

2.1 MSI

The objective of the MSI approach is to identify optimized values and quantified uncertainties for a set of molecular parameters (within theoretical kinetics calculations), rate parameters, and physical model parameters (within simulations of experimental observables) as informed by data from various sources and scales.

In order to fully understand the MSI approach, it is important to understand how data are related to each other across scales, as well as understand the constraints each type of data impose on the model. Specifically, as illustrated in Fig. 1, theoretical kinetics calculations (e.g. TST and/or RRKM-ME [80]), which take active molecular properties (e.g. energies or vibrational frequencies of stationary points on a potential energy surface) as their inputs, yield rate constants for the emergent phenomenological reactions as their outputs. Theoretical kinetics calculations are performed using MESS [81] and/or Variflex [82] software. Kinetic models are then compiled based on these reactions and rate constants from the theoretical kinetics calculations along with any reactions and rate constants which are treated with active rate parameters (instead of molecular parameters within theoretical kinetics calculations). This kinetic model is then used within physical models (e.g. adiabatic/isothermal, isochoric/isobaric reactors), with their own active physical model parameters (e.g. T, P), to predict the observed quantities (e.g. species concentrations, laser absorbance signals) in each experiment and/or other macroscopic quantities of interest. Physical models are simulated using Cantera [83]. Furthermore, as illustrated by dashed arrows in Fig. 1, the constraints imposed by data at any scale can also be projected onto the active parameters via “inverse" uncertainty quantification (UQ) [43], where macroscopic observables impose constraints on reaction rate constants and, likewise, reaction rate constants impose constraints on molecular

properties.

The implementation of the MSI approach can be broken down into two major steps. The first step is to construct an active multi-scale model and to identify a set of targets, $\boldsymbol{y^t}$. This active model consists of a set of active parameters (that will be optimized), $\boldsymbol{x}$, and a structure of physics-based models that make a prediction of the i^{th} target, $f_i(\boldsymbol{x})$, for a given $\boldsymbol{x}$. Active parameters can include active molecular parameters, rate constant parameters, physical model parameters, or any combination of these. Active model parameters are picked by considering the main uncertainty sources in the molecular parameters, kinetic model, as well as experiments. The data to be used as targets will be discussed in greater detail in a following section; target values can include *ab initio* calculations of molecular parameters, experimental rate constant determinations, and macroscopic observables from multi-reaction systems.

The second step of MSI is to impose the constraints from data onto the active model parameters via "inverse" uncertainty quantification (UQ) [43]. Similar to [69, 70, 71], inverse UQ used for the results presented here employs an iterative, uncertainty-weighted, least-squares error minimization.

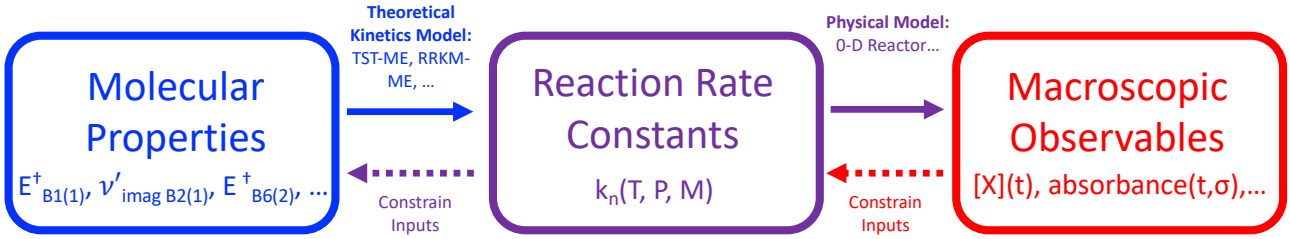


Figure 1: Schematic for the Multi-Scale Informatics approach.

Each iteration of the least-squares error minimization involves construction of a locally linear surrogate model of $\boldsymbol{f}(\boldsymbol{x})$ in the neighborhood of $\tilde{\boldsymbol{x}}$

$$\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{f}(\tilde{\boldsymbol{x}}) + \boldsymbol{S}(\boldsymbol{x} - \tilde{\boldsymbol{x}}) \tag{2.1}$$

where $S_{ij} = (\frac{\partial f_i}{\partial x_j})_{\boldsymbol{x}=\tilde{\boldsymbol{x}}}$; followed by minimization of the uncertainty-weighted least-squares error

$$E(\boldsymbol{x}) = \sum_{i=1} \left(\frac{y_i^{\mathrm{t}} - f_i(\boldsymbol{x})}{z_i} \right)^2 \tag{2.2}$$

to find the optimized values, $\boldsymbol{x}^*$, where y_i^{t} is the i^{th} target value, and z_i is the weighting factor for the i^{th} target, where $z_i = \sigma_i / w_i$ is equal to the uncertainty of the i^{th} target, σ_i, divided by the data set weighting factor, w_i (taken to be $w_i = 1/n^{0.5}$ where n is the number of data points in a data set used from a particular study); followed by setting $\tilde{\boldsymbol{x}}$ to $\boldsymbol{x}^*$ and repeating until converged. Once converged, this nonlinear optimization yields an optimized set of active parameters, $\boldsymbol{x}^*$, and a covariance matrix, $\boldsymbol{\Sigma}$, representing the uncertainties in active parameters and correlations among them as informed by all data (i.e. the joint probability distribution function).

Uncertainties in this work are intended to reflect two standard deviations, though the lack of specification in most studies make such a designation tentative. Ultimately, once sufficient statistics are obtained through the application of this approach to many systems, it can be expected that the results from MSI analysis may permit evaluation of the suitability of the uncertainty assignments for various types of data. Prediction uncertainties were propagated using the covariance matrix. It should also be noted that many other inverse UQ studies with similar numbers of active parameters, e.g. [43, 84, 85], use similar, if not identical, mathematical frameworks for quantifying uncertainty – treating prior and posterior distributions as multivariate normal distributions and using a local first-order surrogate model to construct a covariance matrix.

The previously discussed target values y_i^{t}, can be broken down into four main Target Classes. First, data from *ab initio* calculations of molecular properties belong to Target Class I. Second, data from rate constant determinations belong to Target Class II. Third, macroscopic observables belong to Target class III. Fourth and finally, reported values for the physical model parameters belong to Target class IV. (The mathematical mappings used to implement these individual target classes into the model will be described below.)

Target classes (I) and (IV) serve as regularization terms imposing the prior distributions on all model parameters, meaning the model is sufficiently constrained by Target Classes (I) and (IV) alone and all other data (Target Class II and Target Class III) impose further constraints on the

model. This implies that the prior model is composed of all parameters at their nominal values (*ab initio* calculated values for theoretical kinetics parameters and reported physical model parameters) with uncertainties imposed by Target classes (I) and (IV) alone, while optimized MSI models are composed of all optimized parameters with uncertainties constrained by all Target Classes (I-IV).

2.1.1 Target Class (I): Ab Initio Calculated Values

Including molecular properties from *ab initio* calculations mathematically in the model is quite straightforward:

$$S_{ij} = \delta_{ij} \tag{2.3}$$

where $\delta_{i,j}$ is the Kronecker delta.

2.1.2 Target Class (II): Rate Constant Determinations

Inclusion of rate constant measurements for a reaction n at given $T/P/M$ conditions, $k_{t,n}(T, P, M)$, is treated via a mapping of theoretical kinetics parameters to rate constant predictions, $k_{p,n}(T, P, M)$ (Eq. 2.4). This is done by employing appropriate kinetic theory to calculate the rate constants; for example, transition state theory, master equation simulations, etc.

$$S_{ij} = \frac{\partial \ln k_{p,n}(T_i, P_i, M_i)}{\partial x_j} \tag{2.4}$$

Instead of calculating $k_{t,n}(T, P, M)$ for every T/P/M condition, it was more convenient to use a parametric representation which can be seen in Eq. 2.6.

2.1.3 Target Classes (III) and (IV): Macroscopic Observables and Experimental Conditions

Inclusion of macroscopic observables from combustion experiments is treated via a twofold mapping; by first mapping from theoretical kinetics parameters to rate constant predictions (discussed above), and secondly, mapping from rate constants to combustion predictions, which can

be written for nearly constant temperature systems as shown below in equation 2.5.

$$S_{ij} = \sum_{n} \frac{\partial F_i}{\partial \ln k_{p,n}(T_i, P_i, M_i)} \times \frac{\partial \ln k_{p,n}(T_i, P_i, M_i)}{\partial x_j} \tag{2.5}$$

2.1.4 Parametric Representation of Rate Constants

In practice, rather than calculating $k_{p,n}(T, P, M)$ for every $T/P/M$ condition, it was convenient to represent the $T/P/M$ dependence using some suitable m_n-term parametric representation (Eq. 2.6) of the Chebyshev rate expression (Eq. 2.7).

$$\ln k_{p,n}(T, P, M) = \sum_{m=1}^{m_n} \xi_{n,m}(T, P, M)\hat{k}_{n,m} \tag{2.6}$$

Where $\xi_{n,m}$ and $\hat{k}_{n,m}$ are defined in Eq. 2.10 and Eq. 2.11.

$$\log k(T, P) = \sum_{\tau=1}^{N_\tau} \sum_{\rho=1}^{N_\rho} \alpha_{\tau\rho}\phi_\tau(\tilde{T})\phi_\rho(\tilde{P}) \tag{2.7}$$

In Eq. 2.7 $\alpha_{\tau\rho}$ are the constants defining the rate (represented in a 2-dimensional matrix of shape $N_\tau \times N_\rho$), while $\phi_\tau(\tilde{T})$ and $\phi_\rho(\tilde{P})$ are the Chebyshev polynomial of the first kind of degree ρ or τ evaluated at $\tilde{T}$ or $\tilde{P}$ respectively.

$$\tilde{T} = \frac{2T^{-1} - T_{min}^{-1} - T_{max}^{-1}}{T_{max}^{-1} - T_{min}^{-1}} \tag{2.8}$$

$$\tilde{P} = \frac{2\log P - \log P_{min} - \log P_{max}}{\log P_{max} - \log P_{min}} \tag{2.9}$$

Eq. 2.8 and Eq. 2.9 are the reduced temperatures and pressures which map the ranges (T_{min}, T_{max}) and (P_{min}, P_{max}) to (-1,1). For any reaction being treated with theory parameters, Eq. 2.7 is used to accommodate any functional form the rate constant may take as well as ensure accurate representation over the full temperature and pressure range.

$$\xi_n(T,P,M) = \phi(\tilde{T})_\tau \phi(\tilde{P})_\rho = \{\phi(\tilde{T})_1\phi(\tilde{P})_1, \phi(\tilde{T})_1\phi(\tilde{P})_2, \phi(\tilde{T})_2\phi(\tilde{P})_1, ...\phi(\tilde{T})_{N_\tau}\phi(\tilde{P})_{N_\rho}\} \quad (2.10)$$

$$\hat{k}_n = \alpha_{\tau\rho} = \{\alpha_{11}, \alpha_{12}, \alpha_{21}...\alpha_{N_\tau N_\rho}\} \quad (2.11)$$

In this way equation 2.4 becomes

$$S_{ij} = \frac{\partial \ln k_{p,n}(T_i,P_i,M_i)}{\partial x_j} = \sum_{m=1}^{m_n} \xi_{n,m}(T_i,P_i,M_i)\frac{\partial \hat{k}_{n,m}}{\partial x_j} \quad (2.12)$$

and equation 2.5 becomes the below equation.

$$\begin{aligned} S_{ij} &= \sum_n \frac{\partial F_i}{\partial \ln k_{p,n}(T_i,P_i,M_i)} \times \frac{\partial \ln k_{p,n}(T_i,P_i,M_i)}{\partial x_j} \\ &= \sum_n \frac{\partial F_i}{\partial \ln k_{p,n}} \sum_{m=1}^{m_n} \xi_{n,m}(T_i,P_i,M_i)\frac{\partial \hat{k}_{n,m}}{\partial x_j} \end{aligned} \quad (2.13)$$

2.2 Open Source Software Incorporated Within MSI

2.2.1 Master Equation System Solvers

Two software packages were used for theoretical kinetics calculations, Master Equation System Solver (MESS) [81] and Variflex [82]. Both programs are used to calculate temperature and pressure dependent rate coefficients for complex-forming reactions via solutions of the one-dimensional master equation [81, 82]. MESS and Variflex provide the capability to bridge the gap between the molecular properties block and the reaction rate constants block in Fig. 1. The master equation is a first principles approach to tracking the micro canonical distribution of reacting complexes and calculating phenomenological rate coefficients among the involved particles. Further discussion is beyond the scope of this work, but the reference in this section can provide additional

information.

In this work, MESS and Variflex were used to calculate *a priori* rate constants for certain reactions as well as calculate brute-force sensitivity coefficients (using an automated wrapper [86]). Further detail will be provided regarding this in a later section.

2.2.2 Cantera

Cantera [83] is an open-source suite of tools for problems involving chemical kinetics, thermodynamics, and transport processes. In this work, Cantera is used to bridge the gap between the reaction rate constants and the macroscopic observables (Fig. 1) by solving time dependent chemical kinetic, thermodynamic, and transport calculations. To determine time dependent thermodynamic solutions Cantera relies on the following governing equations,

Mass Conservation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{2.14}$$

Species Conservation

$$\rho \frac{DY_i}{Dt} = w_i - \nabla \cdot (\rho Y_i V_i), \quad \text{i} = 1,...,\text{N} \tag{2.15}$$

Momentum Conservation

$$\rho \frac{D\mathbf{v}}{Dt} = -\nabla \cdot \mathbf{P} + \rho \sum_{i=1}^{N} Y_i \mathbf{f}_i \tag{2.16}$$

Energy Conservation

$$\rho \frac{De}{Dt} = -\nabla \cdot \mathbf{q} - \mathbf{P} : (\nabla \mathbf{v}) + \rho \sum_{i=1}^{N} Y_i \mathbf{f}_i \cdot V_i \tag{2.17}$$

where ρ is the mass density, $\mathbf{v}$ is the mass-averaged velocity of the gaseous mixture, Y_i is the

mass fraction of species i, w_i is the mass rate of production of species i per unit volume, V_i is the diffusion velocity of species i, N is the total number of chemical species, $\mathbf{P}$ is the pressure tensor, $\mathbf{f}_i$ is the external body force per unit mass of species i, e is the specific internal energy of the gaseous mixture, and $\mathbf{q}$ is the heat flux vector.

Equations 2.14-2.17 are dependent on one another and therefore must be solved concurrently. Even for a simple system, solving this system of equations analytically without any simplifying assumptions (in three dimensions) would be infeasible– likely requiring a super computing cluster. However, there is sill ample information that can be gathered from running simulations with simplifying assumptions that use either 0-dimensional or 1-dimensional reactor/flame approximations, which can be run on a personal computer. Depending on the type of simulation, (i.e. ideal gas constant pressure reactor, ideal gas constant volume reactor, etc.) simplifying assumptions (i.e. constant pressure, constant volume, adiabatic, etc.) are used to reduce complexity of equations 2.14-2.17, making them more computationally tractable. Additional information regarding simplifying assumptions applied for specific simulations can be found in Cantera's [83] documentation.

Chapter 3: MSI: Software Implementation

3.1 Overview

The implementation of the Multi-Scale Informatics method required custom software to be written in Python [87], which utilizes the open source software packages discussed in the previous chapter. An overview of the Python software packages that were built (MSI Theory and the MSI Library) to implement the MSI method are discussed in this chapter.

The general flow of information through the sub-modules of the MSI Theory software is shown in Fig. 3, and Fig. 10 shows the general flow of information through the MSI Library software.

3.2 MSI Theory Module

The MSI Theory Module handles all *a priori* theoretical kinetics calculations for reactions that are treated with theoretical molecular parameters in the MSI framework. The MSI Theory Module is executed by running a "main" python file. The "main" file, also known as the input file, defines the path for the MESS or Variflex file that will be executed, the parameters in the MESS or Variflex file that will be perturbed, the perturbation percent, the temperature and pressure list, the rate constant channels that should be extracted, and the shape of the Chebyshev polynomials that are returned. An example input file for the MSI Theory Module can be seen in Fig. 1. The MSI Theory Module contains various sub-modules that function to execute the MESS or Variflex software, perturb parameters, fit Chebyshev polynomials to rate constant data, and calculate sensitivity coefficients. The "main" file for the MSI Theory Module (Fig. 1) is executed, and the variables defined within are passed to the appropriate sub-modules. The sub-modules are described below in the order they are executed.

```
import postprocessor as MSI
import os
import numpy as np

# Initialize MSI.PAPR_MESS:
# -- the first arguement is the file path for the nominal PAPR-MESS input files
# -- the second arguement is the name for the unperturbed (nominal) input file
# -- the third arguement is a dictionary specifying the nominal condition (to allow optimization starting from nonzero perturbations),
#    the value specifies temperature list, pressure list, and nominal perturbation
# -- the fourth arguement is a dictionary with elements specifying the perturbations (sensitivity analysis are based on the nominal conditions defined above),
#    the keys specify names of perturbation runs and values specify temperature, pressure, and perturbation
#    the fifth arguement is a list indicating the channels of interest

temperature_list = np.arange(200,700,10).tolist() + np.arange(700,2500,100).tolist() # K
pressure_list = [[1,2,3,4,5], ['[torr]']] # Torr, will be converted into atm internally
pertubation_percent = 0.05 # percentage of perturbation

channels = ['P1->P2'] # channel-specific rate constants of interest
nominal_MESS_input_path = os.getcwd() + '/ho2+ho2/'
nominal_MESS_input = 'HO2+HO2.inp'


model = MSI.PAPR_MESS(nominal_MESS_input_path, nominal_MESS_input,
                                        {'B1_Energy_1':[temperature_list, pressure_list, 0.00]},
                                        {'B1_Energy_1':[temperature_list, pressure_list, pertubation_percent * 349.759],
                                         'B1_Frequencies_1':[temperature_list, pressure_list, pertubation_percent],
                                         'B1_ImaginaryFrequency_1': [temperature_list, pressure_list, pertubation_percent],
                                         'B1_SymmetryFactor_1': [temperature_list, pressure_list, pertubation_percent],

                                         'P1_Energy_1':[temperature_list, pressure_list, pertubation_percent * 349.759],
                                         'P1_Frequencies_1':[temperature_list, pressure_list, pertubation_percent],
                                         'P1_SymmetryFactor_1': [temperature_list, pressure_list, pertubation_percent],

                                         'W0_Energy_1':[temperature_list, pressure_list, pertubation_percent * 349.759],
                                         'W0_Frequencies_1':[temperature_list, pressure_list, pertubation_percent],
                                         'W0_SymmetryFactor_1': [temperature_list, pressure_list, pertubation_percent]},

                                         'P2_Energy_1':[temperature_list, pressure_list, pertubation_percent * 349.759],
                                         'P2_Frequencies_1':[temperature_list, pressure_list, pertubation_percent],
                                         'P2_SymmetryFactor_1': [temperature_list, pressure_list, pertubation_percent]},
                                          channels, abstraction=False)

# fit rate constants
model.Run() # execute PAPR-MESS perturbations

# define shape of Chebyshev polynomial
n_P = 2
n_T = 8
same_line_result = True
aggregated_sens = False

# fit rates to Chebyshev polynomial
model.fit_Cheb_rates(n_P=n_P, n_T=n_T, P_min=0.0001, P_max=10, T_min=200.0, T_max=3000.0, same_line_result=same_line_result)

# calculate sensitivity coefficients
model.Cheb_sens_coeff(same_line_result=same_line_result, aggregated_sens=aggregated_sens)
```

Figure 1: Example MSI Theory input file.

3.2.1 Pre-processor

The Pre-processor sub-module removes any unnecessary content from a MESS or Variflex input file such as comments or extraneous characters. An example snippet of a MESS input file can be seen in Fig. 2. The nominal MESS or Variflex input file contains information about a specific potential energy surface (PES) and the theoretical molecular parameters used to calculate rate constant channel(s). The Pre-processor sub-module loads the nominal MESS or Variflex input file (from the path provided in the MSI Theory input file) and the loaded file is cleaned by removing extraneous spaces, special characters, and undeclared comments. The cleaning process removes anything from the file that could potentially cause an error when the file is executed by the MESS or Variflex software in a later sub-module. After the file is cleaned, a new version of the file is

saved.

```
EnergyStepOverTemperature               .2
ExcessEnergyOverTemperature             30
ModelEnergyLimit[kcal/mol]              400
CalculationMethod                       direct
WellCutoff                              10
ChemicalEigenvalueMax                   0.0001
MicroRateOutput                         ke.out
MicroEnerMin[kcal/mol]                  -10.
MicroEnerMax[kcal/mol]                  20.
MicroEnerStep[kcal/mol]                 1.0
Model
  EnergyRelaxation
    Exponential
      Factor[1/cm]                       200
      Power                             .85
      ExponentCutoff                    15
    End
  CollisionFrequency
    LennardJones
      Epsilons[1/cm]                    70.  500. !
      Sigmas[angstrom]                  3.3  10.0 !
      Masses[amu]                       28. 66.  !
    End
  Well     W1a   # 3HO2...HO2
    Species
    RRHO
      Geometry[angstrom]                          6
       H                 -2.645806    0.582429    0.031246
       O                 -1.723457    0.496562   -0.270521
       O                 -1.272925   -0.587257    0.342651
       H                  0.518694   -0.639265   -0.113768
       O                  1.473921   -0.532521   -0.312297
       O                  1.788351    0.630321    0.250482
      Core                RigidRotor
        SymmetryFactor                          0.2
      End
      Frequencies[1/cm]                         12
  23.75
  54.05      108.88      205.55      258.39      561.35
       1149.02     1151.48     1445.06     1521.27     3573.95
       3667.68
      ZeroEnergy[kcal/mol]                      -4.22
      ElectronicLevels[1/cm]                    1
        0       3
    End
```

Figure 2: Example snippet of a MESS input file.

3.2.2 Class Generator

The cleaned input file is passed to the Class Generator sub-module. This sub-module reads the cleaned MESS or Variflex input file line by line and builds a dictionary of nested dictionaries to store the file's contents and be accessed by later sub-modules. A dictionary referred to in the context of this chapter is a Python data type more generally known as an associative array. A

dictionary consists of a collection of key-value pairs, and each key-value pair maps the key to its associated value.

3.2.3 File Writer and Value Perturbations

The File Writer and Value Perturbations sub-module takes the dictionary from the Class Generator and creates perturbed MESS or Variflex input files, which are used to calculate perturbed rate constants and sensitivity coefficients. The sub-module uses the dictionary (containing the contents of the original MESS or Variflex input file) and generates various copies. In each copy, one of the molecular parameters (listed in the MSI Theory input file), is perturbed by the perturbation percent (defined in the MSI theory input file). These perturbed dictionaries facilitate sensitivity analysis in a later sub-module. The contents of the original dictionary and each of the perturbed dictionaries are written out as a MESS or Variflex input file and stored in individually created directories. The directories are labeled as either nominal (for the original dictionary/file contents) or have the name of the perturbed parameter (i.e. frequency, energy, symmetry factor, etc.). The file paths for these directories are passed to the next sub-module.

3.2.4 MESS or Variflex Executor

The MESS or Variflex Executor sub-module executes the MESS or Variflex input files written in the previous sub-module. The MESS or Variflex Executor loops over each of the directory paths and executes the MESS or Variflex input file contained within it; the input files are executed using the MESS or Variflex open source software. The MESS or Variflex software calculates the temperature and pressure dependent rate constant(s) from the information in the input file and generates an output file containing the calculated rate constants(s), saving it in the same directory. The next sub-module processes the output files.

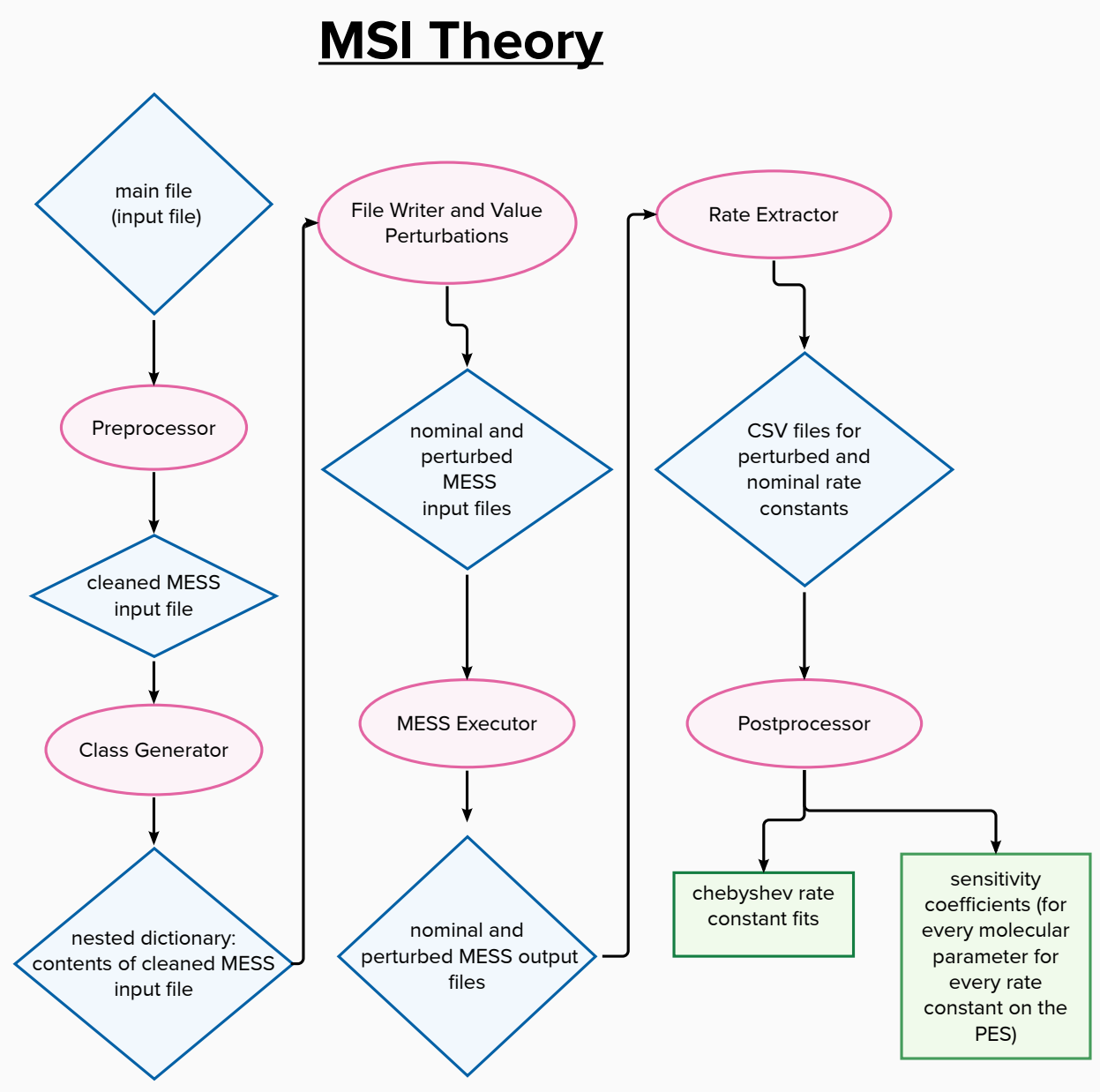


Figure 3: MSI Theory software flow diagram.

3.2.5 Rate Extractor

The Rate Extractor sub-module processes the output files generated in the previous sub-module and extracts the rate data for the pressure and temperature dependent rate constant(s). The output files need to be parsed to extract pressure and temperature dependent rate constant values. The output files may contain rate constants for multiple channels depending on the potential energy

surface (PES) in the original MESS or Variflex input file and the channels listed in the MSI theory input file. The Rate Extractor sub-module loops through the file directories, loads the output files, extracts the necessary rate constant determinations, and writes this information to csv files. Each parsed output file generates a different csv file that is saved in the same directory.

3.2.6 Post-processor

The Post-processor sub-module loops through the directories containing the csv files generated in the previous sub-module and loads the rate data contained within them. The rate data from each parsed csv file is fit to a 2-dimensional Chebyshev polynomial of the first kind, where the coefficients are represented in a 2-dimensional matrix of shape $N_\tau \times N_\rho$. These Chebyshev fits are written to a text file in the nominal directory. The Chebyshev fit(s) for the nominal rate constant(s) and for the perturbed rate constants are used to calculate sensitivity coefficients. The sensitivity coefficients of a given parameter are calculated by finding the difference between the coefficient matrix used to represent the perturbed rate constant fit and the coefficient matrix for the nominal fit. Next, the difference is normalized by the original perturbation value. Sensitivity coefficients are calculated for each parameter and stored in a text file in the main directory. The nominal Chebyshev fit(s) and the sensitivity coefficients (generated for various parameters on the PES) are manually extracted from the text files and used by the main MSI Library.

3.3 MSI Library Overview

The MSI Library contains the source code responsible for the MSI framework and is separated into different modules and sub-modules. Fig. 10 shows the general flow of information through the modules and sub-modules of the MSI Library and can be used to help visualize the discussion in the following section. Throughout this section, the MSI Library may be referred to as MSI or the MSI framework.

```
########
# input file for MSI

working_dir='MSI/data/branching_reaction_study'
iterations=3

begin_yaml_list
['Hong_0.yaml']
['Hong_2.yaml']
'Hong_3.yaml']
['Hong_4.yaml','Hong_4_abs.yaml']
['Hong_1.yaml']
['Troe_4.yaml','Troe_4_abs.yaml']
['Troe_5.yaml','Troe_5_abs.yaml']
end_yaml


begin_model_list
FFCM1.cti
end_model_list

begin_reaction_uncertainty_list
FFCM1_reaction_uncertainties.csv
end_input

begin_master_equation_cti
FFCM1_master_equation_reactions.cti
end_master_equation_cti

begin_master_equation_reactions_list
['H2O2 + OH <=> H2O + HO2',
'2 HO2 <=> H2O2 + O2']
end_master_equation_reactions_list

begin_Chebyshev sensitivity dictionary
{'H2O2 + OH <=> H2O + HO2':[np.array([[-9.24839173e-01,  0.00000000e+00],
       [ 5.70742217e-01,  0.00000000e+00],
       [ 1.89660889e-01,  0.00000000e+00],
       [ 1.43243609e-02,  0.00000000e+00]]),
 np.array([[-3.26819031e+00,  0.00000000e+00],
       [-2.22443669e+00,  0.00000000e+00],
       [-1.11884088e+00,  0.00000000e+00],
       [-4.13092615e-01,  0.00000000e+00]])],
'2 HO2 <=> H2O2 + O2': [np.array([[-5.71717406e-02, -6.99503310e-05,  9.48724272e-05],
        [-1.09795954e-01, -1.37136873e-04,  1.22843294e-04],
        [-9.37120077e-02, -1.18813498e-04,  2.12595928e-05],
        [-7.54332546e-02, -7.43460606e-06, -1.39078738e-05],]),
 np.array([[-6.03208106e-01, -4.85002652e-02,  2.19384705e-02],
        [ 8.63989861e-01,  1.39579175e-02,  1.89981508e-02],
        [-2.75896886e-01,  7.09071614e-02, -7.59035076e-02],
        [-3.01833588e-02, -2.79201341e-02,  1.75491068e-02],])]}
end_Chebyshev sensitivity dictionary

begin_molecular_parameter_uncertainty
molecular_parameter_uncertainty.csv
end_molecular_parameter_uncertainty

rate_constant_target_value
rate_constant_target_values.csv
end_rate_constant_target_value
```

Figure 4: Example MSI input file.

```
units(length='cm', time='s', quantity='mol', act_energy='cal/mol')
ideal_gas(name='gas',
          elements="O H C N Ar He",
          species="""H2      H       O2       OH    O""",
          reactions='all',
          transport='Mix',
          initial_state=state(temperature=300.0, pressure=OneAtm))
#-------------------------------------------------------------------------------
# Species data
#-------------------------------------------------------------------------------
        note=u'tpis78')
species(name=u'O',
        atoms='O:1',
        thermo=(NASA([200.00, 1000.00],
                     [ 3.16826710E+00, -3.27931884E-03,  6.64306396E-06,
                      -6.12806624E-09,  2.11265971E-12,  2.91222592E+04,
                       2.05193346E+00]),
                NASA([1000.00, 6000.00],
                     [ 2.54363697E+00, -2.73162486E-05, -4.19029520E-09,
                       4.95481845E-12, -4.79553694E-16,  2.92260120E+04,
                       4.92229457E+00])),
        transport=gas_transport(geom='atom',
                                diam=2.75,
                                well_depth=80.0),
        note=u'L1/90')
species(name=u'O2',
        atoms='O:2',
        thermo=(NASA([200.00, 1000.00],
                     [ 3.78245636E+00, -2.99673416E-03,  9.84730201E-06,
                      -9.68129509E-09,  3.24372837E-12, -1.06394356E+03,
                       3.65767573E+00]),
                NASA([1000.00, 6000.00],
                     [ 3.66096065E+00,  6.56365811E-04, -1.41149627E-07,
                       2.05797935E-11, -1.29913436E-15, -1.21597718E+03,
                       3.41536279E+00])),
        transport=gas_transport(geom='linear',
                                diam=3.458,
                                well_depth=107.4,
                                polar=1.6,
                                rot_relax=3.8),
        note=u'RUS89')
species(name=u'OH',
        atoms='H:1 O:1',
        thermo=(NASA([200.00, 1000.00],
                     [ 3.99198424E+00, -2.40106655E-03,  4.61664033E-06,
                      -3.87916306E-09,  1.36319502E-12,  3.36889836E+03,
                      -1.03998477E-01]),
                NASA([1000.00, 6000.00],
                     [ 2.83853033E+00,  1.10741289E-03, -2.94000209E-07,
                       4.20698729E-11, -2.42289890E-15,  3.69780808E+03,
                       5.84494652E+00])),
        transport=gas_transport(geom='linear',
                                diam=2.75,
                                well_depth=80.0),
        note=u'IU3/03')
species(name=u'H',
        atoms='H:1',
        thermo=(NASA([200.00, 1000.00],
                     [ 2.50000000E+00,  0.00000000E+00,  0.00000000E+00,
                       0.00000000E+00,  0.00000000E+00,  2.54736600E+04,
                      -4.46682850E-01]),
                NASA([1000.00, 6000.00],
                     [ 2.50000000E+00,  0.00000000E+00,  0.00000000E+00,
                       0.00000000E+00,  0.00000000E+00,  2.54736600E+04,
                      -4.46682850E-01])),
        transport=gas_transport(geom='atom',
                                diam=2.05,
                                well_depth=145.0),
        note=u'L6/94')
#-------------------------------------------------------------------------------
# Reaction data
#-------------------------------------------------------------------------------
# Reaction 1
reaction('H + O2 <=> O + OH', [9.841000e+13, 0.0, 15310.0])
```

Figure 5: Example snippet of a cti file.

```yaml
#PART ONE
file-author:
    name: Carly LaGrotta
    ORCID:
file-version: (1, 0)
reference:
    doi: 10.1016/j.proci.2012.06.108
    authors:
        - name: Zekai Hong
          ORCID:
        - name: Subith S. Vasu
          ORCID:
        - name: David F. Davidson
          ORCID:
        - name: Ronald K. Hanson
    journal: J. Phys. Chem
    year: 2010
    volume: 114
    pages: 5520-5525
    detail: Experimental Study of the rate of OH+HO2->H2O + O2 At high temperatures using reverse reaction rates
#PART TWO
experiment-type: species profile
apparatus:
    kind: batch reactor
    institution: Stanford University
    facility: stainless steel shock tube
    inner-diameter: &diam
        value: 14.13
        units: cm
#PART THREE
common-properties:
    pressure: &pres
        value: 1.909
        units: atm
        relative-uncertainty: .02
    temperature: &temp
        value: 1398
        units: K
        relative-uncertainty: .01
    time: &time
        initial-time: &initialtime
          value: 0
          units: s
        final-time: &finaltime
          value: .0002
          units: s
    composition: &comp
        - species: H2O
          mole-fraction: 0.001234
          relative-uncertainty: .05
        - species: H2O2
          mole-fraction: 0.00254
          relative-uncertainty: .05
        - species: O2
          mole-fraction: 0.00062
          relative-uncertainty: .05
        - species: Ar
          mole-fraction: 0.995606
          relative-uncertainty: .05
    time-shift: &timeshift
        value: 0
        units: s
        absolute-uncertainty:
            value: 0.0000000001
            units: ms
    assumptions: &assumptions
        thermal-boundary: adiabatic #vs. isothermal
        mechanical-boundary: constant pressure #vs. constant volume
        equation-of-state: ideal gas
#PART FOUR
datapoints:
  mole-fraction:
    - csvfile:
      targets:
        - name:
          species:
          absolute-uncertainty:
          relative-uncertainty:
  absorbance:
     - csvfile:
       name:
       wavelength:
         value:
         unit:
       absolute-uncertainty:
       relative-uncertainty:
  concentration:
    - csvfile: 'MSI/data/test_data/hong_oh_0_time_dependent.csv'
      targets:
       - name: OH
         species: OH
         absolute-uncertainty: 10
         relative-uncertainty: .05
    - csvfile: 'MSI/data/test_data/hong_h2o_0_time_dependent.csv'
      targets:
       - name: H2O
         species: H2O
         absolute-uncertainty: 14
         relative-uncertainty: .05
```

Figure 6: Example YAML file for a batch reactor experiment.

3.4 Input File

To run an optimization using the MSI Library, an input file containing certain pieces of information must be constructed. The input file is organized into two parts, necessary and optional. An example of an MSI library input file can be seen in Fig. 4.

Necessary:

- cti file path: The cti file is the Cantera input file. It is the kinetics model containing a list of reactions, associated rate constants, thermodynamic data, transport data, and the kinetic parameters that represent the rate constants for each reaction. A simple example of this file can be seen in Fig. 5. This file should also contain the reactions being treated with theoretical molecular parameters, and the code will identify the reaction strings and replace them with the appropriate theoretical molecular parameter rate constant representations in the cti Core.

- list of YAML file paths: YAML [88] is a data serialization format often used to create configuration files within a programming language. It should be noted the YAML structure for MSI (described below) drew on inspiration from that of the structure presented in Pyked [89]. For the MSI software, the YAML format was extended to other experiment types not previously included in the Pyked format. YAML files for MSI are designed to hold four main pieces of information:

    1. The source of the experimental data.
    2. The information about the type experiment being run (i.e. batch-reactor, jet stirred reactor, etc.).
    3. Additional information needed to simulate the experiment using a physical model in Cantera.
    4. The original experimental data to compare the simulation against. Generally, a YAML file contains all the information about a single experiment.

An example of a YAML file can be seen in Fig. 6. This example contains information about an experiment conducted in a batch reactor. The first part of the YAML file contains information detailing the author of the YAML file as well as the source of the experimental data (i.e. the research publication from which the experimental data is being extracted). The second and third parts of the file contain information regarding the type of experiment that was conducted and information about the apparatus used to conduct the experiment. Part three of the file contains common properties of the experiment along with their associated uncertainties. For a batch reactor, these properties consist of initial pressure, initial temperature, initial time, final time, initial species composition, time shift, and governing assumptions. Governing assumptions describe the conditions under which the experiment was conducted and what equation of state would be suitable for modeling it. The fourth and final part of the YAML file contains paths to (a) csv file(s), which store the real experimental data and the associated uncertainties. For a batch reactor, this data can fall into three categories: mole fractions profiles of species, concentration profiles of species, or absorbance profiles of species. The YAML file in Fig. 6 is a specific template created for a batch reactor physical model. For each physical model implemented within MSI, a slightly different structure for the YAML file was created. All YAML files have the same general structure, but vary slightly based on the specific experiment type with which they are associated. It should be noted that all uncertainties in the YAML file are natural log uncertainties. For example, an uncertainty of a factor of two would be input as 0.693 because ln(2) = 0.693.

- rate constant uncertainty csv path: This file is in csv format and lists all of the reactions in the cti file and the uncertainties associated with the Arrhenius kinetics parameters used to represent them. If a reaction is being treated with theoretical molecular parameters (as discussed later in this chapter), the reaction can be listed in this file and will be removed later by the software. An example of what this file looks like can be seen in Fig 8. It should be noted that all uncertainties for pre-exponential factors are natural log uncertainties, uncertainties for temperature exponent are absolute unit-less quantities, and uncertainties for

activation energy are absolute quantities in units of (1/K).

Reaction	Temperature	Pressure	Mixture	k	ln_unc_k	W
2 OH <=> H2O + O	360	1	Ar:0.98, O2:0.01,H2:0.01	7.71E+11	0.6931472	0.3779645
2 OH <=> H2O + O	340	1	Ar:0.98, O2:0.01,H2:0.01	7.77E+11	0.6931472	0.3779645
2 OH <=> H2O + O	320	1	Ar:0.98, O2:0.01,H2:0.01	8.07E+11	0.6931472	0.3779645
HO2 + OH <=>*	1357	1	Ar:0.98, O2:0.01,H2:0.01	4.82E+13	1.4255151	1.00E-10
HO2 + OH <=>*	1533	1	Ar:0.98, O2:0.01,H2:0.01	2.83E+13	1.4255151	1.00E-10
HO2 + OH <=>*	1511	1	Ar:0.98, O2:0.01,H2:0.01	3.61E+13	1.4255151	1.00E-10
2 HO2 <=> H2O2 + O2	991.9995238	1	Ar:0.98, O2:0.01,H2:0.01	1.09E+12	0.6931472	0.0001
2 HO2 <=> H2O2 + O2	1028.533579	1	Ar:0.98, O2:0.01,H2:0.01	1.30E+12	0.6931472	0.0001
2 HO2 <=> H2O2 + O2	1032.645007	1	Ar:0.98, O2:0.01,H2:0.01	1.50E+12	0.6931472	0.0001
2 HO2 <=> H2O2 + O2	1050.491893	1	Ar:0.98, O2:0.01,H2:0.01	1.30E+12	0.6931472	0.0001

Figure 7: Example snippet of a rate constant target csv file. The “Reaction” column contains the reaction string for the rate constant. The “Temperature”, “Pressure”, and “Mixture” columns contain the temperature (in K), pressure (in atm), and mixture (in mole fraction) for the rate constant target determination. The “k” column contains the value (in cm, mol, and s) for the rate constant target. The “ln_unc_k” column contains the value for the natural log uncertainty of the rate constant target. The “W” column contains the value for the weighting factor of the rate constant target.

Optional:

- Master equation reaction cti: This is a second cti file that contains only the reactions that

are treated using theoretical molecular parameters as active parameters and their associated Chebyshev rate constant fits (calculated in the MSI Theory code).

- List of master equation reactions: This is a list of the strings representing the master equation reactions. This list is used throughout the software to identify the reactions being treated with theoretical molecular parameters.

- Chebyshev sensitivity dictionary: This is a python structured dictionary where the keys consist of the rate constant strings for reactions represented with theoretical kinetics parameters and the values are the sensitivity coefficients associated with the perturbed theoretical kinetics parameters, (calculated in the MSI Theory code).

- Molecular parameter uncertainty csv path: This file is in csv format and lists the uncertainties associated with the theoretical molecular parameters for specific reactions. An example of this file structure can be seen in Fig. 9. It should be noted that all uncertainties in the molecular parameter uncertainty csv file are natural log uncertainties with the exception of energy terms and the scaling factor ($n_{\Delta E_d}$) for the parameters describing a temperature-dependent average energy per down collision, $\Delta E_d = AT^n$, in the exponential-down model. These are absolute uncertainties in units of kcal/mol and a unit-less quantity respectively.

- Rate constant target value csv path: This file is in csv format and lists any experimental rate constant targets being used along with their associated temperature, pressure, and uncertainty. An example of this file format can be seen in Fig. 7. It should be noted all uncertainties in the rate constant target value csv file are natural log uncertainties.

3.5 Input File Parser

The first module of the MSI Library is the Input File Parser. This portion of the code reads in the input file and the information contained within (listed above). The Input File Parser stores

Reaction	A	n	Ea
H + O2 <=> O + OH	0.69314718	1	1000
O + H2 <=> H + OH	0.69314718	2	2000

Figure 8: Example of a reaction uncertainty csv file for rate constant definitions in the cti file which take the form $k = AT^n e^{(-Ea/RT)}$. The "Reaction" column contains a rate constant reaction string, the "A" column contains the natural log uncertainty for the the pre-exponential factor, the "n" column contains the absolute uncertainty for the temperature exponent (a unit less quantity), and the "Ea" column contains the absolute uncertainty for the activation energy divided by the gas constant (in units of 1/T).

the information contained in the input file as the appropriate variable types internally to be used at various points throughout the code.

3.6 cti Core

The cti Core sub-module reads, writes, and updates the cti files associated with the optimization and establishes the simulation processor. On the first iteration of the optimization, the cti Core sub-module creates a copy of the original cti file. The copy of the file is renamed and saved with the original file name with "_updated" appended. This is the cti file that is used throughout the MSI framework and updated each iteration. Next, the cti Core sub-module loads the master equation reaction cti file and puts all of the rate constant expressions contained within it at the end of the "_updated" cti file. All reactions treated with theoretical molecular parameters are grouped together at the end of the reactions list in the "_updated" cti file. Next, the cti Core sub-module uses "_updated" cti file to establish the processor, which is used throughout the optimization. The

H2O2 + OH <=> H2O + HO2	2 HO2 <=> H2O2 + O2	HO2 + OH <=> H2O + O2	CH3 + HO2 <=> CH4 + O2	CH3 + HO2 <=> CH3O + OH
3	0.045	0.7	1	0.262
0.1	2.00E-01	2	0.1	1
0.1	1.35E-01	0.1	1	
0.18	0.2	0.41		
1.00E-06	1.61E+00	0.18		
	1.61E+00	1.00E-06		
	6.93E-01	2		
	1.10E+00	0.1		
	2.00E-01	0.41		
	1.00E-06	0.18		
	1.80E-01	1.00E-06		
	4.50E-02	1.00E-06		
	1.00E-06	0.01		
	4.50E-02	1.00E-06		
	1.35E-01			
	2.00E-01			
	1.10E+00			
	1.61E+00			
	2.00E-01			
	1.35E-01			
	2.00E-01			
	1.00E-10			
	2.00E-01			
	2.00E-01			
	7.00E-02			
	4.50E-02			

Figure 9: Example snippet of a molecular parameter uncertainty csv, where the order of the uncertainties listed under each reaction corresponds to the order sensitivities are listed in the Chebyshev sensitivity dictionary.

processor is established by loading the cti file into Cantera and creating an internal solution object. The internal solution object is a Python class for chemically-reacting solutions, and this Python class contains all the information in the cti file to compute physical models.

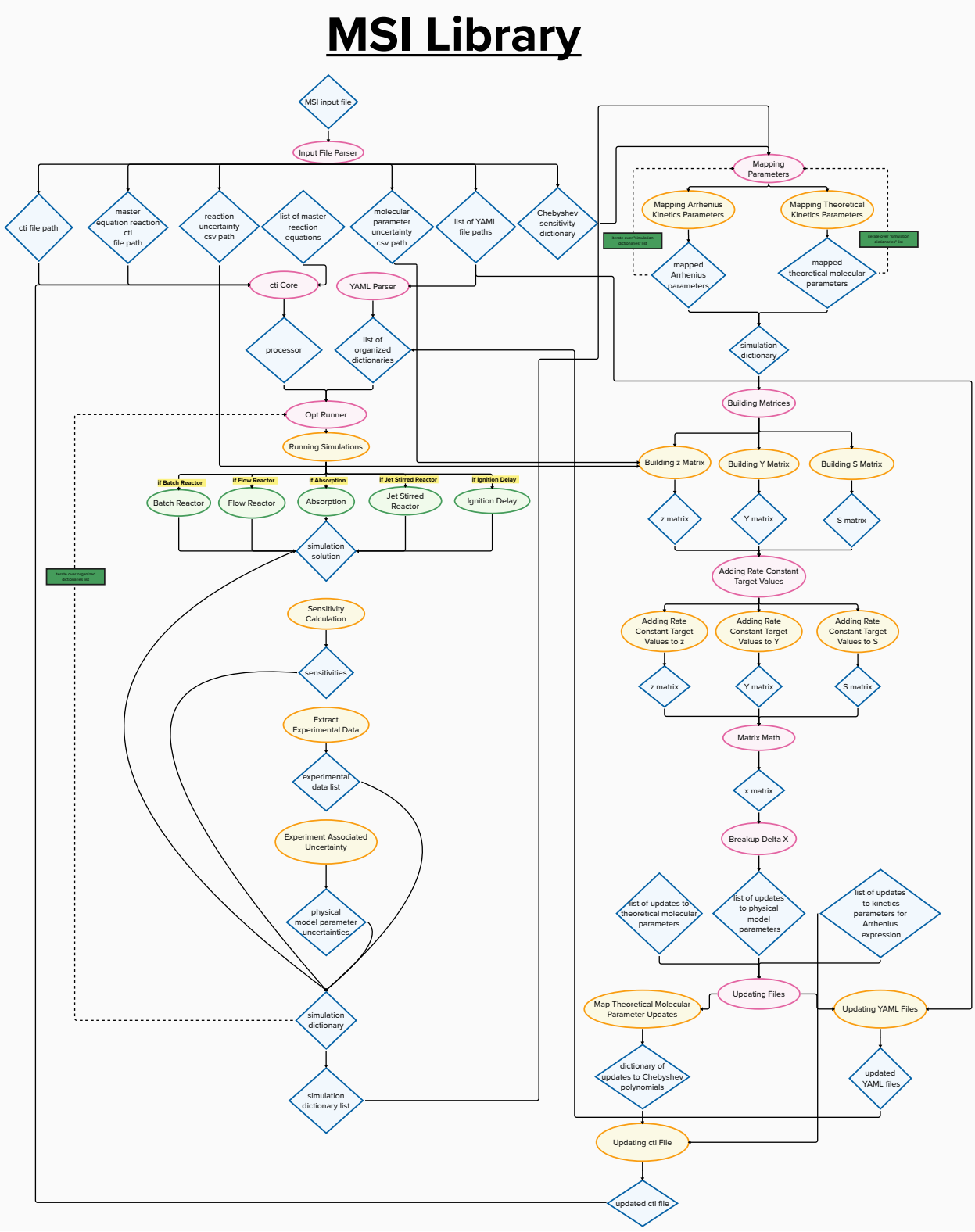


Figure 10: MSI Library software flow diagram. Dashed lines in the diagram represent loops.

3.7 YAML Parser

The YAML Parser sub-module takes the list of the paths of YAML files as input. First, the YAML Parser loads each of the files associated with the paths in the list into a python object – given the complex structure of the YAML files, the loaded object is a series of nested lists and dictionaries. Although generally the YAML file structure for different physicals models is slightly varied, there is one constant portion of the file structure that holds the information about the "experiment type" and "simulation type." The YAML Parser extracts the information associated with the experiment type and simulation type from each object and uses it to pass the object into the appropriate parser function. Each different physical module, batch reactor, jet stirred reactor, etc. has its own parser. The parser functions take the nested object structures for each specific type of YAML file and use a series of functions and methods to untangle the nested object and output the information into an organized dictionary with keys and values for each piece of information originally contained in the YAML file. These organized dictionary objects are returned in a list in the original order the files paths appeared in the list.

3.8 Opt Runner

The next module of the MSI Library is the Opt Runner, which takes the processor created in the cti Core and the list of dictionaries created in the YAML parser as inputs. The Opt Runner loops over the list of input dictionaries and for each dictionary executes a series of sub-modules. The Opt Runner runs physical models, calculates kinetic sensitivities, calculates physical model parameter sensitivities, extracts experimental data from the csv files, and compiles uncertainties for all parameters associated with the original experiments. After each iteration of the opt runner, consisting of execution of all the internal sub-modules, a simulation dictionary is constructed. This object contains all the information returned from the sub-modules within the Opt Runner for a specific input dictionary (which refers to a specific experiment). A list of simulation dictionaries (the same length as the original number YAML files) is returned from the Opt Runner. The sub-

modules internal to the Opt Runner are described below in the order they are executed.

3.8.1 Running Simulations

The Running Simulations sub-module takes the processor created in the cti Core and the list of dictionaries created in the YAML Parser and, using them as inputs, simulates the experiments described in each of the individual YAML files using the appropriate physical models.

The physical model functions implemented in the MSI Library are a batch reactor, a temperature dependent flow reactor, and a jet stirred reactor. Additionally, an absorbance physical model and an ignition delay physical model were written as additions to the batch reactor physical model. The implementation of these physical models (i.e. the portion of the software which solves the time dependent governing equations, Eq. 2.14, 2.15, 2.16, and 2.17) is facilitated by Cantera. However, additional framework was implemented around the Cantera functions to produce organized simulation results and intermediate solutions. Each of these physical models is discussed in detail below.

It should also be noted that the individual physical models contained within the Running Simulations sub-module were all designed to be run within the MSI optimization framework, or independent of the optimization framework. Allowing the end user to run the physical models independent of the MSI framework provides an additional compact tool for simulating physical models.

The batch reactor physical model implemented using Cantera's ideal gas reactor module simulates how a homogeneous chemical composition changes over time given an initial temperature, pressure, and species concentrations. Two versions of the batch reactor can be simulated, constant volume or constant pressure. The batch reactor physical model returns a simulation solution in the form of a Pandas [90] data frame, which contains a time history of temperature, pressure, and mole fraction of every chemical species defined in the original processor. Pandas is a software library written for the Python programming language for data manipulation and analysis, and a data frame is a structure for storing data much like an array, except it provides column labels and row indices among other advantages.

```yaml
Absorption-coefficients:
  - species: H2O2
    wave-lengths:
      - value: 215
        unit: nm
        functional-form: A
        source: Kappel et al.
        parameter-one:
          value: 96000
          units: 'K*cm^2/mol,'
          absolute-uncertainty:
            value: .7
            units: 'K*cm^2/mol,'
        parameter-two:
          value: -19.0
          units: 'cm^2/mol'
          absolute-uncertainty:
            value: .3
            units: 'cm^2/mol'
  - species: HO2
    wave-lengths:
      - value: 215
        unit: nm
        functional-form: B
        source: Kappel et al.
        parameter-one:
          value: 1155000
          units: 'cm^2/mol'
          absolute-uncertainty:
            value: .7
            units: 'cm^2/mol'
        parameter-two:
          value: 1299
          units: 'K'
          absolute-uncertainty:
            value: .3
            units: 'K'
```

Figure 11: Example of a YAML file for absorbance.

An ignition delay physical model function was also implemented using the batch reactor physical model as a base. An ignition delay time is defined as the time needed for a mixture of a fuel and oxidizer to react at a certain temperature and pressure and is a popular experimental method for studying properties of different fuels. The metric for measuring ignition delay time may vary from experiment to experiment, and different definitions needed to be implemented into the physical model to simulate various types of experimental measurements. The definitions for ignition delay times implemented in the physical model include:

- time at which the derivative of pressure with respect to time is at a maximum
- time at which the derivative of temperature with respect to time is at a maximum
- time at which the derivative of a specific species with respect to time is at a maximum
- time at which pressure is at a maximum
- time at which temperature is at a maximum
- time at which a specific species is at its maximum
- time at which temperature is equal to a specified value
- time at which pressure is equal to a specified value
- time at which a specified species is equal to a specified value

A wrapper code was written around the batch reactor to run ignition delay simulations at the required input temperature, pressure, and species concentrations. The time dependent simulation solution from the batch reactor is passed to the appropriate ignition delay function (based on the definition provided in the YAML file or the function input) and the ignition delay time is calculated. The ignition delay times are returned in a Pandas data frame along with the values of the variables (temperature, pressure, and mole fraction of every chemical species defined in the original processor) at the time of ignition.

The batch reactor physical model is used as a building block for the absorption physical model. Many time dependent species profiles are determined experimentally using absorption as the underlying method, however in some cases rather than reporting time dependent species concentrations or mole fractions directly, the quantity of absorption is reported. Therefore, the MSI framework needs to simulate the reported absorption quantity if direct species measurements are not reported in a publication. Absorbance, A, is a dimensionless quantity that is defined as the logarithm of the ratio of incident to transmitted radiant power through a sample $ln(I_0/I_s)$. Simply put, experimenters are able to determine how much of a specific chemical compound exists in a sample by shining a laser at a given wavelength through a sample and monitoring the incident intensity of the light (I_0), the intensity of the light (I_s) after it travels a certain distance (l) and passes through the sample. Light becomes dimmer as it is attenuated due to the properties of the sample through which it is passing. This measurement can be taken continuously as the sample evolves and reacts over time.

In order to relate the measured signal to the concentration of a specific species, the Beer-Lambert law is used. The Beer-Lambert law is a linear relationship between the logarithm of the ratio of incident to transmitted radiant power (absorbance) and the concentration (c), the molar absorption coefficient (σ) and the optical coefficient (also known as the path length) (l) of a solution $A = ln(I_0/I_s) = \sigma lc$. The path length is the physical distance over which the measurement was taken, and the molar absorption coefficient is a known species-dependent property that is a function of the wavelength of light and temperature. The physical model needs to have the ability to simulate when experimenters report absorption time histories that include the absorption of more than one species. For example, if an experiment reports the time dependent absorption for HO_2 and H_2O_2, the physical model would need to determine: $A = \sigma(T,\lambda)_{HO_2} lc_{HO_2} + \sigma(T,\lambda)_{H_2O_2} lc_{H_2O_2}$. This means the physical model needs information about the species that were measured and the molar absorption coefficients associated with them.

This required information is contained in an additional YAML file, which must be provided in association with any YAML file containing an absorption time history as part of the experimental

data. For example, in Fig. 4, it can be noted that the "Hong_4_abs.yaml", "Troe_4_abs.yaml", and "Troe_5_abs.yaml" are listed directly next to the YAML files containing the associated experiment information. If this type of YAML file is passed into the MSI framework, the information contained in it is parsed into an organized dictionary in the YAML parser sub-module to prepare it for use in the absorbance physical model. An example of what this YAML file might look like is provided in Fig. 11. This file contains information for all species that are considered to be absorbing for a specific experiment, the wavelength of light, the value of the molar absorption coefficient(s), and the functional form the molar absorption coefficient(s). There are three functional forms currently built into the absorption physical module to capture the temperature dependence of an absorbing species. The first functional form called "A" assumes the relationship is linear with temperature, $\sigma = p_1 T(t) + p_2$. The second functional form called "B" assumes the relationship is exponential with temperature $\sigma = p_1(1 - e^{-p_2/T(t)})$. The third functional form assumes the relationship is constant with temperature $\sigma = p_1$, where p_1 and p_2 refer to the values defined for the parameters in the YAML file. The absorption YAML file also contains the associated uncertainties for the molar absorption coefficients for each species, which are used later in the optimization.

To simulate absorption time histories, a wrapper was written around the batch reactor physical model to run simulations at the required initial temperature, initial pressure, and initial species concentrations. The batch reactor returns a Pandas data frame, which contains a time history of temperature, pressure, and mole fraction of every chemical species defined in the original processor. The data frame is passed to the absorption physical model where the mole fraction time history of the absorbing species are converted to concentration time histories, and the temperature dependent molar absorption coefficients of the different absorbing species are computed. The concentration of the absorbing species, the associated temperature dependent molar absorption coefficient, and the path length are then multiplied together to compute the absorption of one species. If necessary, the absorption for multiple species are computed and summed together. The absorption physical model returns a simulation solution in the form of a Pandas data frame, which contains a time history of temperature, pressure, and total absorption for each wavelength of light and associated

absorbing species.

The temperature dependent flow reactor physical model is implemented using a series of Cantera's batch reactor modules. The temperature dependent flow reactor simulates the chemical composition in a continuously flowing system given initial conditions for temperature and pressure. The reacting mixture flowing through the reactor can be modeled as a series of infinitely thin coherent "plugs" traveling in the axial direction of the reactor, each with their own uniform composition. As it flows down the reactor, the residence time of the plug is a function of its position in the reactor. A series of batch reactors are simulated for a given temperatures until a given residence time is reached, then a solution object is returned. This is repeated for all temperatures and all residence times to produce a solution mimicking the flow down the reactor. Experimental data is generally reported as a function of residence time, so after all batch reactors are simulated the solution objects are combined. The flow reactor physical model returns a simulation solution in the form of a Pandas data frame for the temperature, pressure, and mole fraction of every chemical species defined in the original processor at pre-specified residence times.

The jet stirred reactor physical model is implemented using Cantera's ideal gas reactor along with Cantera's flow controller and valve modules. The jet stirred reactor (JSR) simulates a perfectly stirred reactor at steady state conditions. The flow controller valves dictate the flow rate of each reactant into the perfectly stirred ideal gas reactor, and the valve holds the pressure in the reactor constant. The reactor is advanced to steady state conditions for a specific residence time and temperature. Experimental data is generally reported as a function of temperature, so the JSR physical model returns a simulation solution in the form of a Pandas data frame for the pressure, temperature, residence time, and mole fraction of every chemical species defined in the original processor at pre-specified temperatures.

3.8.2 Sensitivity Calculation

Sensitivities are calculated during physical model executions to determine sensitivities to various kinetics parameters and physical model parameters that are used later in the MSI framework.

Sensitivity calculations can be broken down into two main categories. First, is the sensitivity of experimental observables (i.e. a species profile or absorbance profile) to kinetics parameters (i.e. a theoretical molecular parameter or a rate constant). Second, is the sensitivity of experimental observables (i.e. a species profile or absorbance profile) to physical model parameters (i.e. initial temperature, initial pressure, etc.). All physical models except for the ignition delay model use Cantera's built in kinetic sensitivity function to calculate the sensitivity of all experimental observables to every rate constant for every chemical reaction in the processor. A wrapper code was written around each physical module that calls Cantera's kinetic sensitivity function and returns the sensitivities. The sensitivities for the simulation are returned in 3-dimensional numpy [91] arrays where the "i" direction represents the time, residence time or temperature; the "j" direction represents the kinetics parameter (rate constant for the reactions in the order the reactions appear in the processor); and the "k" direction represents the experimental observables.

A brute force method is required to calculate the kinetic sensitivities for the ignition delay model because the ignition delay values are derived quantities. A wrapper code was written around the ignition delay physical module to perturb the rate constants in the processor one by one for a small pre-defined percentage. The ignition delay physical model is run once for each perturbed rate constant and the resulting simulation solutions are saved. The brute force kinetic sensitivity values are determined using the nominal ignition delay solution and the perturbed solutions. The difference between the nominal solution and perturbed solution are normalized with the perturbation to determine the sensitivities. The sensitivities are saved in the same 3-dimensional numpy array format described previously.

All physical model parameter sensitivities must be calculated with a brute force method. A wrapper was written around each physical model that perturbs the individual physical model parameters one at a time. The physical model is run once for each perturbed physical model parameter and the simulation solutions are saved. The brute force physical model parameter sensitivity values are determined in the same manner as ignition delay model, using the nominal solution, perturbed solutions, and perturbation amount. The sensitivities are returned in a list of 3-dimensional

numpy arrays where each array in the list represents one physical model parameter. In an individual array, the "i" direction represents the time, residence time, or temperature; the "j" direction represents the kinetics parameter (rate constant for the reactions in the order the reactions appear in the processor); and the "k" direction represents the physical experimental observables.

The kinetic sensitivities and physical model parameter sensitivities are added to the simulation dictionary.

3.8.3 Extract Experimental Data

The next sub-module of the Opt Runner extracts experimental data from csv files. In the dictionary passed from the YAML parser into the opt runner there is a list of csv file paths. The csv files contain experimental data associated with the experiments represented in the YAML files. The extract experimental data sub-module loops over the list of csv file paths and loads them into Pandas data frames. The Pandas data frames vary in size but have up to five columns. The first column is the independent variable, such as time, temperature, or residence time. The second column is the dependent variable, such as a species concentration, species mole fraction, absorption, or ignition delay time. The last three columns are optional, but if included, the third column is the absolute uncertainty associated with each individual measurement contained in the csv file. The fourth column is the relative uncertainty associated with each individual measurement contained in the csv file. The fifth column is the weighting factor for each individual measurement contained in the csv file. The Extract Experimental Data sub-module loads the data into Pandas data frames, places them in a list, and adds the list to the simulation dictionary.

3.8.4 Interpolate Simulation Data to Experimental Data

This sub-module is only executed if the physical model being executed (for a given iteration) is a batch reactor or absorbance. This sub module take the sensitivities, simulation solution and extracted experimental data as inputs. This sub-module contains a series of sub-functions that interpolate the simulation solution and all the sensitivity coefficients to the time grids of the exper-

imental data set(s). These interpolated simulation solutions are placed in a list , and the interpolated sensitivities are also placed in a list. These lists are added to the simulation dictionary.

3.8.5 Experiment Associated Uncertainty

The final sub-module of the opt runner takes the uncertainties associated with the experimental parameters (i.e. initial temperature, initial pressure, initial species concentrations, time shift, etc.) as well as uncertainties associate with the experimental data contained in the dictionary passed from the YAML parser and adds them to the the simulation dictionary.

3.9 Mapping Parameters

The next module of the MSI framework is the mapping parameters module. The Mapping Parameters module takes the list of simulation dictionaries generated from the opt runner module, the Chebyshev sensitivity dictionary, processor, and master equation reactions list as inputs and returns a list of simulation dictionaries, with additional items appended to each dictionary. This module of the MSI framework contains a sub-module that maps the kinetics sensitivities calculated for the total rate constants to sensitivities for the Arrhenius parameters, and a sub-module that maps the kinetic sensitives for the total rate constant to the sensitivities for the the theoretical molecular parameters.

3.9.1 Mapping Arrhenius Kinetics Parameters

The mapping Arrhenius Kinetics parameters sub-module takes the simulation dictionaries list as an input and iterates over each simulation dictionary in the list, to generate kinetic sensitives mapped to the Arrhenius parameters. For any reactions that are not represented by theoretical molecular parameters, the MSI framework represents them using traditional Arrhenius parameters (i.e. pre-exponential factor (A), temperature exponent (n) and activation energy (Ea)). To treat the Arrhenius parameters as active parameters, the kinetic sensitivities calculated in the previous module (sensitivities of experimental observables to the total rate constant) are mapped onto the

individual Arrhenius parameters using a parametric representation. The parametric representation in Eq. 3.1 is applied to the 3-dimensional array stored in the simulation dictionary, where $\hat{k}$ and $\xi_{n,m}$ are defined in Eq. 3.2 and Eq. 3.3 respectively. Three new 3-dimensional arrays are returned and stored in a list, each one representing one of the Arrhenius parameters. This list of arrays is added to the simulation dictionary.

$$\ln k_{p,n}(T,P,M) = \sum_{m=1}^{m_n} \xi_{n,m}(T,P,M)\hat{k}_{n,m} \tag{3.1}$$

$$\hat{k} = \{\ln(A_n), n_n, -Ea_n/R\} \tag{3.2}$$

$$\hat{k} = \xi_{n,m} = \{1, \ln(T), 1/T\} \tag{3.3}$$

3.9.2 Mapping Theoretical Kinetics Parameters

The Mapping Theoretical Kinetics Parameters sub-module takes the simulation dictionaries, master equations reactions list, and processor as inputs and iterates over each simulation dictionary in the list to generate sensitives mapped to the theoretical molecular parameters. For any reactions that are represented by theoretical kinetics parameters in the MSI framework, it is necessary to map the kinetic sensitives for the rate constants to sensitivities for the theoretical molecular parameters. First, the index of the reactions in the processor being treated with theoretical molecular parameters are identified using the the master equations reactions list. Next, the columns corresponding to the index values are sliced from the 3-dimensional array, which contains the kinetic sensitivities to the total rate constants. Finally, the sensitivities sliced out of the array are mapped to theoretical molecular parameters using Eq.3.6. Where the first term in the equation represents the sensitivities calculated by Cantera and the second term in the equation represents the sensitivities calculated by MESS or Variflex. A parametric representation (to accommodate the Chebyshev reaction format) in the form of Eq. 3.1 is applied where $\hat{k}$ and $\xi_{n,m}$ are defined in Eq. 3.4 and Eq. 3.5 respectively.

Application of this mapping yeilds equation Eq. 3.6.

$$\hat{k}_n = \alpha_{\tau\rho} = \{\alpha_{11}, \alpha_{12}, \alpha_{21} ... \alpha_{N_\tau N_\rho}\} \tag{3.4}$$

$$\xi_n(T,P,M) = \phi(\tilde{T})_\tau \phi(\tilde{P})_\rho = \{\phi(\tilde{T})_1\phi(\tilde{P})_1, \phi(\tilde{T})_1\phi(\tilde{P})_2, \phi(\tilde{T})_2\phi(\tilde{P})_1, ...\phi(\tilde{T})_{N_\tau}\phi(\tilde{P})_{N_\rho}\} \tag{3.5}$$

$$\begin{aligned} S_{ij} &= \sum_n \frac{\partial F_i}{\partial \ln k_{p,n}(T_i,P_i,M_i)} \times \frac{\partial \ln k_{p,n}(T_i,P_i,M_i)}{\partial x_j} \\ &= \sum_n \frac{\partial F_i}{\partial \ln k_{p,n}} \sum_{m=1}^{m_n} \xi_{n,m}(T_i,P_i,M_i) \frac{\partial \hat{k}_{n,m}}{\partial x_j} \end{aligned} \tag{3.6}$$

The mapping in Eq. 3.6 is applied to each column sliced out of the original kinetic sensitivity array, and a list of three dimensional arrays is returned, where each array represents the theoretical kinetics parameters and observables for a single reaction. This list of arrays is added to the simulation dictionary.

3.10 Building Matrices

The next module of the MSI code is the matrix builder, which takes the list of simulation dictionaries, reaction uncertainty csv path, and molecular parameter uncertainty csv path as inputs and assemble the $\boldsymbol{S}$, $\boldsymbol{Y}$ and z matrices. These matrices are used to compute the weighted least squares optimization. The sub-modules are be described in the order they are executed.

3.10.1 Building z Matrix

The z matrix, also known as the z column vector, is the matrix that contains the uncertainty values for the experimental target data as well as the active parameters in the optimization. Building the z matrix is the first sub-module to be executed, and takes the list of simulation dictionaries, re-

action uncertainty csv path, list of master equations, and theoretical kinetic parameters uncertainty csv path as inputs.

The first portion of the z matrix is constructed by looping over the simulation dictionaries and extracting the list of Pandas data frames containing the experimental data. As previously mentioned, these data frames have the option of containing the relative uncertainty, absolute uncertainty, and weighting factor for the experimental data. If any of these values are contained in the data frame, they are used as the default uncertainty values for the experiment. If the data frame does not contain values for absolute uncertainty, and/or relative uncertainty, the code defaults to the values originally listed in the YAML file under the csv path for experimental data – in this case, these values are applied to every data point in the experimental data set.

After a relative and absolute uncertainty value are assigned for every data point for a given experiment, a single uncertainty value must be determined for each data point. First, the absolute uncertainty value for each data point is divided by the actual experimental measurement to convert it to a unit-less absolute uncertainty. Next, the relative uncertainty and the new unit-less absolute uncertainty are combined for each data point by taking the square root of the sum of the squares. The mathematical expression for this operation is: $\sigma_i = \sqrt{\sigma_{relative,i}^2 + (\sigma_{absolute,i})/Y_i^t)^2}$ where Y_i^t is a singular experimental data point at a given time or temperature.

After total uncertainty (σ_i) for each experimental data point is determined for a given data set, the weighting factor is calculated. The weighting factor (w_i) is calculated for a given data set as $w_i = 1/n^{0.5}$ where n is the number of points in the data set. Finally, the total uncertainty calculated for each experimental data point for a given data set is normalized by the value calculated for the weighting factor for that data set, $z_i = \sigma_i/w_i$. This process is repeated for every data set in every simulation dictionary, and the resulting values are vertically stacked in the z column vector.

The next portion of the z column vector is constructed by adding the uncertainties for the kinetic parameters. The kinetic uncertainties file contains the list of reactions in the processor and their associated Arrhenius parameter uncertainties (Fig. 8). This file is loaded into a Pandas data frame, and any rows of the data frame for reactions in the master equation reaction list (these reactions

are being treated with theoretical kinetics parameters not Arrhenius parameters) are removed. The uncertainties for the Arrhenius parameters for the remaining rows in the data frame are added to the z column vector. First, the uncertainties for the pre-exponential factor are added to z matrix. Next, the uncertainties for the temperature exponent are added to the z matrix. Finally, the uncertainties for the activation energy are added to the z matrix.

Next, the uncertainties for the theoretical molecular parameters are added to the z column vector. The theoretical molecular parameter uncertainty csv (Fig. 9) contains the uncertainties associated with the theoretical molecular parameters representing the reactions in the master equation reactions list. This csv is loaded into a Pandas data frame and the uncertainties are added to the z column vector for the theoretical molecular parameters one reaction at a time, in the order they appear in the master equations reactions list.

The final part of z column vector is constructed by adding the uncertainties for all physical model parameters for each experimental data set. To add these uncertainties the simulation dictionaries, are looped over and the physical model parameter uncertainties associated with specific keys in the dictionary (i.e. temperature, pressure, species concentration, time shift, residence time) are added to the z column vector for every experiment. The z column vector is returned from this sub-module.

3.10.2 Building $\boldsymbol{Y}$ Matrix

The $\boldsymbol{Y}$ matrix (column vector) contains the difference between target data and predictions made by the model. Building the $\boldsymbol{Y}$ matrix sub-module takes the list of simulation dictionaries and the $\boldsymbol{x}$ vector as inputs. The first portion of the $\boldsymbol{Y}$ column vector is constructed by determining the natural log difference between the experimental data and the model predictions. Simulation dictionaries are looped over, and the list of Pandas data frames storing experimental data and the simulation solution is extracted. The units in the simulation solution are checked for compatibility with the experimental data; if the units are not compatible, the simulation units are converted to those of the experimental data. The experimental target data and the simulation solution are then passed into a

function where a natural log difference between the two quantities is calculated for each data point, $Y_i = \ln(Y_i^t) - \ln(f_i(\boldsymbol{x}))$. This process is repeated for every data set in every simulation dictionary and the resulting Y_i values are vertically stacked in the Y column vector.

The next portion of the $\boldsymbol{Y}$ column vector is constructed by determining the difference between the *a priori* active parameters (which are also considered to be targets) and the active parameters from the previous iteration of the optimization. In the first iteration of the optimization, (provided the optimization was started from the nominal values) this portion of the $\boldsymbol{Y}$ column vector is filled with zeros because there have been no updates to the active parameter values. For any subsequent iterations of the optimization, this portion of the $\boldsymbol{Y}$ column vector is populated with $\tilde{\boldsymbol{x}}$ vector (after it is multiplied by negative one). The $\tilde{\boldsymbol{x}}$ vector contains values that the simulations and sensitivities were computed at for the current iteration of the optimization. Thus, the negative of the total change in these values is the same as the difference between original target value and the model prediction. The $\boldsymbol{Y}$ column vector is returned from this sub-module.

3.10.3 Building $\boldsymbol{S}$ Matrix

The $\boldsymbol{S}$ matrix contains sensitivity values. This sub-module takes the list of simulation dictionaries, list of master equations, and processor as inputs.

The first step in constructing the $\boldsymbol{S}$ matrix is assembling all the sensitivities of experimental observables to kinetic parameters. The simulation dictionaries are looped over and the list containing sensitivities mapped to Arrhenius kinetic parameters are extracted. The data from the arrays are added to the $\boldsymbol{S}$ matrix in the following manner: first, the sensitivity of the first experimental observable for the pre-exponential factor is added to $\boldsymbol{S}$ matrix, and then the sensitivity of the first experimental observable for the temperature exponent is stacked horizontally adjacent to it, and finally the sensitivity of the first experimental observable for the activation energy is stacked horizontally adjacent to that. For the next experimental observable, the sensitivities for the pre-exponential factor, temperature exponent, and activation energy are stacked horizontally relative to each other but stacked below the values for the first experimental observable. This pattern is

repeated for every simulation dictionary and every experimental observable. Next, the sensitivities for the theoretical kinetic parameters are added to the $\boldsymbol{S}$ matrix. All of the simulation dictionaries are looped over, and the list containing the sensitivities of experimental observables to the theoretical kinetics parameters are extracted. The theoretical molecular parameters for the first experimental observable and the first reaction are added to the $\boldsymbol{S}$ matrix adjacent to the sensitivities for the activation energy for the first experimental observable, for the first experiment. Sensitivities for the theoretical molecular parameters (for the first experimental observable for the first experiment) are stacked horizontally adjacent to the values from the first reaction. The theoretical molecular parameters for the second observable follows the same pattern, except they are stacked below the the sensitivities from the first observable. This pattern continues for the remainder of the simulation dictionaries.

The next portion of the $\boldsymbol{S}$ matrix is constructed by adding the sensitivities of the physical model parameters for each experiment. All of the simulation dictionaries are looped over, and the list of sensitivities to physical model parameters are extracted. The sensitivities of the first observable to the first physical model parameter are stacked horizontally adjacent to the theoretical kinetics parameter sensitivities for the first experiment and first observable. The sensitivities of the second experimental observable to the first physical model parameter of the first experiment is added to the matrix below the first observable, and this pattern continues for every experimental observable in the first experiment. The remainder of the column is then filled with zeros because physical model parameters only have sensitivities to their own experimental data. This pattern is repeated for the remainder of the physical model parameters and observables in the first experiment. The sensitivity of the observables in the second experiment to the physical model parameters are added in the same general structure; they are appended horizontally adjacent to the physical model parameters for the first experiment, but vertically adjacent to the theoretical kinetics parameters for the first observable of the second experiment. The values in the column above and below the sensitivities for the physical model parameters of the second experiment are filled with zeros. This pattern is repeated for the remainder of the experiments, observables, and physical model parameters. The

only time a new column is not appended to the $\boldsymbol{S}$ matrix, when a physical model parameter is added, is if two experiments are using the same set of absorption coefficients for the same species at the same wavelength, and a column has already been added to the $\boldsymbol{S}$ matrix to account for existence. In this case, the sensitivity of an experimental observable to the absorption coefficient is placed in the column corresponding to that absorption coefficient and vertically aligned with its corresponding experiment and experimental observable. The next portion of the $\boldsymbol{S}$ matrix is constructed by appending an identity matrix to the $\boldsymbol{S}$ matrix with the "i" by "j" dimensions of the number of active parameters in the model. This is to represent the sensitivity of each active parameter to itself. The $\boldsymbol{S}$ matrix is then returned from this sub-module.

3.11 Adding Rate Constant Target Values

The next module of the MSI framework takes $\boldsymbol{Y}$, z, and $\boldsymbol{S}$ matrices as inputs and appends rate constant target values. The sub-modules are be described in the order they are executed.

3.11.1 Adding Rate Constant Target Values to z

This sub-module takes in the z matrix and the rate constant target value csv file path as inputs. The rate constant target value csv file is loaded into a Pandas data frame. The uncertainty column and the weighting factor column are extracted from the data frame. The uncertainties column is divided by the weighting factor column. The resulting values are appended to the z column vector. The final z column vector is returned from this sub-module.

3.11.2 Adding Rate Constant Target Values to $\boldsymbol{Y}$

This sub-module takes in the $\boldsymbol{Y}$ matrix, the rate constant target value csv file path, and the processor as input. The rate constant target value csv file is loaded into a Pandas data frame. Each row of the data frame is iterated over and using the temperature, pressure, mixture composition (in the row), and processor, a model prediction for the rate constant at the given conditions is made. Similar to the experimental data, the natural log difference of the experimentally determined rate

constant target and the model prediction for the rate constant is determined ($Y_i^t = \ln(Y_i^t) - \ln(f_i(\boldsymbol{x}))$) and the value is appended to the $\boldsymbol{Y}$ column vector. This is repeated for every row of the Pandas data frame.

It should also be noted that this module is capable of handling more complicated rate constant target value definitions (which are often times presented in publications) by using specific syntax in the rate constant target csv file. The following examples explain the syntax for more complicated rate constant target values.

- HO_2 + HO_2 = *: This represents the summation of all the rate constant values (at a specific temperature, pressure, and mixture composition) that have HO_2 + HO_2 as reactants.

- HO_2 + HO_2 = H_2O_2 + O_2 / HO_2 + HO_2 = *: This represents the rate constant value of HO_2 + HO_2 = H_2O_2 + O_2 (at a specific temperature, pressure, and mixture composition) divided by the summation of all the rate constant values (at the same specific temperature, pressure, and mixture composition) that have HO_2 + HO_2 as reactants.

- HO_2 + HO_2 = H_2O_2 + O_2 (+) H + HO_2 = H_2 + O_2: This represents the summation of the rate constant value for the reaction of HO_2 + HO_2 = H_2O_2 + O_2 and H + HO_2 = H_2 + O_2(at a specific temperature, pressure, and mixture composition).

These different syntax methods for creating more complicated rate constant target expressions can be combined in any way the user needs to represent the experimental data they are attempting to input into the model structure. The final $\boldsymbol{Y}$ column vector is then returned from the sub-module.

3.11.3 Adding Rate Constant Target Values to $\boldsymbol{S}$

This sub-module performs the appropriate mapping on the rate constant target values and appends the results to the $\boldsymbol{S}$ matrix. This sub-module takes in the $\boldsymbol{S}$ matrix, the rate constant target value csv file path, and the processor as inputs. The rate constant target value csv file is loaded into a Pandas data frame. The columns of the data frame relevant to this sub-module are the columns that contain the rate constant target values, the reaction string, the temperature, the pressure, and

the mixture composition. Each row of the data frame is iterated over, and using the target temperature and pressure, the mapping in Eq. 3.7 is computed for a given target rate constant. The mapped sensitivities are appended to the $\boldsymbol{S}$ matrix in the columns corresponding to the target reaction's kinetic parameters. The final $\boldsymbol{S}$ matrix is then returned from this sub-module.

$$S_{ij} = \frac{\partial \ln k_{p,n}(T_i, P_i, M_i)}{\partial x_j} \tag{3.7}$$

3.12 Matrix Math

This module of MSI is responsible for taking the fully assembled $\boldsymbol{z}$, $\boldsymbol{Y}$, and $\boldsymbol{S}$ matrices and computing the $\boldsymbol{x}$ matrix and the covariance matrix ($\boldsymbol{\Sigma}$). Eq. 3.8 is solved to compute the optimal $\boldsymbol{\Delta x}$.

$$\boldsymbol{S\Delta x} = \boldsymbol{Y} \pm \boldsymbol{Z} \tag{3.8}$$

The first step of solving Eq. 3.8 is dividing $\boldsymbol{S}$ and $\boldsymbol{Y}$ element-wise by $\boldsymbol{z}$, where $y_i = Y_i/z_i$ and $s_{ij} = s_{ij}/z_i$. After they are normalization these matrices are referred to as $\boldsymbol{y}$ and $\boldsymbol{s}$. Next, the pseudo-inverse of $\boldsymbol{s}$ matrix is computed and dotted with the $\boldsymbol{y}$ matrix yielding $\boldsymbol{\Delta x}$ (Eq. 3.9).

$$\boldsymbol{\Delta x} = (\boldsymbol{s}^{\mathrm{T}}\boldsymbol{s})^{-1}\boldsymbol{s}^{\mathrm{T}} \cdot \boldsymbol{y} \tag{3.9}$$

For context, each iteration starts from some $\boldsymbol{x_\gamma} = 0$, where γ refers to the iteration number (the starting values at $\gamma = 0$ for $\boldsymbol{x}$ are generally taken to be zeros). For each iteration, the simulations and the sensitivities are evaluated at $\tilde{\boldsymbol{x}}$, which is equal to $\boldsymbol{x_\gamma}$, and Eq. 3.9 is used to find the optimal $\boldsymbol{\Delta x}$. The optimal values to compute the simulations and sensitivities for next iteration are determined via Eq. 3.10. Thus, in the following iteration, $\boldsymbol{x_{\gamma+1}}$ becomes $\tilde{\boldsymbol{x}}$, where $\boldsymbol{x}$ represents the total update to active parameters in the model from their nominal values. As the model converges, the values in the $\Delta\boldsymbol{x}$ matrix computed for each iteration become closer and closer to zero, and no further updates to parameters are needed. The covariance matrix is also computed in this sub-module.

The covariance matrix is computed with Eq. 3.11. The $\Delta \boldsymbol{x}$ matrix, $\boldsymbol{x}$ matrix, and covariance matrix are returned from this sub-module.

$$\boldsymbol{x}_{\boldsymbol{\gamma}+\mathbf{1}} = \boldsymbol{\Delta x} + \tilde{\boldsymbol{x}} \tag{3.10}$$

$$\boldsymbol{\Sigma} = \left(\boldsymbol{s}^T \boldsymbol{s}\right)^{-1} \tag{3.11}$$

3.12.1 Breakup $\Delta \boldsymbol{x}$

This sub-module of MSI is responsible for taking the values contained in $\Delta \boldsymbol{x}$ matrix and separating them into different objects and data types. These objects and data types are passed into the appropriate modules and updates are applied to active model parameters (Eq. 3.10), meaning in the next iteration of the optimization simulations and sensitivities can be computed with $\boldsymbol{x}_{\boldsymbol{\gamma}+\mathbf{1}}$. This sub-module iterates over the $\Delta \boldsymbol{x}$ matrix and separates the values into three main categories, namely: a list of dictionaries containing physical model parameter updates for each experiment, a dictionary containing Arrhenius parameter updates, and a dictionary containing theoretical molecular parameter updates. These dictionaries are returned and passed to the appropriate function to apply the updates.

3.13 Updating Files

This module of MSI is responsible for updating the active kinetic parameters in the cti file and the active physical model parameters in the YAML files. The Updating Files module takes the dictionary containing the Arrhenius parameter updates, the dictionary containing the theoretical molecular parameter updates, the list of dictionaries containing the physical model parameter updates for each experiment, the Chebyshev sensitivity dictionary, and the list of YAML files paths as inputs.

3.13.1 Map Theoretical Molecular Parameter Updates

Reactions that are represented by theoretical kinetics parameters require that updates to those parameters be mapped to the coefficients of the Chebyshev polynomial used to represent the rate constants. This step must be completed before updating the active kinetic parameters in the cti file. The mapping between kinetic parameters and rate constants relies on the locally linear surrogate model. In the current version of the MSI framework, the MESS or Variflex code is not run every iteration. For each rate constant represented with theoretical molecular parameters, the update to the Chebyshev coefficients are calculated by multiplying the $\Delta \boldsymbol{x}$ values by their respective sensitivity coefficients. The total change in the rate constant is calculated by summing the series of updates found in the prior steps. After multiplication, these arrays are summed, resulting in the $\Delta \boldsymbol{x}$ for a given rate constant based on the updates to its theoretical kinetics parameters. Mathematically, $\Delta k = \Sigma_{m=0}^{p} \Delta x_m x \frac{\partial k}{\partial x_m}$, where p is the number of molecular parameters representing a rate constant, k. The updates to the Chebyshev polynomial coefficients are computed for each reaction represented by theoretical molecular parameters and saved to a dictionary. This sub-model returns a dictionary containing the updates for all reactions represented by theoretical kinetics parameters in the form of Chebyshev polynomials.

3.13.2 Updating YAML Files

This sub-module is responsible for applying the updates to the active physical model parameters. The original YAML files store all the values for the physical model parameters. This sub-module takes the list of dictionaries containing the updates to the physical model parameter as an input. On the first iteration of the optimization, this sub-module also makes a copy of every original YAML file. The copies of YAML files are saved with the original file name and the addition of "_updated". The updates to the new physical model parameter values are stored within the "_updated" files. These files are then read into MSI for all subsequent iterations. Then, this sub-module loops over all the dictionaries in the list passed in (each dictionary in that list refers a single experiment, corresponding to a single YAML file) and applies the physical model parameter

updates to the corresponding YAML files. If an experiment has an associated absorbance YAML file, the molar absorption coefficients in the file are also updated when the dictionary associated with that experiment is iterated over.

3.13.3 Updating cti File

This sub-module is responsible for applying the updates to the rate constant expressions (made up of kinetics parameters) in the cti file. This sub-module takes the dictionaries of updated Chebyshev polynomials, the updated Arrhenius parameters, and the processor as inputs. All the reactions in the processor are iterated over, and the appropriate updates are applied to each rate constant. The updated processor is then written back out to the updated cti file.

After the YAML files and the cti file are updated, another iteration begins using the updated active parameters. The iterations continue until the number of predefined iterations is completed.

3.14 Test Scripts

The following scrips are not integrated into the MSI Library source code, rather they are test scripts used to help validate the software is working correctly. If major modifications are made to the software, it is suggested these tests are run again to validate no major issues were introduced during the modification.

3.14.1 S_{ij} Test

This test script was created to validate the sensitivity matrix generated by the MSI code. This code compares $\boldsymbol{S}$ matrix calculated by the MSI code to that calculated by a wrapper. The wrapper calculates $\boldsymbol{S} = \partial \boldsymbol{Y}/\partial \boldsymbol{x}$ using a brute force method. The wrapper loops through all x_i parameters and perturbs each one slightly, one by one, then uses MSI's calculated $\boldsymbol{Y}$ vector and the definition of the local sensitivity coefficient ($S_{ij} = \partial Y_i/\partial x_i$) to calculate S_{ij}. The elements of S_{ij} calculated by the wrapper are compared to the elements of S_{ij} calculated internally by MSI for $\boldsymbol{x}_{\gamma=0}$ (the first iteration before any updates are applied). This script tests the code's ability to calculate each S_{ij}

element correctly and the code's ability to use each *xi* correctly when updating values.

3.14.2 Reproduction of Target Values Test

This test script was created to validate the calculation of the covariance matrix and the ability of MSI to reproduce target (Y_i^t) values. This script computes the optimization using only one data point of fake target data with a very small uncertainty value ($1x10^{-6}$). The script has the ability to add a piece of fake target data for rate constant targets (for any reaction), macroscopic observable targets (for any experiment type, i.e. batch reactor, flow reactor, jet stirred reactor, etc.), and each observable type (mole fractions, concentrations, absorbance profiles, ignition delay times, etc.).

First, a selection must be made by the user for the type of target to test against, and then the MSI code is executed. The index of the optimized $\boldsymbol{Y}$ vector that refers to the difference between the test target data and the model prediction is extracted. This value should be very close to zero, or numerically, something less than the input uncertainty. This tests the code's ability to reproduce target data and calculate the covariance matrix.

3.14.3 Weighting Factor Testing

The Weighting Factor test script is nearly identical to the Reproduction of Target Values Test script but includes multiple (100) identical versions of the same data point. This tests the codes ability to calculate the weighing factor (W_i) correctly.

Chapter 4: Towards a High-Accuracy Kinetic Database Informed by Theoretical and Experimental Data: CH_3 + HO_2 as a Case Study

4.1 General Notes to the Reader

The remainder of this work will discuss three specific MSI case studies; the first study focuses on the rate constants of the CH_3 + HO_2 and associated reactions, the second study focuses on the rate constants of the HO_2 + HO_2 and associated reactions and the third study focuses on a variety of reactions linking together smaller subsystems. The scientific motivation behind selecting these specific reaction systems will be discussed in their individual sections in detail. In addition to the individual scientific motivations for each of these systems, they also made good candidates for demonstrating key elements of the MSI approach for the following reasons: 1) Calculations involving theory parameters could be completed with reasonable computational resources and time. 2) There exists enough of a selection of raw experimental data from experiments involving small subsystems, homogeneous environments, and dilute conditions. 3) These subsystems have overlapping reactions and thus could be used to begin to grow a network of strongly linked, redundant data.

Before beginning the discussions and presenting results for each individual subsystem it is perhaps worth noting the following which applies to each kinetic system discussed in this document. Each chapter discussing a subsystem of reactions contains a Table labeling and numbering key reactions and parameters of interest; individual reactions or parameters may change number from chapter to chapter. When reading the results and discussion please note the label and number refer to the specific table in that chapter. Additionally, if supplemental material or supporting material is mentioned in a specific chapter the reader can find this material (labeled with its corresponding chapter title) at the end of the document.

4.2 Scientific Motivation: $CH_3 + HO_2$

In this study, an MSI optimization to determine the channel-specific $CH_3 + HO_2$ rate constants as compared to a previous rate-parameter optimization will be presented. Temperature dependence for this system is not well established from experimental data alone and recent studies have indicated that uncertainties in H_2O_2/CH_3 and H_2O_2/CH_4 reactions have important ramifications for modeling a variety of combustion systems. As natural gas becomes an increasingly viable energy source, understanding methane and hydrogen combustion systems and the uncertainties associated with them, becomes increasingly important [92, 93]. For example, the uncertainties in $CH_3 + HO_2 = CH_4 + O_2$ and $CH_3 + HO_2 = CH_3O + OH$ were found to be among the largest contributors to the uncertainties in NO_x production in a natural gas turbine relevant flame [94] such that improved quantification, and reduction of, uncertainties in these reactions are of interest to computational engine design.

All to say that this system can certainly benefit from establishing more refined rate constants and uncertainties by using MSI. Notably, this study found the optimization performed using the MSI approach applied to carefully selected (mostly raw) experimental data yielded an opposite temperature dependence for the channel-specific $CH_3 + HO_2$ rate constants as compared to a previous rate-parameter optimization. While both optimization studies use the same theoretical calculations to constrain model parameters, only the present optimization, which incorporates theory directly into the model structure, yields results that are consistent with theoretical calculations.

4.3 Implementation: $CH_3 + HO_2$

Key elements of the MSI approach are demonstrated here for the $CH_3 + HO_2$ reactions as a case study. This case study leverages Burke's previous MSI analysis of a pure H_2O_2 system [69] for analysis of H_2O_2/ CH_4 experiments.

For system one, ($CH_3 + HO_2$) kinetic parameters for R1-R8 and physical model parameters were assigned to represent uncertainties in the kinetic and physical models (Table 2). In this work

other “secondary” reactions were taken from the Foundational Fuel Chemistry Model (FFCM) [84]. Rate constants and brute-force sensitivity coefficients were calculated via TST-ME [80] in VARIFLEX [82] and PAPR-MESS [81]. Macroscopic observables and sensitivity coefficients were calculated via homogeneous, isochoric or isobaric models in Cantera [83].

It is perhaps worth noting the following which applies to this system and all following kinetic systems discussed in this document. The exact parametric and structural uncertainties to be considered naturally depend on the specifics of the system of interest. For example, when the systems involve larger kinetic mechanisms, where the potential for missing reactions is higher, greater consideration of “missing" reactions may be required (e.g. by using RMG as in [70]). A broader discussion of various sources of parametric and structural uncertainty can be found elsewhere [71], but as discussed elsewhere [71] and in the results section below, consideration of all relevant parametric and structural uncertainties is important for obtaining physically realistic parameter adjustments.

It was determined that the rate constants for R1 and R8, were sufficiently constrained over the relevant temperature range by macroscopic observables and rate constant targets, therefore Arrhenius parameters (as opposed to molecular parameters within theoretical kinetics models) were treated as active parameters. For R2-R7, molecular parameters within theoretical kinetics models were treated as active parameters.

For R2-R5, the molecular parameters were taken to be the same as Burke’s previous work [69]. The active molecular parameters included transition state barrier heights ($E^{\dagger}$); scaling factors for all harmonic frequencies of reactants and products (ν_{all}); scaling factors for all (transition-state) harmonic frequencies (ν_{TS}), transitional mode frequencies (ν_{tr}), symmetric stretching mode frequencies (ν_{ss}), and imaginary frequencies (ν_{imag}) in tight transition states; VRC-TST correction factors for loose transition states ($f_{\mathrm{VRC-TST,c}}$); the smaller of the two well depths for pre- and post-transition state complexes (E_{w}); and scaling factors for hindered rotor potentials (η).

For R6-R7, a more limited set of parameters was necessary to represent the uncertainties in the theoretical kinetics calculations. For R6, which proceeds through a barrierless entrance transition

state to form a van der Waals complex that dissociates to products via an exit transition state with a barrier submerged by only ∼1.9 kcal/mol relative to the reactants, the rate constant over the temperature range of interest is essentially controlled by only the exit transition state, to which active $E_6^\dagger$ and $\nu_{\mathrm{low}(6)}$ (a scaling factor for the lowest four harmonic frequencies in the transition state) are ascribed. (Variation of $\eta_{\mathrm{TS}(6)}$ and $\nu_{\mathrm{imag}(6)}$ within their uncertainty limits yielded negligible differences in the rate constants over the relevant temperature range.) For R7, which proceeds via a barrierless entrance transition state to form a CH_3OOH complex that dissociates to products via another barrierless exit transition state that lies ∼25 kcal/mol below the reactants, the rate constant is essentially controlled by the barrierless entrance transition state, to which an active $f_{\mathrm{VRC-TST,c}}$ is ascribed. (Calculations using roaming estimates from [95] for J = 0 show negligible impact of roaming on the present results, so roaming was not included here.)

Table 1: List of active model parameters considered in the optimization.[a,b]

Reaction		Kinetic Parameters	
R1	H_2O_2(+M) = OH + OH(+M)	$A'_{(1)}, n_{(1)}, E_{(1)}$	
R2	H_2O_2 + OH = HO_2 + H_2O	$E^\dagger_{(2)}, \nu'_{\mathrm{all}(2)}, \nu'_{tr(2)}, \nu'_{\mathrm{ss}(2)}, \nu'_{\mathrm{imag}(2)}, E_{\mathrm{w}(2)}, \eta'_{\mathrm{H_2O_2}}, \eta'_{\mathrm{TS}(2)}$	
R3	HO_2 + HO_2 = H_2O_2 + O_2	$E^\dagger_{(3)}, \nu'_{\mathrm{all}(3)}, \nu'_{\mathrm{tr}(3)}, \nu'_{\mathrm{ss}(3)}, \nu'_{\mathrm{imag}(3)}, E_{\mathrm{w}(3)}, \eta'_{\mathrm{TS}(3)}$	
R4	HO_2 + OH = H_2O + O_2	$E^\dagger_{(4g)}, \nu'_{\mathrm{all}(4)}, \nu'_{\mathrm{tr}(4g)}, \nu'_{\mathrm{ss}(4g)}, \nu'_{\mathrm{imag}(4g)}, E_{\mathrm{w}(4g)}, \eta'_{\mathrm{TS}(4g)}$	
		$E^\dagger_{(4e)}, \nu'_{\mathrm{TS}(4e)}, \nu'_{\mathrm{tr}(4e)}, \nu'_{\mathrm{ss}(4e)}, \nu'_{\mathrm{TS}(4e)}, f'_{\mathrm{VRC-TST,c}(4)}$	
R5	OH + OH = O + H_2O	$E^\dagger_{(5g)}, \nu'_{\mathrm{all}(5)}, \nu'_{\mathrm{tr}(5g)}, \nu'_{\mathrm{ss}(5g)}, \nu'_{\mathrm{imag}(5g)}, E_{\mathrm{w}(5g)}$	
		$E^\dagger_{(5e)}, \nu'_{\mathrm{TS}(5e)}, \nu'_{\mathrm{tr}(5e)}, \nu'_{ss(5e)}$	
R6	CH_3 + HO_2 = CH_4 + O_2	$E^\dagger_{(6)}, \nu'_{\mathrm{low}(6)}, (Q_{\mathrm{anh,c,CH_3}})$	
R7	CH_3 + HO_2 = CH_3O + OH	$f'_{\mathrm{VRC-TST}(7)}, (Q_{\mathrm{anh,c,CH_3}})$	
R8	CH_4 + OH = CH_3 + H_2O	$A'_{(8)}, n_{(8)}, E_{(8)}$	
Macroscopic Observables		Physical Model Parameters	
E1-E2	Shock-heated H_2O_2/H_2O/O_2/Ar	$T'_e, P'_e, M'_{\mathrm{H_2O_2},o,e}, M'_{\mathrm{H_2O},o,e}, M'_{\mathrm{O_2},o,e}$	$e = 1...2$
E3	Shock-heated H_2O_2/H_2O/O_2/Ar	$T'_e, P'_e, M'_{\mathrm{H_2O_2},o,e}, M'_{\mathrm{H_2O},o,e}, M'_{\mathrm{O_2},o,e}$	$e = 3$
E4	Shock-heated H_2O/O_2/H/Ar	$T'_e, P'_e, M'_{\mathrm{H},o,e}, M'_{\mathrm{H_2O},o,e}, M'_{\mathrm{O_2},o,e}$	$e = 4$
E5-E9	Shock-heated H_2O_2/Ar	$T'_e, P'_e, M'_{\mathrm{H_2O_2},o,e}, \sigma'_{1,\mathrm{H_2O_2}}, \sigma'_{2,\mathrm{H_2O_2}}, \sigma'_{1,\mathrm{HO_2}}, \sigma'_{2,\mathrm{HO_2}}$	$e = 6...10$
E10	Shock-heated H_2O_2/H_2O/O_2/Ar	$T'_e, P'_e, M'_{\mathrm{H_2O_2},o,e}, M'_{\mathrm{H_2O},o,e}, M'_{\mathrm{O_2},o,e}, \hat{\sigma}'_{1,\mathrm{H_2O_2}}, \hat{\sigma}'_{1,\mathrm{HO_2}}$	$e = 10$
E11	Shock-heated CH_4/H_2O_2/H_2O/O_2	$T'_e, P'_e, M'_{\mathrm{CH_4},o,e}, M'_{\mathrm{H_2O_2},o,e}, M'_{\mathrm{H_2O},o,e}, M'_{\mathrm{O_2},o,e}$, $\hat{\sigma}'_{1,\mathrm{H_2O_2}}, \hat{\sigma}'_{1,\mathrm{HO_2}}$	$e = 11$

[a] Note that $'$ indicates ln() of that quantity.
[b] Note that parameters in () are represented by only a single value but influence multiple reactions.

Previous theoretical calculations for R6 and R7 neglect anharmonicity of the CH_3 umbrella

mode [96, 97]. To date, estimates using different assumptions indicate anharmonic corrections ranging from nearly zero to 30% [98, 99] (Jasper, personal communication) – precluding definitive conclusions for its role. We retain the nominal values from [96, 97] but assign another active parameter for R6 and R7 corresponding to a scaling factor for the anharmonic correction from [98], $Q_{\mathrm{anh,c,CH_3}}$ – effectively serving as a parameterization of the structural uncertainty related to anharmonicity in the CH_3 umbrella mode.

Physical model parameters from experiments included as macroscopic observable targets were also among the active model parameters. These parameters included initial temperatures, T_e, initial pressures, P_e, initial species mole fractions, $M_{i,o,e}$, and coefficients for temperature-dependent cross sections of H_2O_2 and HO_2 at 215 nm ($\sigma_{1,\mathrm{H_2O_2}}$, $\sigma_{2,\mathrm{H_2O_2}}$, $\sigma_{1,\mathrm{HO_2}}$, and $\sigma_{2,\mathrm{HO_2}}$) and 227 nm ($\hat{\sigma}_{1,\mathrm{H_2O_2}}$ and $\hat{\sigma}_{1,\mathrm{HO_2}}$).

Target class (I): *ab initio* calculations

Molecular properties from *ab initio* calculations were used as targets for the molecular parameters for R2-R5 from (Harding and Klippenstein, Unpublished) [69], for R6 from [96], and R7 from [97] with uncertainties listed in Table 2 reflecting the level of electronic structure theory and the level of detail in treatment [69, 70, 71, 80]. In particular, barrier height uncertainty estimates for the various method/basis set combinations were based on quoted errors from a recent paper by Klippenstein and Cavallotti [80].

Target class (II): Rate constant determinations

Rate constant determinations (from studies other than those treated as macroscopic observables) were included as targets for R2, R4, R5, and R8. Given that uncertainties in rate constant determinations are often higher than typically estimated and prone to systematic errors [69, 71, 44], conservative uncertainties of a factor of two (or the scatter in the data if larger) were applied.

Table 2: List of targets used in the optimization.[a,b]

I. *Ab initio* calculations			Source
$E_2^\dagger$(3 kcal/mol), $E_3^\dagger$(2.3 kcal/mol), $E_{4g}^\dagger$(2 kcal/mol), $E_{4e}^\dagger$(2 kcal/mol), $E_{5g}^\dagger$(1.4 kcal/mol), $E_{5e}^\dagger$(1.4 kcal/mol), ν'_{all}(0.03), ν'_{TS}(0.03), ν'_{tr}(0.1), ν'_{ss}(0.18), ν'_{imag}(0.18), E_w(1 kcal/mol), $\eta'_{H_2O_2}$(0.1), $\eta'_{TS(2)}$(0.26), $\eta'_{TS(3)}$(0.41), $\eta'_{TS(4g)}$(0.41), $\eta'_{TS(4e)}$(0.41), $f'_{VRC-TST,c(4)}$(0.7), $E_6^\dagger$(2.3 kcal/mol), $\nu'_{low(b)}$(0.1), Q_{anh,c,CH_3}(1), $f'_{VRC-TST(7)}$(0.262)			(Harding and Klippenstein Unpublished) [96, 97]
II. Rate constant measurements			
k'_2, k'_4, k'_5, k'_8 (see text)			[69, 100, 101]
III. Global exp.	IV. Exp. conditions		
$M'_{OH,e}(t)(0.05)$, $M'_{H_2O,e}(t)(0.05)$	$T'_e(0.01)$, $P'_e(0.02)$, $M'_{H_2O_2,o,e}(0.05)$, $M'_{H_2O,o,e}(0.05)$, $M'_{O_2,o,e}(0.05)$	$e = 1...2$	[102]
$M'_{H_2O,e}(t)(0.05)$	$T'_e(0.01)$, $P'_e(0.02)$, $M'_{H_2O,o,e}(0.1)$, $M'_{O_2,o,3}(0.01)$, $M'_{H_2O_2,o,3}(0.01)$,	$e = 3$	[103]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01)$, $P'_e(0.02)$, $M'_{H_2O,o,e}(0.1)$, $M'_{O_2,o,e}(0.01)$, $M'_{H,o,e}(2.3)$,	$e = 4$	[104]
$abs'_e(t)(0.1)$	$T'_e(0.02)$, $P'_e(0.04)$, $M'_{H_2O_2,o,e}(0.1)$ $\sigma'_{1,H_2O_2}(0.7)$, $\sigma'_{2,H_2O_2}(0.3)$ $\sigma'_{1,HO_2}(0.7)$, $\sigma'_{2,HO_2}(0.3)$	$e = 5...9$	[105]
$abs'_e(t)(0.1)$, $M'_{H_2O,e}(t)(0.05)$, $M'_{OH,e}(t)(0.05)$,	$T'_e(0.02)$, $P'_e(0.04)$, $M'_{H_2O_2,o,e}(0.05)$, $M'_{H_2O,o,e}(0.2)$, $M'_{O_2,o,e}(0.2)$, $\hat{\sigma}'_{1,H_2O_2}(0.7)$ $\hat{\sigma}'_{1,HO_2}(0.7)$	$e = 10$	[106]
$abs'_e(t)(0.1)$ $M'_{O_2,e}(t)(0.05)$, $M'_{H_2O,e}(t)(0.05)$	$T'_e(0.02)$, $P'_e(0.04)$, $M'_{H_2O_2,o,e}(0.05)$, $M'_{H_2O,o,e}(0.05)$, $M'_{O_2,o,e}(0.05)$, $M'_{CH_4,o,e}(0.05)$, $\hat{\sigma}'_{1,H_2O_2}(0.7)$ $\hat{\sigma}'_{1,HO_2}(0.7)$	$e = 11$	[107]

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

Target classes (III) and (IV): Macroscopic observables and experimental conditions

Raw data for the macroscopic observables from experiments indicated in Table 2 were included as targets using uncertainties as reported or estimated using typical values [103, 102, 104, 108, 106]. (While other studies examined R6 and R7 [92, 96, 109], further analysis, indicated strong influences of several secondary reactions and, therefore, they were not included as targets.) Macroscopic observables used as targets included time-resolved H_2O mole fractions, OH mole

fractions, and absorbance at 215 and 227 nm. Reported values for the experimental conditions were used as targets for physical model parameters with uncertainties as reported or using typical values [103, 102, 104, 108, 106].

4.4 Results: CH_3 + HO_2

The set of optimized molecular parameters, rate parameters, and physical model parameters constrained by the respective targets in Table 2 yield values and predictions in reasonable consistency with the targets from *ab initio* calculations, rate constant determinations, macroscopic observables and reported experimental conditions and other experiments used as targets for the CH_3 + HO_2 system. The discussion below will focus on the macroscopic observables for shock-heated $CH_4/H_2O_2/H_2O/O_2$/Ar mixtures, that emphasize R1-R8 in Table 2 and k_6 and k_7 in particular

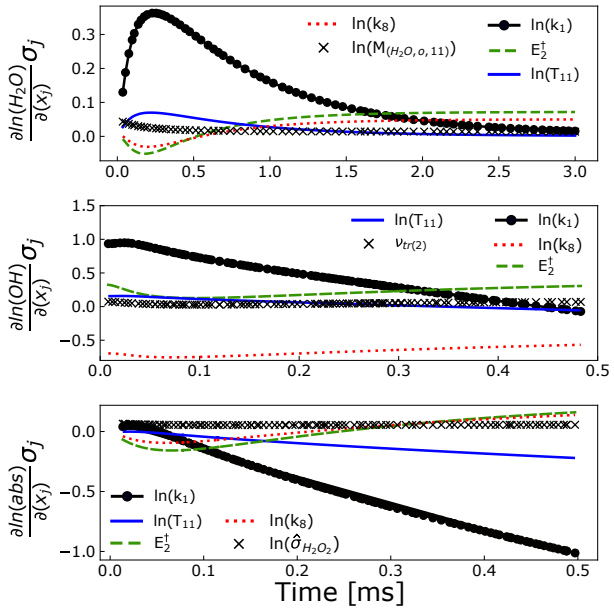


Figure 1: Uncertainty-weighted sensitivity analysis for the OH, H_2O, and absorbance time profiles of Hong et al. [108] For the figure, k_1 and k_8 are shown with uncertainties of a factor of two and $\hat{\sigma}_{H_2O_2}$ and $\hat{\sigma}_{HO_2}$ are shown with uncertainties of 10%.

Uncertainty-weighted sensitivity analysis of the time profiles for H_2O, OH, and absorbance at 227 nm are shown in Fig. 1. Uncertainty-weighted sensitivity coefficients indicate the parameters whose uncertainties influence the uncertainties in the model predictions and, likewise, which pa-

rameters can be constrained by inclusion of data for those observables. Inspection of the five most influential parameters shown in Fig. 1 reveals that predictions of all three observables are primarily influenced by uncertainties in parameters related to k_1, k_2 ($E_2^\dagger$ and $\nu_{\mathrm{tr}(2)}$), k_8, and physical model parameters (T_{11}, $M_{\mathrm{H_2O},o,11}$, and $\hat{\sigma}_{\mathrm{H_2O_2}}$).

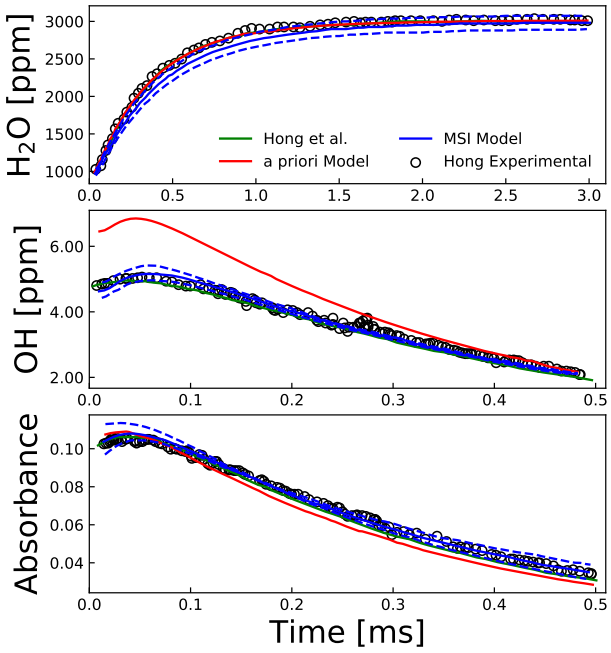


Figure 2: H_2O, OH and absorbance time profiles at 1103 K in shock-heated $H_2O/H_2O_2/CH_4/Ar$. Symbols represent experimental data from Hong et al. [108]; lines represent model predictions.

The time histories from theses experiments were used to determine k_6 and k_7 individually [108] and exhibit some sensitivity to their active parameters. However, it is clear that after including constraints from *ab initio* calculations [96, 97], uncertainty-weighted sensitivity coefficients for parameters related to k_6 and k_7 are comparatively minimal. These results demonstrate the implicit interconnections among rate constant determinations for R6-R8 and those for R1-R5; these results also highlight the advantage of analyzing raw data to capture the information content and unravel the inherent interconnections among rate constant determinations, rather than relying on derived data.

Predicted time profiles for H_2O, OH and absorbance at 227 nm are shown in Fig. 2 for the *a priori* model and MSI model, as compared to the experimental data from Hong et al. [108]. While

the *a priori* model predicts lower absorbance and higher OH than observed experimentally, the MSI model predictions for H_2O, OH and absorbance at 227 nm are consistent with experimental data within uncertainties. Given that the three time profile predictions are primarily influenced by k_1 and k_2, which are already constrained by our H_2O_2 experiments, it is not surprising that only including H_2O_2 experiments as targets (without any CH_4-containing experiments as targets) yields optimized predictions remarkably close to the optimized profiles including all data from Table 2. However, including the CH_4/H_2O_2 experiments as targets did yield minor improvements in the OH peak and absorbance profile (of 10%). Additionally, inclusion of the CH_4/H_2O_2 experiments as targets did also yield some minor adjustments in parameters for R6-R8, though predictions with these parameters at their nominal values (and R1-R5 parameters optimized to only H_2O_2 experimental targets) are already reasonably consistent with experimental data.

The ability of the model to predict these time profiles so well without including the CH_4/H_2O_2 experiments as targets is a testament to the advantage of leveraging data from multiple sources. Theoretical calculations for R6-R7, data for pure H_2O_2 decomposition, and k_8 determinations are leveraged here in the interpretation of the raw experimental data to extract more accurate information for R6-R7.

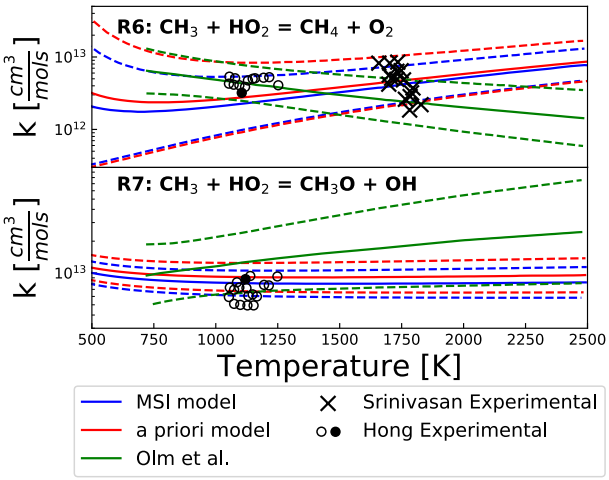


Figure 3: Rate constants for R6 and R7. Original experimental interpretations [108, 96] are shown by symbols (filled in for the conditions of Fig. 2). Dashed lines show prior and posterior uncertainties.

Fig. 3 shows comparisons of k_6 and k_7 from the *a priori* model (which are simply calculated rate constants from [96] and [97] respectively), MSI model, and model of Olm et al. [85] along with experimental rate constant determinations from Hong et al. [108] (from the experiments shown in Fig. 2) and [96]. One can note from Fig. 3 the values and uncertainties of the rate constants in the *a priori* model and MSI model are very similar, thus indicating that the MSI rate constants are primarily controlled by theory.

It should be noted that while both the present MSI optimization and the Arrhenius parameter optimization from Olm et al. [85] use theoretical calculations [96, 97] as targets, the Arrhenius parameter optimization yields a temperature dependence for k_6 and k_7 which is inconsistent with theory [96, 97]. It can be hypothesized that these differences are rooted in: first, the selection of experimental data used as targets and second the way that theoretical calculations are included as targets.

In regard to the first, many targets used in Arrhenius parameter optimization [85] exhibit (neglected) structural uncertainties comparable to parametric uncertainties; these structural uncertainties can include chemically termolecular reactions [45], prompt dissociation [47], mixture rules [78] and missing collision efficiencies [110]. Using targets influenced by structural uncertainties would then impose un-physical constraints on the model parameters. Inspection of the covariance matrix presented [85] reveals that k_6 and k_7 parameters are strongly correlated to several other reactions whose optimized rate constants are either at the edge of prior uncertainties, have very different temperature dependence than the prior model, or are inconsistent with theoretical calculations.

In regard to the second, direct inclusion of theoretical calculations into the MSI model structure ensures that optimized rate constants are inherently consistent with theory. Altogether, the present use of dilute experiments minimizes structural uncertainties in multi-component reactive mixtures [45, 78, 110] and direct incorporation of theory into the MSI framework minimizes the potential that kinetic parameters will be optimized to un-physical values.

As an independent test of the MSI model, predictions using the MSI model were compared to

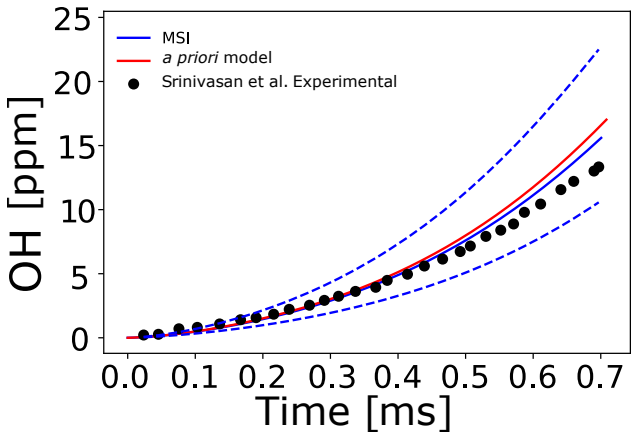


Figure 4: OH time-history in shock heated O_2/CH_4/Kr mixture near 1795 K and 0.948 atm. Symbols represent experimental data from Srinivasan et al. [96]; lines model predictions.

the OH profile of Srinivasan et al. [96] (Fig. 4). This measurement was not included as an optimization target because although it was used to make determinations for the rate constant of CH_3 + HO_2 = CH_4 + O_2, and exhibits some sensitivity to parameters describing this rate constant, after uncertainty weighted sensitivity analysis it was revealed that largest contributors to uncertainty for this observable are parameters describing other reactions. However, the optimized MSI model reproduce the OH time profile within uncertainties. Validating the model with independent test sets is a way to confirm all of the optimized kinetics parameters are physically realistic values.

4.5 Concluding Remarks: CH_3 + HO_2

An MSI analysis focused on the CH_3 + HO_2 reaction is performed – leveraging our previous analysis of pure H_2O_2 [69] for analysis of H_2O_2/CH_4 experiments. Results and discussion highlight the key elements of any approach to developing a high-accuracy kinetic database, including: the importance of raw data for quantifying the information content of experimental data, and the utility of theoretical kinetics calculations for constraining experimental interpretations and providing an independent data source and finally, the subtleties of target data selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties.

The studies in the next sections will focus on more highly studied systems in order to achieve sufficient data redundancy to test the validity of the optimized values, and quantified uncertain-

ties. The uncertainty weighted sensitivity results here give an initial indication of the difficultly in achieving data redundancy in kinetics – such that a large web of interconnected systems may be required. Specifically in the third case study, exploring branching reactions, the validity of using MSI to explore a large web of interconnected reactions to to help achieve data redundancy will be tested and discussed.

Chapter 5: Resolving Discrepancies Between State-of-the-art Theory and Experiment for HO_2 + HO_2 via Multiscale Informatics

5.1 Scientific Motivation: HO_2 + HO_2

The self reaction of HO_2 radicals ($HO_2 + HO_2 = H_2O_2 + O_2$) (R1) has been extensively studied across a variety of temperature and pressure conditions due to its importance in both combustion and the Earth's atmosphere. In combustion, R1 is known to be important for accurate predictions of flames [111, 42, 112] and ignition [113, 37] of many fuels, particularly at high pressures. In the Earth's atmosphere, R1 plays an important role in HO_x budgets—with implications for formation rates of H_2O_2, O_3, and organic peroxides [114] and oxidation of inorganic and organic pollutants [115].

At higher temperatures relevant to combustion, R1, a radical-radical reaction, has historically proven difficult to isolate experimentally. Experimental determinations of its rate constant are often entangled with those for another radical-radical reaction $OH + HO_2 = H_2O + O_2$ (R5, cf. Table 1), which is also important to high-pressure flames [111, 42, 112] and similarly difficult to isolate experimentally. Consequently, rate constants for both of these reactions were largely uncharacteristic at combustion-relevant temperatures (aside from rough estimates from flame modeling studies) until the late 20^{th} century.

Specifically, in 1995 and 2002, Troe and co-workers [128, 105] reported the first rate constants of the two reactions at combustion-relevant temperatures by taking advantage of the fact that OH and HO_2 are two of the main intermediates in the H_2O_2 decomposition system. The rate constants were determined by fitting rate parameters within a four-reaction kinetic model to UV laser absorbance time profiles, which provided a combined measure of HO_2 and H_2O_2, in shock-heated H_2O_2 mixtures. Both of these sets of rate coefficients exhibited unusually deep minima at 1000-

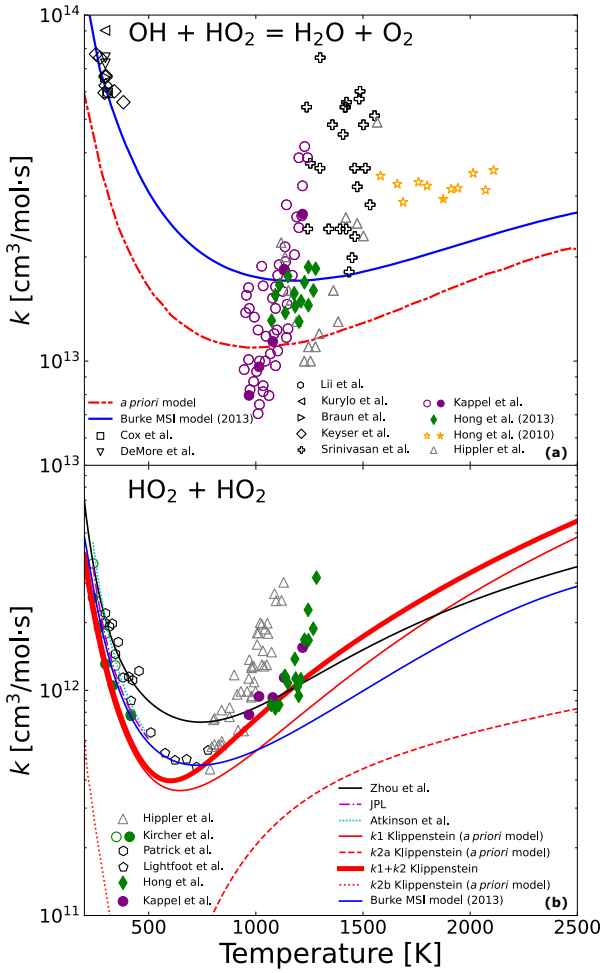


Figure 1: Rate constants at 1 atm for (a) HO_2 + OH = H_2O + O_2 (R5), (b) HO_2 + HO_2 = H_2O_2 + O_2 (R1) and HO_2 + HO_2 = products (R1 + R2). Symbols represent experimental determinations [116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 104] as indicated in the legend. Solid symbols specifically designate the experimental determinations of Hong et al. [106], Kappel et al. [105], and Kircher et al. [107] that correspond to raw experimental data in Figures 5, 4, and 15 (where the MSI model reproduces the raw data). Solid and dashed lines represent various model predictions, as indicated in the legend [69, 113, 71, 127], with the exception of blue dashed lines in (a) which represent uncertainties. (Note: given that R2 was not considered in the original experimental interpretations, the total rate constant k_1+k_2 for the experimental determinations simply corresponded to the reported k_1 values).

1200 K followed by a sharp rise with increasing temperature (see Fig. 1). Later modeling studies would show that the existence, exact value, and exact temperature of these minima had significant implications for high-pressure flame predictions [111, 42, 112].

Then, in 2006 and 2010, two determinations by Srinivasan et al. [126] and Hong et al. [104]

based on reflected shock experiments with initial reactants other than H_2O_2 provided independent evaluations of R5 at high temperatures. Their determined rate constants instead showed little temperature dependence and, in fact, were remarkably consistent with a simple constant-activation-energy extrapolation from the low-temperature experimental data. These determinations led many to suspect that R5 did not exhibit a minimum at all, though drawing any definitive conclusions about the rate constant for R5 (or R1) at intermediate to high temperatures remained difficult.

In 2012, two independent studies on R5 were presented at the 34th Combustion Symposium that used novel techniques to shed new light on the debate: one based on a MultiScale Informatics analysis of all previous theoretical and experimental data from Burke et al. [69] and another based on multi-species measurements of shock-heated H_2O_2 mixtures from Hong et al. [106]. Both of these studies found rate constants that exhibited minima but showed a much milder temperature dependence. Furthermore, the rate constants proposed in each study were in essentially perfect agreement at intermediate temperatures (see Fig. 1 a), thus appearing to resolve the previous long-standing debate—at least for R5.

On the other hand, for R1, the multi-species measurements from Hong et al. [106] appeared to uphold the rate constants from Kappel et al. [105] that showed a sharp rise with increasing temperature. By contrast, the MSI rate constants from Burke et al. [69] showed a milder temperature dependence, but the theoretical calculations for R1 (on which the MSI rate constants were partially based) were suspected to be affected by uncertainties in the treatment of anharmonicities, whose impact would increase with increasing temperatures. However, more recent theoretical calculations from Zhou et al. [127] and Klippenstein et al. [113], both of which employ advanced treatments of anharmonicity, also show a milder temperature dependence—suggesting that inconsistencies still remain among theoretical and experimental data for R1 (see Fig. 1 b). Furthermore, calculations of Klippenstein et al. [113] indicate a previously unknown channel $HO_2 + HO_2 = O_2 + 2OH$ (R2b), which appears to be prominent at higher temperatures and may therefore confound all earlier analyses.

These discrepancies in the temperature dependence in k_1 and the difference of a factor of

three in k_1 (which greatly exceeds the estimated uncertainties [106] in the experimentally derived k_1) determined by Hong et al. [106] and calculated by Klippenstein et al. [113] are particularly noteworthy, given that their respective experimental and theoretical methods are a tour de force in rate constant quantification techniques. For example, Hong et al. [106] simultaneously measured H_2O, OH, and absorbance at 227 nm (providing a combined measure of HO_2 and H_2O_2) with microsecond time resolution following reflected shock waves at precisely defined thermodynamic conditions. As they indicate, these time-resolved measurements for three species, together with element balance, would appear to completely capture the evolution of the four key chemical species involved in the H_2O_2 decomposition mechanism (H_2O_2, OH, HO_2, and H_2O). Likewise, Klippenstein et al. [113] not only characterized the torsional dependence of the vibrational frequencies (similar to Zhou et al. [115]) but also included a multi-reference-based variable reaction coordinate transition state theory (VRC-TST) treatment of the barrierless channels and high-level coupled-cluster-based TST treatment of the saddle points. Notably, all stationary points were characterized using the ANL0 method [129], which is among a collection of recent high-level *ab initio* schemes [130, 131, 132, 133] capable of 0.2 kcal/mol accuracies in energy calculations.

Altogether, resolution of the discrepancies at high temperatures is therefore important to both (1) accurate modeling of high-pressure ignition and flame propagation and (2) understanding the limitations and capabilities of state-of-the-art methods for quantifying rate constants.

Naturally, in light of the above discussion, one might wonder about the following questions. For example, can the discrepancy between theoretically calculated and experimentally derived rate constants at high temperatures be attributed to a new reaction channel (R2) not considered in the original experimental interpretations? Alternatively, can the discrepancy be attributed to remaining uncertainties in even state-of-the art theory? Or finally, can the discrepancy be attributed to remaining ambiguities in experimental interpretations of even state-of-the-art multi-species measurements of complex multi-reactions systems?

Compared to high temperatures, at lower temperatures relevant to the Earth's atmosphere, R1 is easier to isolate experimentally and has correspondingly been characterized much more precisely.

For example, many of the experimental studies (e.g. Refs [120, 122, 119, 121, 123, 124, 125, 107, 116, 117]) involve flash photolysis to generate HO_2 followed by highly sensitive, time-resolved laser absorbance measurements under very dilute conditions [134]. Among these, Kircher et al. [107] presented what is known to be one of the most extensive data sets to date at atmospheric conditions—with measurements spanning a wide range of temperatures (240-417 K), pressures (79-700 torr), and bath gases (Ar, N_2). Their experimental determinations of k_1 are generally in good agreement with other experimental determinations [116, 117] (with differences of less than $\sim$25%). Notably, experimental determinations over varied pressure (including those of Kircher et al. [107]) show a pressure dependence of k_1, which has been usually attributed to a weakly bound HO_2..HO_2 intermediate complex mediating the reaction [135, 136].

However, the recent theoretical calculations by Klippenstein et al. [113] reveal the presence of weakly bound and *chemically bound* complexes on the triplet and singlet surfaces. Their master equation equation calculations indicate that stabilization (or partial stabilization) of each of these complexes contributes significantly to the pressure dependence at low temperatures. While the weakly bound complexes are likely to decompose on timescales much faster than experimental timescales, the chemically bound HOOOOH complex could have relatively slower decomposition timescales that could obfuscate k_1 determinations from the measured absorbance time profiles. Depending on the how much HOOOOH is formed, whether it decomposes back to reactants, and how fast it decomposes, its evolution may be of some influence to experimental interpretations of the measured absorbance time profiles.

Therefore, the objectives of the present work are two-fold. The primary objective is to achieve resolution of the remaining inconsistencies among theoretical calculations and experimental determinations for k_1 at the highest temperatures – relevant to both combustion modeling and understanding the capabilities and limitations of state-of-the-art theory and experiments – with a focus on the questions outlined above. A secondary objective is to characterize the impact (if any) of the chemically bound intermediate complexes in the low-temperature experiments. Here, we present a comprehensive analysis of data from a wide variety of sources for the relevant reaction systems us-

ing the MultiScale Informatics approach [112], which is a multi-physics uncertainty quantification framework capable of simultaneously considering theoretical and experimental data on equal footing according to their respective uncertainties. We then demonstrate that the present MSI model is capable of explaining all theoretical and experimental data—including data across the full temperature range from Kappel et al. [105] and Hong et al. [106] and data across the full temperature and pressure range from Kircher et al.[107]—thereby resolving the abovementioned discrepancies among theoretically calculated and experimentally derived k_1 at high temperatures. We then use this model for subsequent analyses to ascertain the role of new production channels in experimental interpretations and, more generally, uncover the origins of the *apparent* discrepancies among theoretical and experimental data.

5.2 Implementation: HO_2 + HO_2

The MSI analysis focused on the reactions in the H_2O_2 decomposition mechanism (R1-R6 in Table 2), with particular attention to the HO_2 + HO_2 reaction. The nominal rate constants used for each of these reactions are based on theoretical calculations [113, 137] (Harding and Klippenstein, Unpublished) and associated active kinetic parameters were assigned to represent their uncertainties as discussed further below. The overall kinetic model used for data analysis included the reactions listed in Table 2 along with other "secondary" reactions that are (or may be) involved in interpretations of the raw experimental data. The nominal rate constants and uncertainty factors for these secondary reactions were taken from the published and previously optimized Foundational Fuel Chemistry Model (FFCM) [84] for all other H_2/C_1-C_2 oxidation kinetics, the JPL evaluation [138] for a selected list of chlorine reactions for modeling radical precursor chemistry (cf. Supporting Information), and a H_2 oxidation model from Konnov [139] for all other electronically excited species kinetics (cf. Supporting Information). The pre-exponential factors in the rate constant expressions for each of these secondary reactions were assigned as active parameters. However, the results shown below are virtually identical when the analysis is performed with or without realistic uncertainties (as reported in Refs. [84, 139, 138])—indicating that uncertainties in all

Table 1: List of parameters used in the optimization.[a]

Reaction		Kinetic Parameters
R1	$HO_2 + HO_2 = H_2O_2 + O_2$	$\sigma'_{Z(1)}$, $n_{\Delta E_d(1)}$, $A'_{\Delta E_d(1)}$, $\epsilon'_{\Delta E_d(1)}$, $\nu'_{mr(1)}$, $\eta'_{mr(1)}$ $\nu'_{\mathrm{B1(1)}}$, $\mathrm{E_{B1(1)}}$, $\mathrm{E_{P1(1)}}$, $E_{\mathrm{W1a(1)}}$, $\rho'_{\mathrm{corr\ W1b(1)}}$, $\rho'_{\mathrm{corr\ W1a(1)}}$, $\nu'_{\mathrm{W1a(1)}}$, $f'_{\mathrm{VRC\text{-}TST,c\ B1(1)}}$ $E_{\mathrm{W1b(1)}}$, $\nu'_{\mathrm{imag\ B2(1)}}$, $\nu'_{\mathrm{P1(1)}}$, $f'_{\mathrm{VRC\text{-}TST,c\ P1(1)}}$, $\nu'_{\mathrm{P2(1)}}$, $\nu'_{\mathrm{B8(1)}}$, $\rho'_{\mathrm{corr\ B8(1)}}$ $E_{\mathrm{P2(1)}}$, $\nu'_{\mathrm{W1b(1)}}$, $E_{\mathrm{B8(1)}}$, $E_{\mathrm{B2(1)}}$, $E_{\mathrm{P1(1)}}$
R2a	$HO_2 + HO_2 = O_2 + 2OH$	$\sigma'_{Z(2)}$, $n_{\Delta E_d(2)}$, $A'_{\Delta E_d(2)}$, $\epsilon'_{\Delta E_d(2)}$
R2b	$HO_2 + HO_2 = HOOOOH$	$E_{\mathrm{B4(2)}}$, $\nu'_{\mathrm{W2b(2)}}$, $\nu'_{\mathrm{B7_{1(2)}}}$, $\nu'_{\mathrm{B9(2)}}$, $f'_{\mathrm{VRC\text{-}TST,c\ B5(2)}}$, $\nu'_{\mathrm{B4(2)}}$
R2c	$HO_2 + HO_2 = H_2O + O_3$	$f'_{\mathrm{VRC\text{-}TST,c\ B7(2)}}$, $\nu'_{\mathrm{W2a(2)}}$, $E_{\mathrm{W2b(2)}}$, $\nu'_{\mathrm{imag\ B3(2)}}$, $f'_{\mathrm{VRC\text{-}TST,c\ P1(2)}}$
R2d	$HO_2 + HO_2 = H_2O_2 + O_2(X)$	$\mathrm{E_{B3(2)}}$, $E_{\mathrm{B7(2)}}$, $E_{\mathrm{P4(2)}}$, $\nu'_{\mathrm{P3(2)}}$, $\rho'_{\mathrm{corr\ B9(2)}}$
R2e	$H_2O_2 + O_2(X) = HOOOOH$	$\nu'_{\mathrm{B3(2)}}$, $\mathrm{E_{B6(2)}}$, $\nu'_{\mathrm{B6(2)}}$, $\nu'_{\mathrm{P5(2)}}$, $E_{\mathrm{W3(2)}}$,
R2f	$H_2O_2 + O_2(X) = H_2O + O_3$	$E_{\mathrm{P3(2)}}$, $\nu'_{\mathrm{P1(2)}}$, $\rho'_{\mathrm{corr\ W2a(2)}}$, $\rho'_{\mathrm{corr\ P5(2)}}$, $\nu'_{\mathrm{P4(2)}}$, $\nu'_{\mathrm{W3(2)}}$
R2g	$H_2O_2 + O_2(X) = O_2 + 2OH$	$E_{\mathrm{P5(2)}}$, $E_{\mathrm{B5(2)}}$, $\rho'_{\mathrm{corr\ W2b(2)}}$, $\nu'_{\mathrm{imag\ B6(2)}}$, $E_{\mathrm{W2a(2)}}$,
R2h	$H_2O + O_3 = HOOOOH$	$E_{\mathrm{P1(2)}}$, $\eta'_{\mathrm{B7_1(2)}}$, $\eta'_{\mathrm{B7_2(2)}}$, $\eta'_{W3_1(2)}$, $\eta'_{W3_2(2)}$, $\eta'_{\mathrm{W3_3(2)}}$, $\eta'_{\mathrm{P5(2)}}$
R2i	$H_2O + O_3 = O_2 + 2OH$	$E_{\mathrm{B7_1(2)}}$, $f'_{\mathrm{VRC\text{-}TST,c\ B7_2(2)}}$, $\nu'_{\mathrm{B7_{2(2)}}}$
R2j	$O_2 + 2OH = HOOOOH$	$f'_{\mathrm{VRC\text{-}TST,c\ B4(2)}}$, $E_{\mathrm{B9(2)}}$, $E_{\mathrm{P1(2)}}$, $E_{\mathrm{W2a(2)}}$, $\nu'_{\mathrm{B5(2)}}$
R3	$H_2O_2 + OH = H_2O + HO_2$	$E_{\mathrm{B1(3)}}$, $\nu'_{\mathrm{B1(3)}}$, $\eta'_{\mathrm{B1(3)}}$, $\nu'_{\mathrm{imag\ B1(3)}}$
R4	$OH + OH = H_2O + O$	$E_{\mathrm{B1(4)}}$, $E_{\mathrm{B2(4)}}$, $\nu'_{\mathrm{B1_1(4)}}$, $\nu'_{\mathrm{B1_2(4)}}$
R5a	$OH + HO_2 = H_2O + O_2$	$f'_{\mathrm{VRC\text{-}TST,c\ B1(5)}}$, $E_{\mathrm{B2(5)}}$, $\nu'_{\mathrm{B2(5)}}$, $\eta'_{\mathrm{B2(5)}}$, $\nu'_{\mathrm{imag\ B2(5)}}$, $E_{\mathrm{B3(5)}}$
R5b	$OH + HO_2 = H_2O + O_2(X)$	$\nu'_{\mathrm{B3(5)}}$, $\eta'_{\mathrm{B3(5)}}$, $\nu'_{\mathrm{imag\ B3(5)}}$, $\nu'_{\mathrm{W1(5)}}$
R6	H_2O_2 (+M) = OH + OH (+M)	$A'_{(6)}$, $n_{(6)}$, $Ea_{(6)}$

Macroscopic Observables		Physical Model Parameters	
E1-E3	Shock-heated $H_2O_2/H_2O/O_2/Ar$	T'_e, P'_e, $M'_{\mathrm{H_2O_2},o,e}$, $M'_{\mathrm{H_2O},o,e}$, $M'_{\mathrm{O_2},o,e}$	$e = 1...3$
E4	Shock-heated $H_2O/O_2/H/Ar$	T'_e, P'_e, $M'_{\mathrm{H},o,e}$, $M'_{\mathrm{H_2O},o,e}$, $M'_{\mathrm{O_2},o,e}$	$e = 4$
E5-E10	Shock-heated H_2O_2/Ar	T'_e, P'_e, $M'_{\mathrm{H_2O_2},o,e}$, $\sigma'_{1,\mathrm{H_2O_2}}$, $\sigma'_{2,\mathrm{H_2O_2}}$, $\sigma'_{1,\mathrm{HO_2}}$, $\sigma'_{2,\mathrm{HO_2}}$	$e = 5...10$
E11-E25	Shock-heated $H_2O_2/H_2O/O_2/Ar$	T'_e, P'_e, $M'_{\mathrm{H_2O_2},o,e}$, $M'_{\mathrm{H_2O},o,e}$, $M'_{\mathrm{O_2},o,e}$, $\hat{\sigma}'_{1,\mathrm{H_2O_2}}$, $\hat{\sigma}'_{2,\mathrm{H_2O_2}}$, $\hat{\sigma}'_{1,\mathrm{HO_2}}$, $\hat{\sigma}'_{2,\mathrm{HO_2}}$	$e = 11...25$
E26-E35	Flash-photolyzed $Cl/Cl_2/CH_3OH/O_2/Ar$	T'_e, P'_e, $M'_{\mathrm{Cl},o,e}$, $M'_{\mathrm{Cl_2},o,e}$, $M'_{\mathrm{CH_3OH},o,e}$, $M'_{\mathrm{O_2},o,e}$, $\hat{\sigma}'_{1,\mathrm{H_2O_2}}$, $\hat{\sigma}'_{2,\mathrm{H_2O_2}}$, $\hat{\sigma}'_{1,\mathrm{HO_2}}$, $\hat{\sigma}'_{2,\mathrm{HO_2}}$	$e = 26...35$

[a]Note that $'$ indicates ln() of that quantity.

other secondary reactions (and the fact that FFCM has been previously optimized) appear to have a negligible impact on the present interpretations and conclusions.

The theoretical kinetics calculations of Klippenstein et al. [113] and Harding and Klippenstein (Unpublished) were used for the active kinetic treatment of the HO_2 + HO_2 reactions (R1 and R2) and other H_2O_2 decomposition system reactions (R3-R5), respectively. Active parameters were assigned for each of the following: energies of each stationary point (E), scaling factors for all harmonic frequencies of each stationary point (ν), scaling factors for each hindered rotor potential (η),

scaling factors for the imaginary frequency used in the Eckart tunneling correction for each tight transition state (ν_{imag}), state density correction factors to account for any significant uncertainties in the treatment of anharmonicity of some stationary points (ρ_{corr}), VRC-TST correction factors to account for uncertainties in the treatment of barrierless transition states ($f_{\text{VRC-TST}}$), and scaling factors for the parameters describing a temperature-dependent average energy per down collision, $\Delta E_d = AT^n$, in the exponential-down model ($A_{\Delta E_d}$ and $n_{\Delta E_d}$) and the cross section of the well in the Lennard-Jones collision model (σ_Z) to account for uncertainties in the treatment of collisional energy transfer. For B2 in R1, for which the torsional dependence of the harmonic frequencies was considered, the same harmonic frequency scaling factor was applied to all frequencies at all torsional angles. For R6, whose rate constant was sufficiently constrained by macroscopic observables over the relevant temperature ranges, Arrhenius parameters were simply treated as active parameters (similar to our earlier work [69, 72]).

Theoretical calculations of HO_2 + HO_2 reactions (R1 and R2) indicate that the weakly bound complexes (W1a, W1b, W2a, W2b) decompose and/or isomerize sufficiently quickly (faster than 10^6 s^{-1}) that they are quasi-steady-state species on experimentally relevant timescales at all experimental conditions considered and, in fact, do not exist as chemical species under many conditions [81] [140, 141]. By contrast, the chemically bound HOOOOH complex decomposes on distinct, experimentally relevant timescales under some conditions such that it is important to consider HOOOOH among the resolved chemical species [142]. At other conditions, however, HOOOOH reacts on timescales comparable to internal energy relaxation such that it is not a well-defined chemical species. In light of the above discussion, HOOOOH is the only complex considered among the chemical species in the present phenomenological description of R1 and R2, where rate constants are obtained across all temperatures and pressures (including those where HOOOOH reacts on internal energy relaxation timescales) using the Well Extension method in MESS [81].

Physical model parameters from experiments included as macroscopic observable targets were also among the active model parameters. These parameters included initial temperatures, T_e, initial pressures, P_e, initial species mole fractions, $M_{i,o,e}$, and coefficients for temperature-dependent

cross sections of H_2O_2 and HO_2 at 215 nm (σ_{1,HO_2}, σ_{2,HO_2}, σ_{1,H_2O_2}, σ_{1,H_2O_2}) and 227 nm ($\hat{\sigma}_{1,HO_2}$, $\hat{\sigma}_{2,HO_2}$, $\hat{\sigma}_{1,H_2O_2}$, $\hat{\sigma}_{2,H_2O_2}$).

Rate constants and brute-force sensitivity coefficients were calculated using an automated wrapper [86] for Variflex [82] (for R3-R5) and MESS [81] (for R1-R2). Macroscopic observables and sensitivity coefficients were calculated using homogeneous, isochoric or isobaric models in Cantera [83].

Table 2: List of targets used in the optimization.[a,b]

I. *Ab initio* calculations			Source
$\sigma'_{Z(1)}(1.1)$, $n_{\Delta E_d(1)}(0.2)$, $A'_{\Delta E_d(1)}(1.1)$, $\epsilon'_{\Delta E_d(1)}(1\text{x}10^{-10})$ $\sigma'_{Z(2)}(1.1)$, $n_{\Delta E_d(2)}(0.2)$, $A'_{\Delta E_d(2)}(1.1)$, $\epsilon'_{\Delta E_d(2)}(1\text{x}10^{-10})$, (See Fig. 2 for remainder of parameter uncertainties for R1 and R2.) $E_{B1(3)}(3.0$ kcal/mol), $\nu'_{B1(3)}(0.1)$, $\eta'_{B1(3)}(0.1)$, $\nu'_{\text{imag B1(3)}}(0.18)$ $E_{B1(4)}(1.4$ kcal/mol), $E_{B2(4)}(1.4$ kcal/mol), $\nu'_{B1(4)}(0.1)$, $\nu'_{\text{imag B1,B2(4)}}(0.18)$ $f'_{\text{VRC-TST,c B1(5)}}(0.7)$, $E_{B2(5)}(2$ kcal/mol), $\nu'_{B2(5)}(0.1)$, $\eta'_{B2(5)}(0.41)$, $\nu'_{\text{imag B2(5)}}(0.18)$, $E_{B3(5)}(2$ kcal/mol) $\nu'_{B3(5)}(0.1)$, $\eta'_{B3(5)}(0.41)$, $\nu'_{\text{imag B3(5)}}(0.18)$, $\nu'_{W1(5)}(0.01)$			Klippenstein [113], (Harding and Klippenstein Unpublished)
II. Rate constant measurements			
k'_3, k'_4, k'_{5a}+k'_{5b} (see text)			[69, 100]
III. Global exp.	IV. Exp. conditions		
$M'_{OH,e}(t)(0.05)$, $M'_{H_2O,e}(t)(0.05)$	$T'_e(0.01)$, $P'_e(0.02)$, $M'_{H_2O_2,o,e}(0.05)$, $M'_{H_2O,o,e}(0.05)$, $M'_{O_2,o,e}(0.05)$	$e = 1...2$	[102]
$M'_{H_2O,e}(t)(0.05)$	$T'_e(0.01)$, $P'_e(0.02)$, $M'_{H_2O,o,e}(0.1)$, $M'_{O_2,o,e}(0.01)$ $M'_{H_2O_2,o,e}(0.01)$,	$e = 3$	[103]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01)$,$P'_e(0.02)$, $M'_{H_2O,o,e}(0.1)$, $M'_{O_2,o,e}(0.01)$, $M'_{H,o,e}(2.3)$	$e = 4$	[104]
$\text{abs}'_e(t)(0.1)$	$T'_e(0.02)$, $P'_e(0.04)$, $M'_{H_2O_2,o,e}(0.1)$, $\sigma'_{1,H_2O_2}(0.7)$, $\sigma'_{2,H_2O_2}(0.3)$, $\sigma'_{1,HO_2}(0.7)$, $\sigma'_{2,HO_2}(0.3)$	$e = 5...10$	[105]
$\text{abs}'_e(t)(0.1)$, $M'_{H_2O,e}(t)(0.05)$, $M'_{OH,e}(t)(0.05)$,	$T'_e(0.02)$, $P'_e(0.04)$, $M'_{H_2O_2,o,e}(0.05)$, $M'_{H_2O,o,e}(0.2)$, $M'_{O_2,o,e}(0.2)$, $\hat{\sigma}'_{1,H_2O_2}(0.7)$, $\hat{\sigma}'_{2,H_2O_2}(0.7)$, $\hat{\sigma}'_{1,HO_2}(0.7)$, $\hat{\sigma}'_{2,HO_2}(0.7)$	$e = 11...25$	[106]
$\text{abs}'_e(t)(0.1)$	$T'_e(0.02)$, $P'_e(0.04)$, $M'_{Cl,o,e}(0.1)$, $M'_{Cl_2,o,e}(0.05)$, $M'_{CH_3OH,o,e}(0.05)$, $M'_{O_2,o,e}(0.05)$, $\hat{\sigma}'_{1,H_2O_2}(0.7)$, $\hat{\sigma}'_{2,H_2O_2}(0.7)$, $\hat{\sigma}'_{1,HO_2}(0.7)$, $\hat{\sigma}'_{2,HO_2}(0.7)$	$e = 26...35$	[107]

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

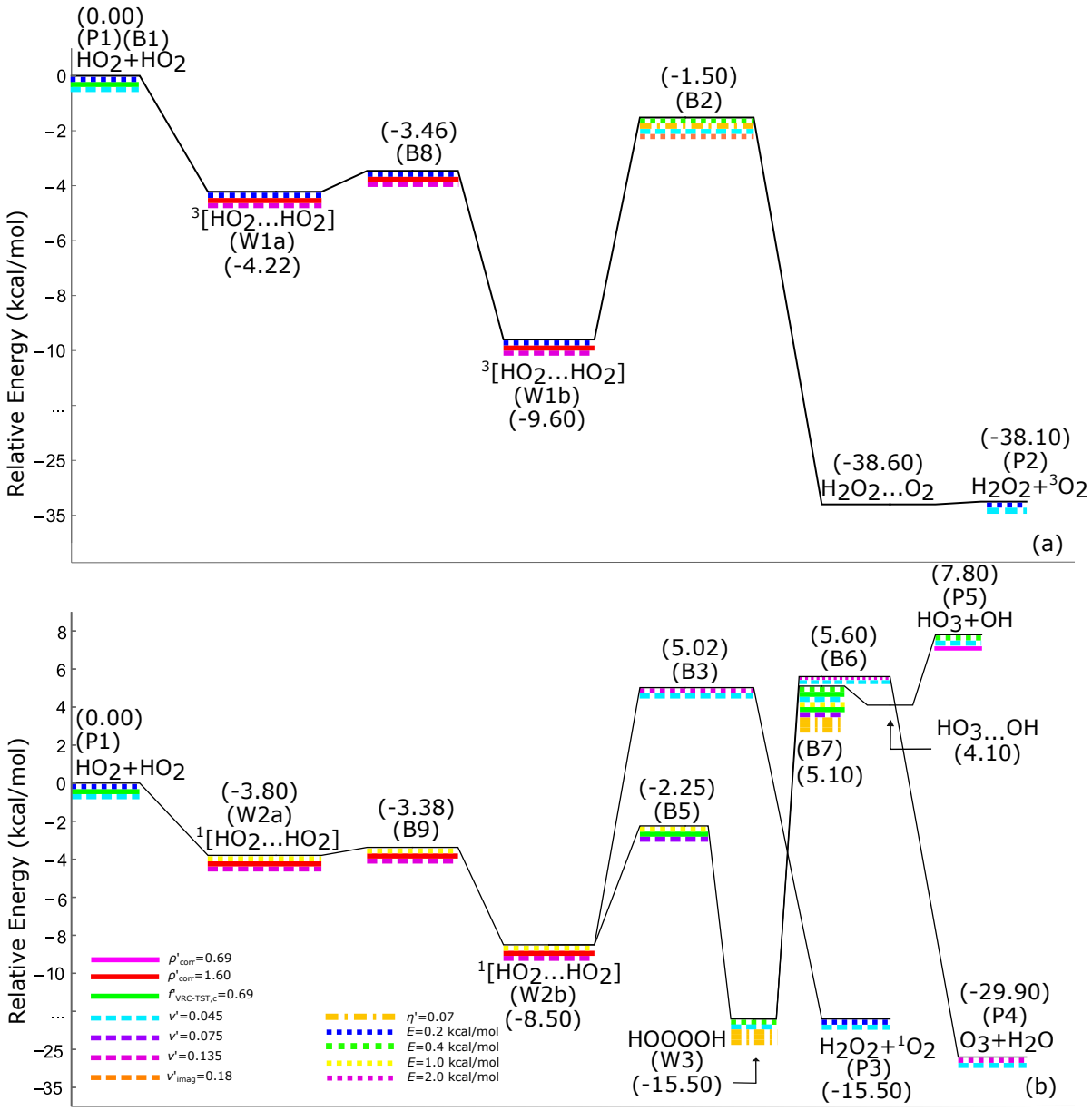


Figure 2: Potential energy surface for the (a) triplet and (b) singlet surfaces of HO_2 + HO_2. The colored lines refer to the uncertainty assigned to a specific barrier or well for a parameter as indicated in the legend.

Target class (I): *ab initio* calculations

Molecular properties from *ab initio* calculations were used as targets for the molecular parameters for R1-R2 [113] and R3-R6 (Harding and Klippenstein, Unpublished) [69]. Uncertainties for

each active parameter were chosen on the basis of the level of treatment used in each aspect of the theoretical calculations, as described in further detail below.

For R1-R2, as recommended in Klippenstein et al. [113], the energies of stationary points (E) calculated using the ANL0 method [129] were assigned uncertainties of 0.2 kcal/mol if the CCSDT(Q) corrections were less than $\sim$1 kcal/mol and CBS extrapolations were less than $\sim$0.5 kcal/mol (P1, W1a, B8, W1b, P2, $^3O_2$ + OH + OH), 0.4 kcal/mol if the CCSDT(Q) corrections or CBS extrapolations were close to those thresholds (B2, W4), and 2 kcal/mol if CCSDT(Q) corrections were much higher (B3, B6, P4). Energies calculated using CASPT calculations were assigned typical uncertainties of 1 kcal/mol [80] (though comparisons against Davidson-corrected MRCI calculations suggest that somewhat lower uncertainties may be possible for the calculations [113]). For P3 and P4, whose energies were obtained via experimental measurements of the $^1O_2/^3O_2$ splitting and HO_3 dissociation energy combined with ANL0 calculations of H_2O_2 + $^3O_2$ and $^3O_2$ + OH + OH, uncertainties in the energies were simply taken to representative of the uncertainties in the accompanying ANL0 calculations (0.2 and 0.4 kcal/mol, respectively).

For all stationary points calculated using the ANL0 method, the uncertainties in harmonic frequency scaling factors (ν) of $\sim$5% and hindered rotor potential scaling factors (η) of $\sim$8% were chosen to each yield propagated uncertainties of $\sim$20% in k_1 at high temperatures (on the basis of limited explorations of rovibrational treatment effects in Klippenstein et al. [113]). For all stationary points calculated using multi-reference calculations, larger uncertainties in harmonic frequency scaling factors (ν) of $\sim$8% were used. Higher harmonic frequency (ν) uncertainties of $\sim$15% and state density correction factors (ρ_{corr}) with uncertainties of a factor of 5 were used for each of the four weakly bond complexes (W1a, W1b, W2a, W2b), which are expected to have significant anharmonic effects but were treated with rigid-rotor harmonic oscillator approximations. For context, these uncertainties encompass the empirical adjustments of the state densities of a factor of 2 to 2.5 used in the theoretical calculations [113]. Consistent with our earlier work [69], uncertainties in the VRC-TST correction factors ($f_{\text{VRC-TST}}$) of a factor of 2 are used for all transition states treated with VRC-TST and uncertainties in the imaginary frequencies (ν_{imag}) of 25% are used for all tight

transition states.

For R3-R6, barrier height (E) uncertainties were chosen on the basis of the quoted errors for the various method/basis set combinations from a recent assessment [80]; uncertainties in the harmonic frequencies (ν) were chosen to yield similar total propagated harmonic frequency uncertainties as our earlier work [69] (which had multiple active parameters for harmonic frequency uncertainties); and the uncertainties in the hindered rotor potentials (η), imaginary frequencies (ν_{imag}), and VRC-TST correction factors ($f_{\text{VRC-TST}}$) were taken to be the same as our earlier work [69].

Target class (II): Rate constant determinations

While the most accurate way to capture the information content of kinetic experiments is to include the raw data as macroscopic observables, it is often convenient to include the derived data from some kinetic experiments as rate constant determinations instead (e.g. if the raw data are not available, the experimental determinations are reasonably direct, or the reactions are of lesser direct interest). Here, select rate constant determination data for k_3, k_4, and k_{5a} + k_{5b} (listed in Table 2) were used as targets. However, these data were assigned conservative uncertainties of a factor of two or the experimental scatter, if larger, based on the fact that rate constant determination uncertainties are often thought to be larger than common estimations and/or contain systematic errors [69, 71, 44].

Target classes (III) and (IV): Macroscopic observables and experimental conditions

The raw data for the macroscopic observables from the experiments listed in Table 2 were included as targets using reported uncertainties, if available, or typical values, otherwise [103, 102, 104, 105, 106, 107]. Macroscopic observables included as targets consisted of time-resolved H_2O mole fractions, OH mole fractions, and absorbance at 215 and 227 nm in shock-heated and flash-photolyzed mixtures. The reported values of the experimental conditions were imposed as targets for the physical model parameters using the reported uncertainties, if available, or typical values, otherwise [103, 102, 104, 106, 105, 107].

Among the low-temperature determinations of k_1, the experimental data of Kircher et al. [107] are of particular interest – given that the data span both temperature and pressure and often form the basis for atmospheric chemistry evaluations [138, 136]. In light of the questions raised in the Introduction, it is worth assessing what the main products of HO_2 + HO_2 are at atmospheric conditions and whether the timescales and fate of stabilized complexes have any bearing on the derived rate constants (in addition to any other secondary reactions). However, the raw data used for the experimental determinations of k_1 are not reported in the original paper. Therefore, to assess the potential influence of interpretation uncertainties (e.g. as related to the new reaction channels, R2), the present analysis makes use of a procedure to (approximately) regenerate the raw data from experimental rate constant determinations from the information reported. Overall, the procedure is similar in spirit to that of Najm and co-workers [143] and, in fact, our own limited testing suggests that this procedure yields posterior probability distribution functions that can be remarkably similar to their full Bayesian inference method [143].

For reference, in the experiments of Kircher et al. [107], Cl_2 was photolyzed in the presence of CH_3OH, O_2, and an Ar or N_2 bath to yield a sequence of reactions (Cl_2 = Cl + Cl, Cl + CH_3OH = HCl + CH_2OH, CH_2OH + O_2 = CH_2O + HO_2) that ultimately forms approximately one HO_2 per Cl. Absorbance at 227 nm, which provide a combined measure of HO_2 and H_2O_2, was then monitored as a function of time [107]. Using previously reported absorbance cross sections for HO_2 and H_2O_2, the rate constant for HO_2 + HO_2 was then fit to the data after the short period during which HO_2 is formed (under the assumption that HO_2 + HO_2 is the only reaction contributing to absorbance time profile) [107].

The specific concentrations of Cl_2, CH_3OH, O_2, and Cl for each experiment were not reported, but typical concentration ranges of Cl_2 (1.5 to 15 x 10^{15} molec cm^{-3}), CH_3OH (1.0 to 5.0 x 10^{15} molec cm^{-3}), and O_2 (1.4 to 4.0 x 10^{17} molec cm^{-3}) and the photolysis laser pulse energy (500-1000 J/flash) were reported and can be used to estimate the typical initial concentrations. In our regenerated raw data, we assume [Cl_2], [CH_3OH], and [O_2] to take values at the midpoint of each range. We then estimate the initial Cl concentration, $[Cl]_0$, to be equal to $[HO_2]_{g,L}$ x

$[Cl_2]_0/[Cl_2]_{0,L}$ x $E_p/E_{p,L}$, where $[Cl_2]_0$ and E_p are the initial Cl_2 concentration and laser pulse energy used in Kircher et al. [107] and $[HO_2]_{g,L}$, $[Cl_2]_{0,L}$, $E_{p,L}$ are the reported HO_2 concentration generated, initial Cl_2 concentration, and laser pulse energy used in another study by Lightfoot et al.[117] using the same experimental approach. Similarly, while the raw absorbance time profiles were not reported, the derived rate constants at each temperature and pressure were reported [107] such that the absorbance time profiles that result in a given k_1 value can be approximated. Using this estimated initial [Cl] and the same cross sections and assumptions as in Ref. [107], the original interpretation procedure [107] used in deriving k_1 is then reversed to recreate the absorbance time profiles that would have resulted in the reported k_1 being derived. Examples of these absorbance time profiles can be found in Fig. 15. Of course, since the actual raw data and complete experimental information are not available, the data generated in this manner are inherently approximate. Nevertheless, these data are still useful for assessing any potential systematic biases in experimental interpretations [143], including those related to the lifetimes of chemically bound reaction intermediates (HOOOOH) and/or other secondary reactions.

5.3 Results and discussion: HO_2 + HO_2

Overall, the values of and predictions using the optimized set of molecular parameters, rate parameters, and physical model parameters informed by the target data listed in Table 2 are consistent with the target data from *ab initio* calculations, rate constant determinations, measured macroscopic observables, and reported experimental conditions (as well as data from other experiments not included among the target data). In particular, as shown below, the MSI model is consistent with the theoretical calculations of Klippenstein et al. [113], four different macroscopic observables during H_2O_2 decomposition from Kappel et al. [105], Hong et al. [106], and Alquaity et al. [144] at high temperatures, and macroscopic observables following photolytic generation of HO_2 at low temperatures from Kircher et al. [107]—across the full temperature and pressure ranges of each of those studies.

5.3.1 High temperatures

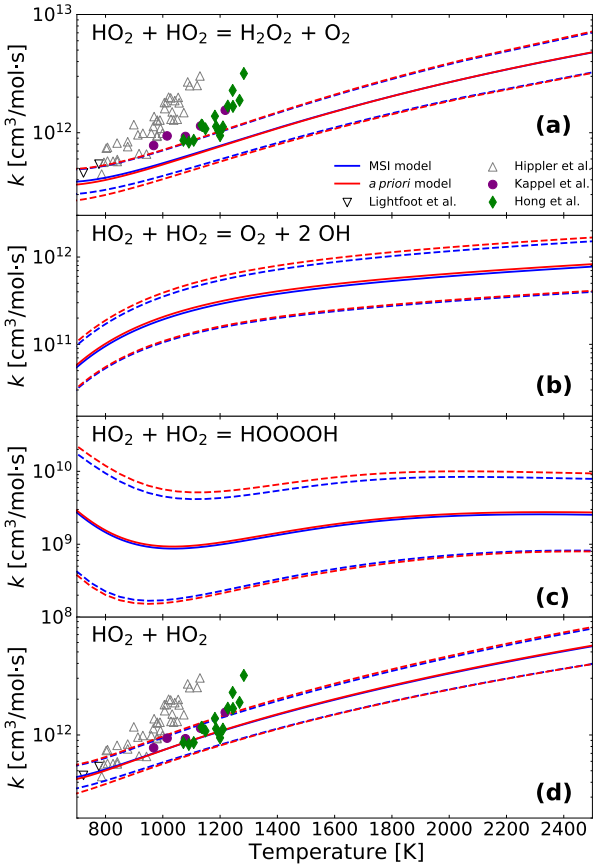


Figure 3: High-temperature rate constants at 1 atm for (a) $HO_2 + HO_2 = H_2O_2 + O_2$ (R1), $HO_2 + HO_2 = O_2 + 2OH$ (R2a), (c) $HO_2 + HO_2 = HOOOOH$ (R2b), (d) $HO_2 + HO_2$ = products (R1 + R2a-d). Symbols represent experimental determinations [118, 117, 106, 105] as indicated in the legend. Solid symbols specifically designate the experimental determinations of Hong et al. [106] and Kappel et al. [105] that correspond to raw experimental data in Figs. 4 and 5 (where the MSI model reproduces the raw data). Solid and dashed lines represent the values and uncertainties, respectively, for the *a priori* and MSI models. (Note: given that R2 was not considered in the original experimental interpretations, the total rate constants $k_1 + k_{2a\text{-}d}$ for the experimental determinations simply correspond to the reported k_1 values).

As shown in Fig. 3, the high-temperature rate constants (and uncertainties) for $HO_2 + HO_2$ in the MSI model are nearly identical to those in the *a priori* model, which are simply theoretically calculated rate constants from Klippenstein [113]. Furthermore, in addition to being completely consistent with the theoretical calculations [113], the MSI model is also consistent with macroscopic observables from the high-temperature k_1 determinations [105, 106] (Figs. 4 and 5) across

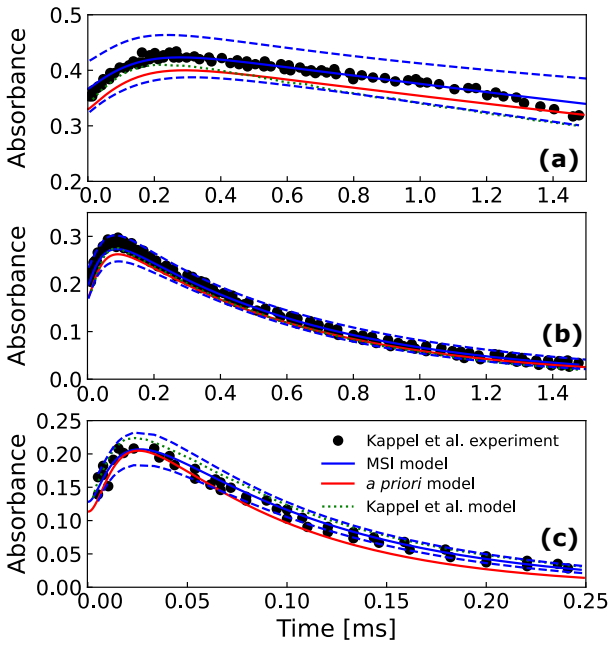


Figure 4: Time profiles of absorbance at 215 nm in shock-heated H_2O_2/Ar mixtures near (a) 968 K and 3.6 atm, (b) 1079 K and 3.3 atm, and (c) 1218 K and 3.0 atm. Symbols represent experimental data from Kappel et al. [105]; lines represent model predictions using the *a priori* model, MSI model, and the original model of Kappel et al. [105] derived from the experimental data shown.

the *entire* temperature ranges of each of those studies—thus implying that the theoretical and experimental data are actually entirely consistent, at least when considering the raw data from the experimental studies.

For example, Fig. 4 compares the experimental measurements [105] and model predictions using the *a priori* model, MSI model, and the original model derived from the measurements [105] for time profiles of absorbance at 215 nm at the lowest, middle and highest temperatures of Kappel et al. [105]. While the *a priori* model predicts lower absorbance than observed experimentally, the MSI model predictions for all three profiles are consistent with experimental data within uncertainties. Similarly, Fig. 5 compares the experimental measurements [106] and model predictions using the *a priori* model, MSI model, and the original model derived from the measurements [106] for time profiles of H_2O, OH, and absorbance at 227 nm at the highest and lowest temperatures of Hong et al. [106]. While the *a priori* model predicts higher OH and lower absorbance than observed experimentally, the MSI model predictions for all three profiles are consistent with ex-

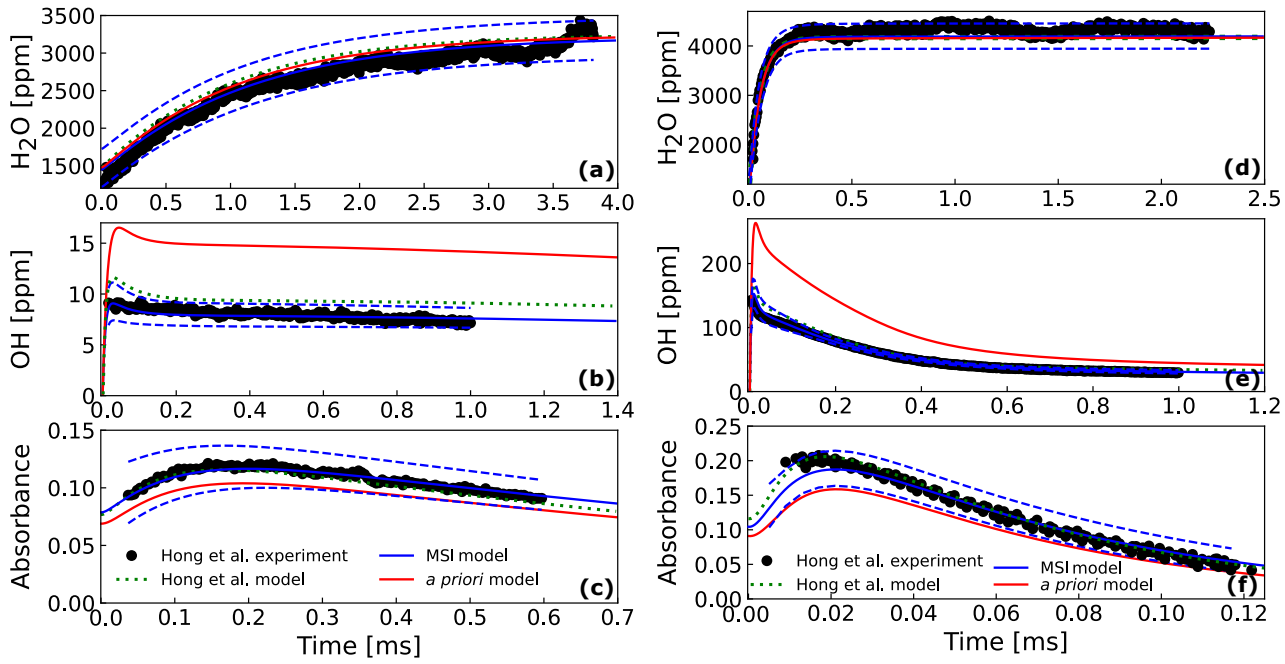


Figure 5: Time profiles of H_2O, OH, and absorbance at 227 nm in shock-heated $H_2O_2/H_2O/O_2/Ar$ mixtures near (a-c) 1073 K and 1.786 atm and (d-f) 1283 K and 1.635 atm. Symbols represent experimental data from Hong et al. [106]; lines represent model predictions using the *a priori* model, MSI model, and the original model of Hong et al. [106] derived from the experimental data shown.

perimental data within uncertainties. In both cases (Figs. 4 and 5), the MSI interpretations show at least equally good agreement with the measured time profiles as the original interpretations.

In fact, the same trend is observed across the entire multi-species measurement dataset [106]. In particular, as shown in Fig. 7, the mean and variance of the deviations between measured and predicted profiles are similar when using the MSI model or the original model derived from the data [106]—indicating that each appear to be an equally valid interpretation of the measured time profiles. However, as indicated above, the MSI model interpretation employs values for k_1 and k_2 that are nearly identical to the theoretical kinetics calculations [113] and is, therefore, completely consistent with both the theoretical calculations [113] and the raw experimental data [105, 106]—thereby resolving the *apparent* discrepancies among theoretically calculated [113] and experimentally derived [105, 106] k_1 at high temperatures.

Consequently, a series of additional analyses using this MSI model was performed to identify

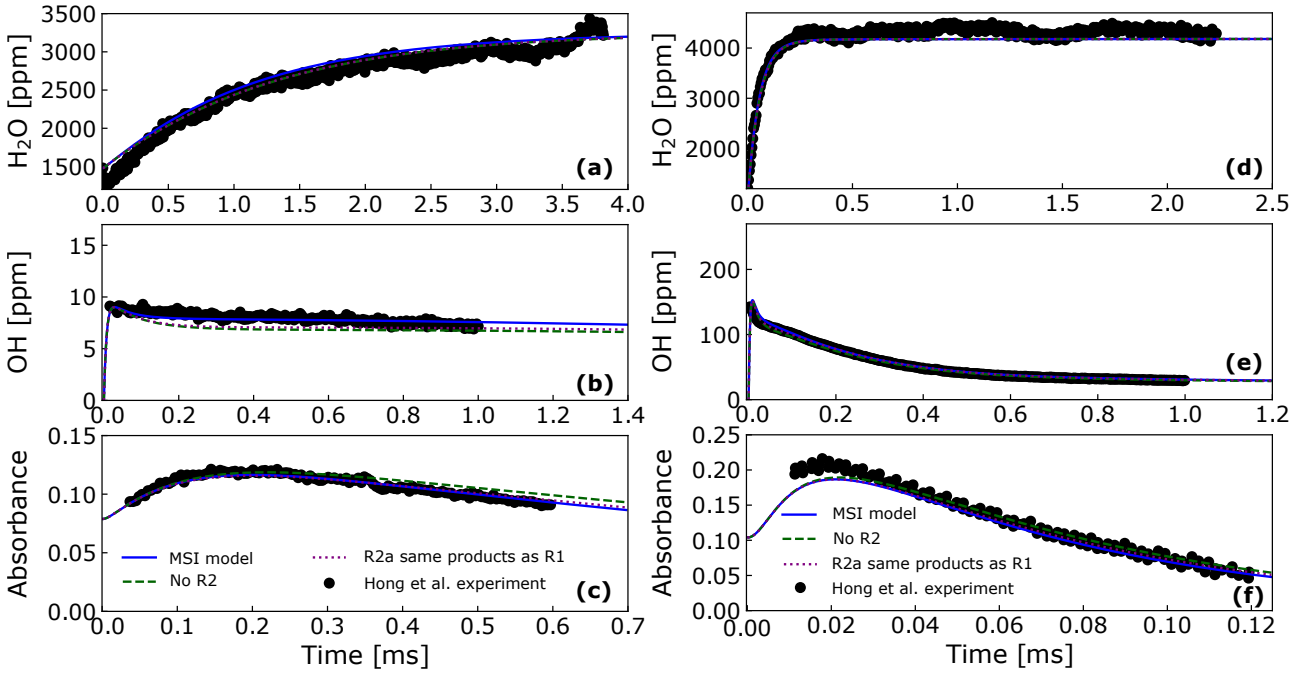


Figure 6: Time profiles of H_2O, OH, and absorbance at 227 nm in shock-heated $H_2O_2/H_2O/O_2/Ar$ mixtures near (a-c) 1073 K and 1.786 atm and (d-f) 1283 K and 1.635 atm. Symbols represent experimental data from Hong et al. [106]; lines represent model predictions varied treatment of R2.

an explanation for the apparent inconsistency between theory and experiment—through exploration of the questions outlined in the Introduction.

The first possibility explored is that the discrepancy between theoretically calculated and experimentally derived rate constants at high temperatures could be attributed to the new reaction channel (R2), which was identified in the recent theoretical kinetics calculations [113] but not considered in the original experimental interpretations conducted prior. To this end, Fig. 6 compares model predictions using the MSI model and two alternative versions of the MSI model, one in which R2 is completely excluded and another in which the optimized rate constant values for k_1 and k_{2a} are summed and exclusively attributed to k_1 (which is equivalent to assuming R2a were to instead produce the same products as R1). Note that, in the results shown, neither of the two modified MSI models were re-optimized against the data. These modifications yielded modest differences in the predicted profiles, but both of these modified models still yield predictions that

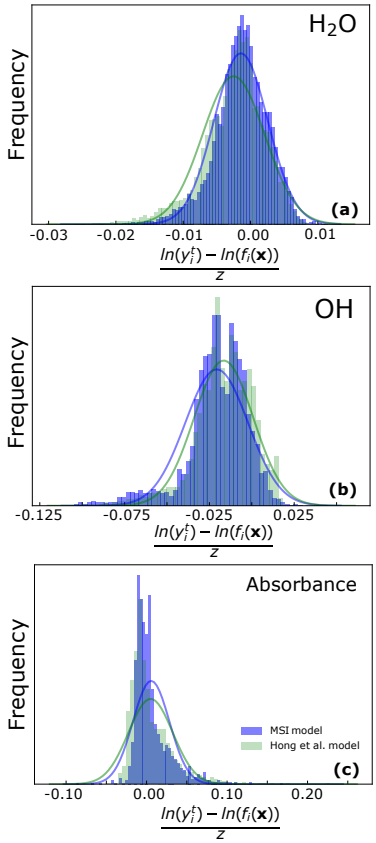


Figure 7: Histograms for the deviation of model predictions using the MSI model and the model of Hong et al.[106] from the measured time profiles for (a) OH, (b) H_2O, and (c) absorbance at 227 nm used that were originally used to derive the experimentally determined rate constants in the model of Hong et al. [106].

are consistent within the experimental data within uncertainties. Furthermore, these modest differences are even smaller at the higher temperature than at the lower temperature—such that the role of R2 is actually less pronounced at the highest temperatures where the discrepancies between theoretical and experimental determinations are largest. At the highest temperatures, the observed lack of sensitivity to the products of HO_2 + HO_2 can be attributed to the rapid timescale for H_2O_2 decomposition such that the difference between the R1 and R2a is effectively negated immediately, such that only the total rate constant $k_1 + k_{2a}$ has any noticeable influence. As indicated by the bottom panel of Fig. 3, the total rate constant $k_1 + k_{2a\text{-}d}$ in the MSI model still differs from the experimentally derived k_1. In fact, a separate analysis where the MSI model was re-optimized with R2 artificially excluded yields similar agreement with the measured observables

and reaches nearly identical values for k_1 (within 1%). Altogether, while the additional reaction channel has some modest influence on experimental interpretations, it does not appear to explain the discrepancy between theoretically calculated and experimentally derived rate constants at high temperatures.

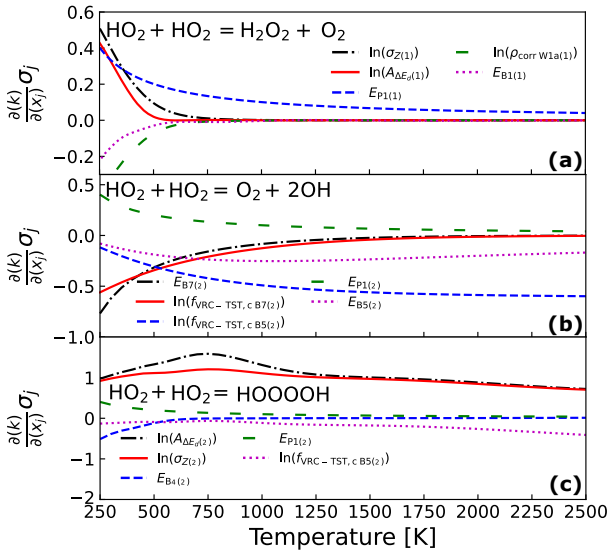


Figure 8: Uncertainty-weighted sensitivity analysis for the rate constants of R1, R2a, and R2b at 1 atm.

The second possibility explored is that the discrepancy can be attributed to remaining uncertainties in even state-of-the art theory. To this end, it is worth noting that the experimental determinations of k_1 at the highest temperatures [105, 106] lie outside the propagated uncertainties in the theoretical calculations [113] for k_1 and k_1 + $k_{2a\text{-}d}$ (i.e. the *a priori* model in Fig. 3)—such that the discrepancies do not appear to stem from (known) uncertainties in the theory. As shown in Fig. 8, while the uncertainties in the calculated k_1 due to uncertainties in the stationary point energies ($E_{\mathrm{P1(1)}}$, $E_{\mathrm{B1(1)}}$) and properties of the intermediate complexes ($\rho_{corrW1a(1)}$, $A_{\Delta E_d(1)}$, σ_Z) can be quite large at low temperatures, these uncertainties decrease with increasing temperatures as the effect of energies in the Boltzmann factor decreases and as the most precisely characterized transition state in the theoretical calculations [113] (B2) becomes the dominant reaction bottleneck. In addition to the parametric uncertainties considered in the MSI model and presented in Fig. 8, we also briefly explored the influence of potential rapid intersystem crossings between the singlet

and triplet surfaces via the weakly bound intermediate complexes—which would comprise a potential structural uncertainty [71, 70] in the theoretical model. However, limited explorations with a combined master equation for the combined singlet and triplet systems that assume infinitely fast intersystem crossing rates among W1a, W1b, W2a, and W2b indicate only minor differences in the calculated rate constants and their temperature dependence. Altogether, these results indicate the discrepancies among theoretical and experimental determinations of k_1 do not appear to be the result of (known) uncertainties in the theoretical calculations and, in fact, resolution of the discrepancies can be achieved using rate constants nearly identical to the theoretical kinetics calculations [113], cf. Fig. 3.

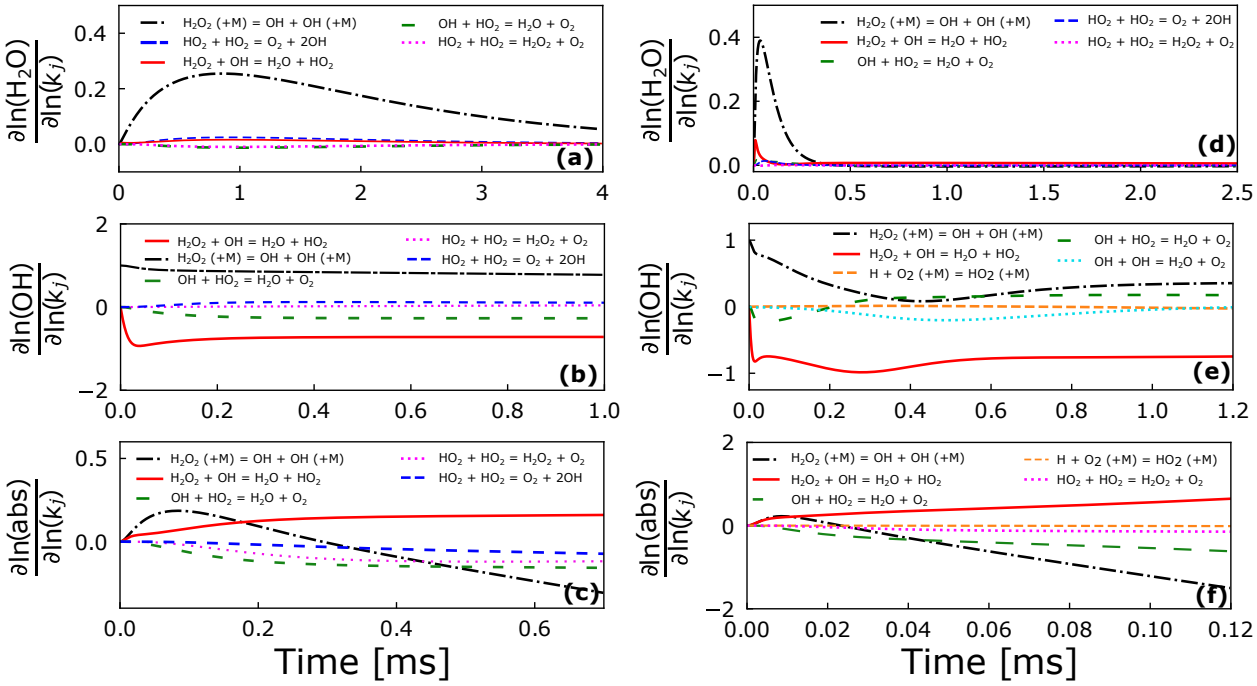


Figure 9: Conventional kinetic sensitivity analysis for predictions of the OH, H_2O, and absorbance time profiles shown in Fig. 5 using the MSI model.

The third remaining possibility to be explored is that even state-of-the-art multi-species measurements that completely describe the evolution of the key species do not preclude the possibility of multiple interpretations of H_2O_2 decomposition data, which have proven to be notoriously difficult to interpret historically. Conventional rate constant sensitivity analysis at the lowest and

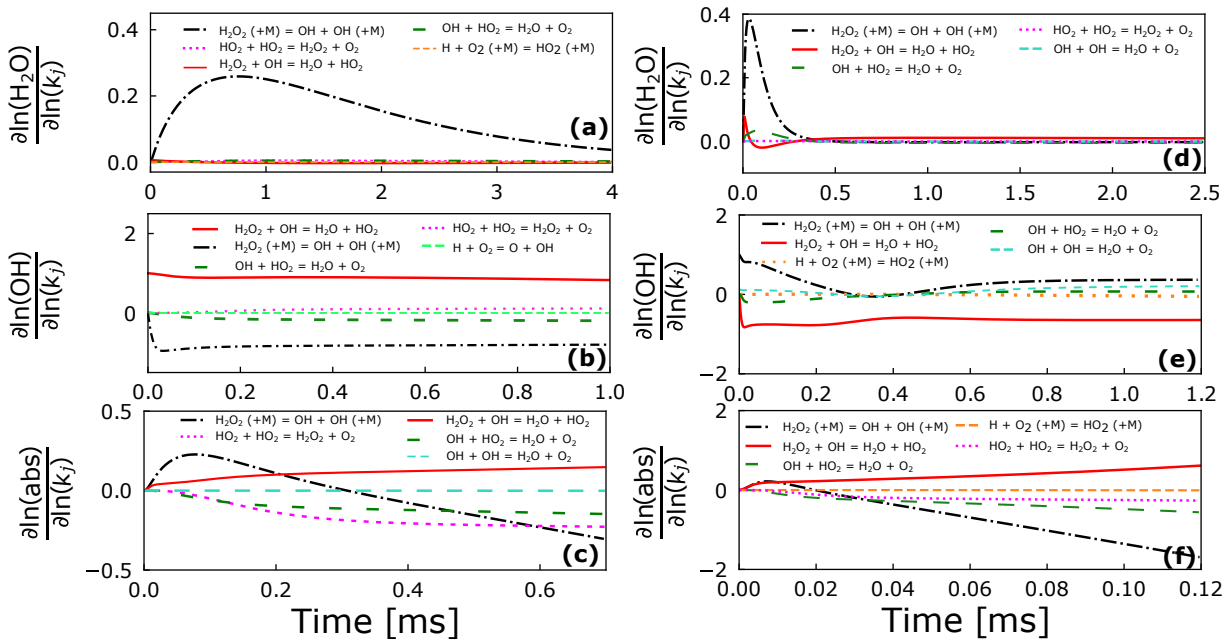


Figure 10: Conventional kinetic sensitivity analysis for predictions of the OH, H_2O, and absorbance time profiles shown in Fig. 5 using the model of Hong et al. [106].

highest temperatures of Hong et al. [106] for the MSI model (Fig. 9) and the model of Hong et al. [106] (Fig. 10) provide some initial glimpse into this idea. (Similarly, based on the fact that other experimental determinations [105] of k_1 also hinge on the absorbance time profiles, the same ideas also apply to those as well.) Overall, Figs. 9 and 10 indicate R1, R3, R4, R5, and R6 are the five reactions with the strongest influence on the experimental observables—supporting the notion that measurements of these observables during H_2O_2 decomposition provide information about their rate constants, as has been leveraged in many previous studies [128, 105, 103, 102, 106]. Of course, as recognized previously, the observables are generally influenced by rate constants for multiple reactions—such that the information about the rate constants of multiple reactions are often coupled [128, 105, 102, 106].

The sensitivity analysis results also reveal two further important subtleties in the interpretation of the experimental observables. First, the sensitivity coefficients obtained using the MSI model (Fig. 9) and the model of Hong et al. [106] (Fig. 10) are different (i.e. the sensitivity coefficients are nonlinear). Consequently, conclusions about the relative importance of each reaction

for each observable for one model may not hold for other values of rate constants in the model. For example, the relative importance of R1 in predictions using the MSI model is considerably less than in predictions using the model of Hong et al. [106]. Second, the sensitivity coefficients, and their relative rankings, are strongly temperature dependent. Consequently, conclusions about the relative importance of each reaction for each observable at one temperature may not hold for other temperatures. For example, the relative importance of R1 in predictions of both models is considerably smaller at higher temperatures, where the largest discrepancies between theoretical calculations and experimental determinations of k_1 are observed. This reduced sensitivity at higher temperatures serves to explain the fact that the MSI model and the model of Hong et al. [106] can both reproduce the measured time profiles equally well despite differences in their k_1 values that exceed experimental uncertainty estimates [106] based on analyses at lower temperatures (and with other choices of rate constants).

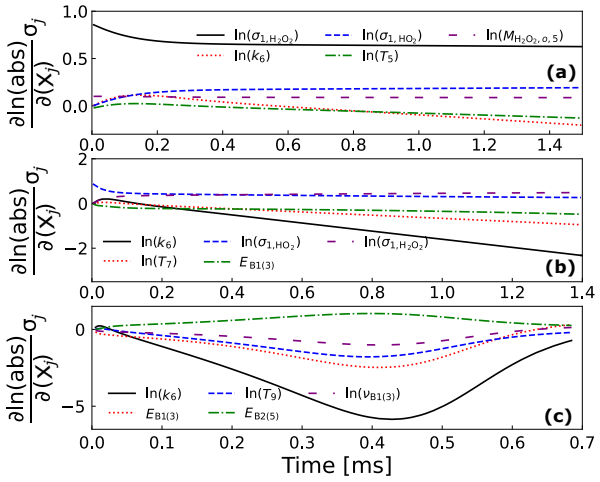


Figure 11: Uncertainty-weighted sensitivity analysis for absorbance time profiles of Kappel et al. [105] shown in Fig. 4.

In fact, once the constraints from *ab initio* theory are imposed, uncertainty-weighted sensitivity analysis (Figs. 11 and 12) indicates that the measured time profiles provide minimal additional information about k_1 beyond the theoretical calculations but, importantly, do provide very tight constraints on several other reactions within the H_2O_2 system. For context, while traditional sensitivity coefficients, $\frac{\partial f_i}{\partial x_j}$, can reveal the model parameters whose values have the strongest influence

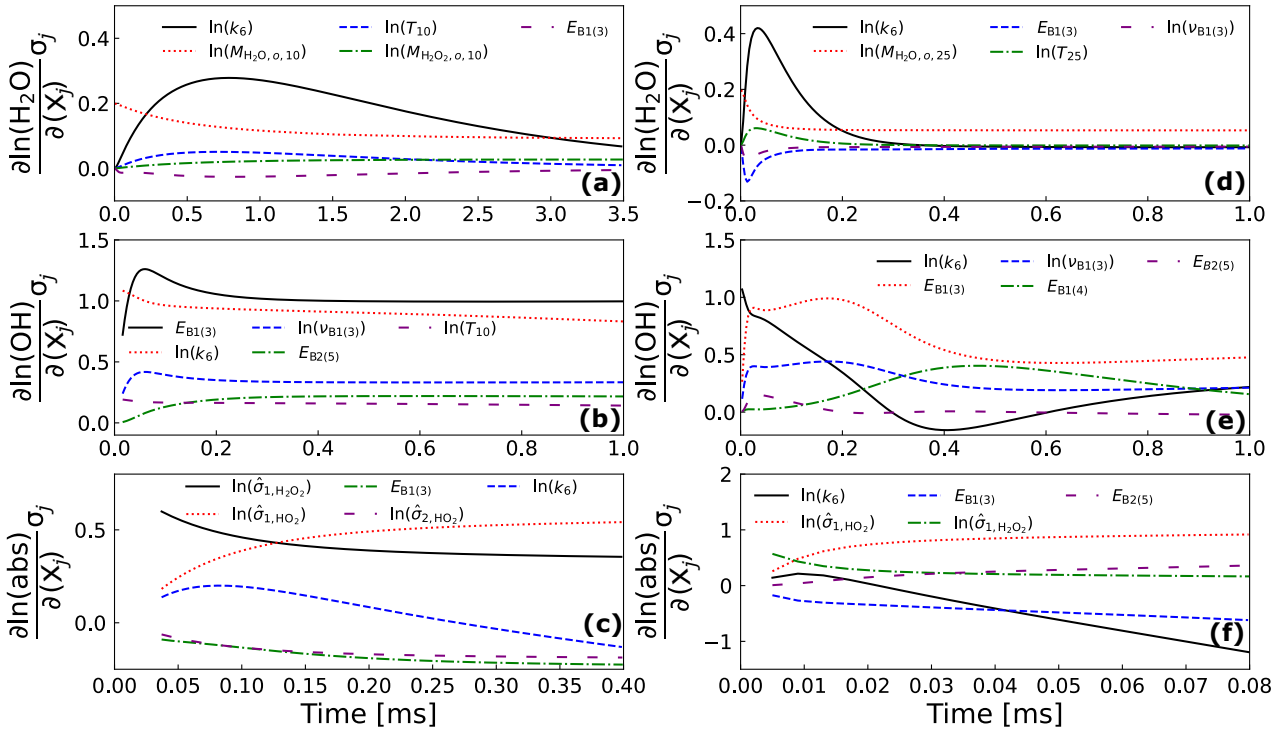


Figure 12: Uncertainty-weighted sensitivity analysis for OH, H_2O, and absorbance time profiles of Hong et al. [106] shown in Fig. 5.

on model predictions, uncertainty-weighted sensitivity coefficients, $\frac{\partial f_i}{\partial x_j}\sigma_j$, can reveal the model parameters whose prior uncertainties, σ_j, contribute the most to the uncertainty in the model predictions. Likewise, the model parameters with the highest uncertainty-weighted sensitivity coefficients are among those most informed by data for that particular target, y_i^{t}.

In particular, uncertainty-weighted sensitivity coefficients of the time profiles for absorbance 215 nm and H_2O, OH, and absorbance at 227 nm are shown in Fig. 11 and 12 respectively. Inspection of the most influential parameters in Figs. 11 and 12 reveals that theory-constrained predictions of the macroscopic observables are predominately influenced by uncertainties in parameters related to k_6, k_3 ($\mathrm{E}_{B1(3)}$, ν_3), k_4 ($\mathrm{E}_{B1(4)}$), k_5 ($\mathrm{E}_{B1(5)}$), and physical model parameters, including the initial mole fractions ($M_{\mathrm{H_2O_2}}$ and $M_{\mathrm{H_2O}}$), initial temperature (T), and absorbance cross sections (e.g. σ_{1,H_2O_2}, σ_{1,HO_2}). Therefore, the measurements of these time profiles provide important constraints on those parameters describing the rates of those other reactions (and the physical model).

In this regard, the MSI model yields better agreement with the experimental data than the *a priori* model via adjustments to parameters describing R3-R6. Similarly, the fact that the MSI model and the model of Hong et al. [106] can both achieve agreement with experimental data with different k_1 values can be explained in large part by the fact that the MSI model contains different rate constants for other reactions. Notably, k_6, which is the most influential rate constant in Figs. 9 and 12, in the MSI model is ∼10-30% lower than that of Hong et al. (see Fig. S42) and k_3, which is generally the second most influential rate constant in Figs. 9 and 12, in the MSI model is ∼10-25% lower than that of Hong et al. (see Fig. S42). Of note, while both the MSI model and the model of Hong et al. [106] are consistent with the measured time profiles of Hong et al. [106], the MSI model is also consistent with all other data for the H_2O_2 decomposition and its constituent reactions, including *ab initio* theoretical calculations (Harding and Klippenstein, Unpublished) [113] and other experimental data for the reactions in Table 2 across broad ranges of thermodynamic conditions (cf. Supporting Information).

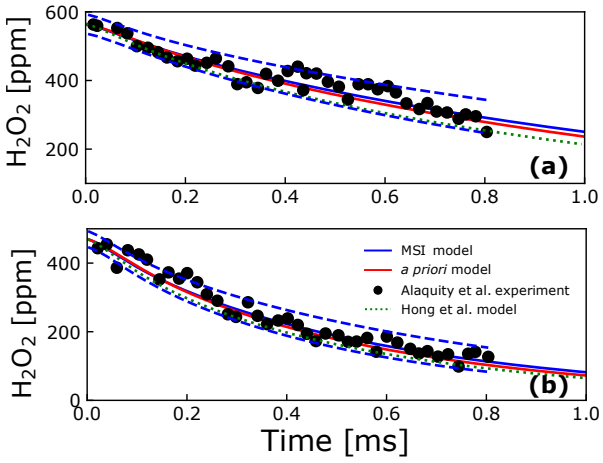


Figure 13: Time profiles of H_2O_2 in shock-heated H_2O_2/Ar mixtures near (a) 1074 K and 1.6 atm and (b) 1127 K and 1.5 atm. Symbols represent experimental data from Alquaity et al. [144]; lines represent model predictions using the *a priori* model, MSI model, and model of Hong et al. [106].

As an independent test of the MSI model, predictions using the MSI model were compared to the recent time profiles of H_2O_2 from Alquaity et al. (Fig. 13). While these measurements were not included as optimization targets, simulations using the MSI model reproduce the H_2O_2 time

profiles within uncertainties. Therefore, this performance is indicative of the predictive capability of the overall MSI model and provides confidence that the parameter adjustments within the MSI model are physically realistic.

5.3.2 Low temperatures

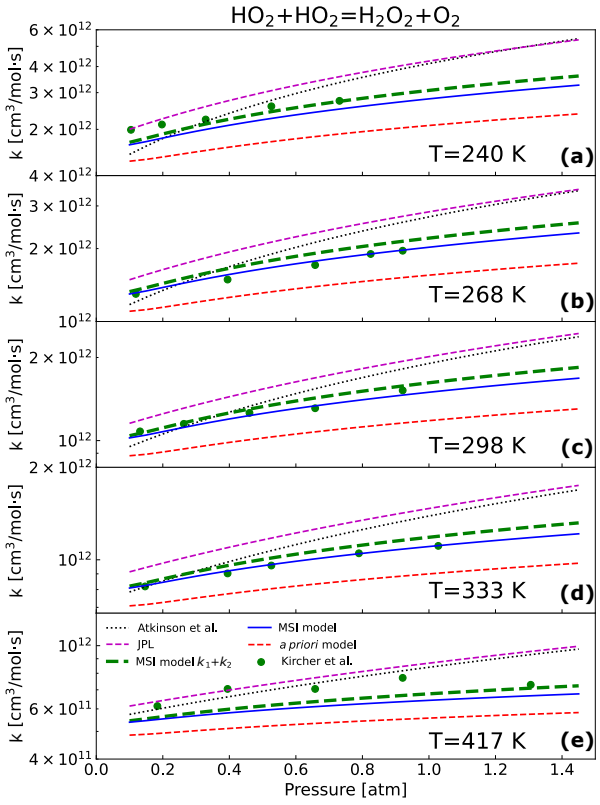


Figure 14: Low-temperature rate constants for the HO_2 + HO_2 reaction across varied temperature and pressure in Ar. Symbols represent the experimental determinations of Kircher et al. [107]. Lines represent the values in the *a priori* model (i.e. the nominal theoretical calculations [113]), MSI model, rate evaluation of Atkinson et al. [136], and rate evaluation of JPL [138].

This MSI model is also consistent with and provides improvements over the *a priori* model for the low-temperature data of Kircher et al. [107]. In particular, predictions using the *a priori* model and MSI model are compared to the k_1 determinations of Kircher et al. [107] in Fig. 14 and the re-generated time profiles based on their k_1 determinations (as discussed above in the Implementation section) in Fig. 15. Overall, the *a priori* model yields lower rate constants than the experimentally determined k_1 and, correspondingly, predicts slower HO_2 decay than the regenerated experimental

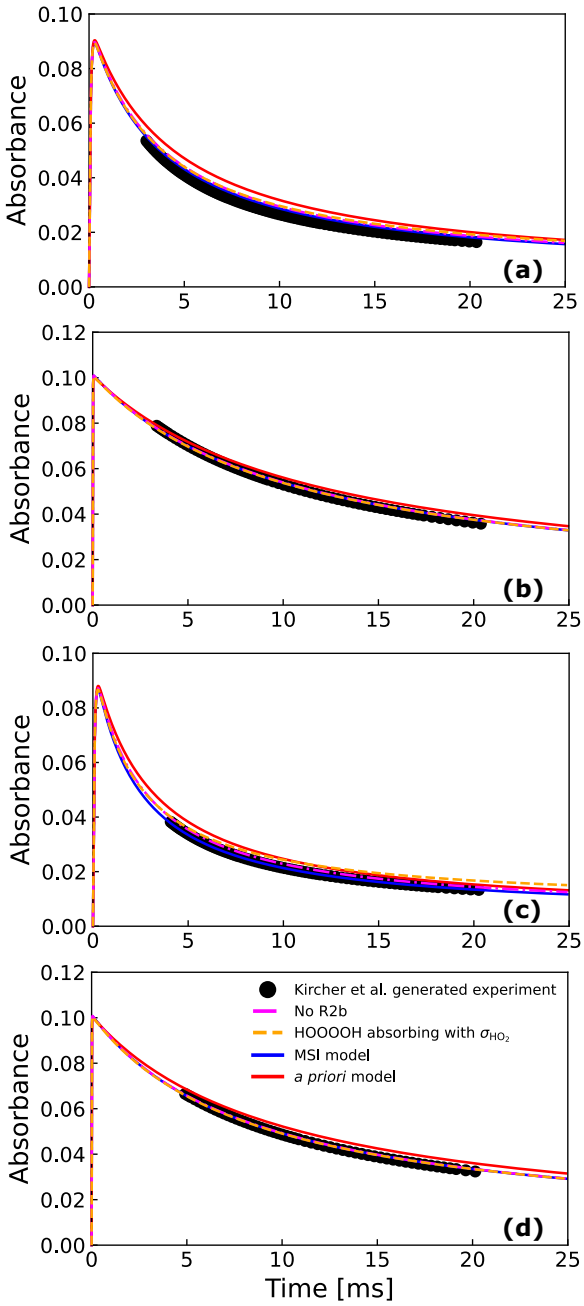


Figure 15: Time profiles of absorbance at 227 nm in flash-photolyzed $Cl/Cl_2/CH_3OH/O_2/Ar$ mixtures near (a) 240 K and 0.10 atm, (b) 417 K and 0.18 atm, (c) 240 K and 0.73 atm, and (d) 417 K and 1.3 atm. Symbols represent (re)generated raw experimental data from Kircher et al. [107]; lines represent model predictions using the *a priori* model, the MSI model, and a model that incorporates the reported experimentally derived value of k_1 at each condition [107] and secondary reactions of the *a priori* model.

time profiles (despite crude empirical adjustments in the properties of weakly bound complexes in the original calculations [113] to improve agreement with low-temperature experiments.) On the whole, the MSI model yields rate constants that are much closer to the experimentally determined k_1 and regenerated experimental time profiles. In contrast to experimental measurements at higher temperatures, sensitivity analysis of these time profiles at low temperatures revealed minimal influence from reactions other than HO_2 + HO_2.

However, as shown by model predictions employing alternative treatments of HOOOOH from R2b in Fig. 15 and differences between k_1 and k_1 + $k_{2a\text{-}d}$ in Fig. 14, HOOOOH stabilization appears to be of some modest relevance to the experimental interpretations. Noting that the differences between k_1 and k_1 + $k_{2a\text{-}d}$ are essentially exclusively attributable to R2b, the differences indicate that R2b is responsible for up to $\sim$8% of the total HO_2 + HO_2 reaction rate constant at atmospheric pressure (while being negligible at the lowest experimental pressures). Correspondingly, simulations where R2 is excluded from the MSI model yield a slower decay of the absorbance. However, these differences are only noticeable at higher pressures and lower temperatures, based on the way HOOOOH evolves on experimental timescales. At higher pressures and lower temperatures, HOOOOH decomposition is sufficiently slow (e.g. $\sim$3 s^{-1} at 240 K and 0.73 atm) that HOOOOH is stable and R2b is effectively irreversible on experimental timescales (tens of milliseconds). By contrast, at higher pressures and higher temperatures, HOOOOH decomposes rapidly (e.g. $\sim$4 x 10^5 s^{-1} at 417 K and 1.3 atm) relative to experimental timescales. Given that HOOOOH decomposes primarily back to HO_2 + HO_2, R2b rapidly reaches partial equilibrium and therefore does not contribute significantly to the decay rate of the absorbance signal at higher temperatures. Consequently, the experimental rate constant determinations of Kircher et al. [107] more closely correspond to the total rate constant k_1 + $k_{2a\text{-}d}$ at lower temperatures and more closely correspond to the triplet channel rate constant k_1 at higher temperatures.

Of course, in addition to affecting the absorbance time profiles by contributing to HO_2 disappearance, HOOOOH formed via R2b may also affect the absorbance signal by absorbing at 227 nm along with HO_2 and HOOH, but its absorption cross section is not known. Simulations that

assume an HOOOOH absorption cross section equal to that of HOOH yield negligible differences from those using the MSI model (which ignores HOOOOH absorption). However, simulations that assume an HOOOOH absorption cross section equal to that of HO_2 (which is ten times higher than that for HOOH) yield noticeably slower decay of the absorbance signal, as shown in Fig. 15. If this higher absorption cross section were assumed in the MSI model, the optimized k_1 values would be up to 12% higher than those of the present analysis—suggesting that uncertainties in the HOOOOH cross section may also contribute somewhat to uncertainties in the experimental interpretations. Given lack of actual raw data from the experiments and lack of information about the HOOOOH absorption cross section, it is not possible to make any definitive conclusions, though it is possible that complications related to HOOOOH formation (discussed in this paragraph and the preceding paragraph) may contribute to the scatter of 25% among k_1 determinations from various experimental studies.

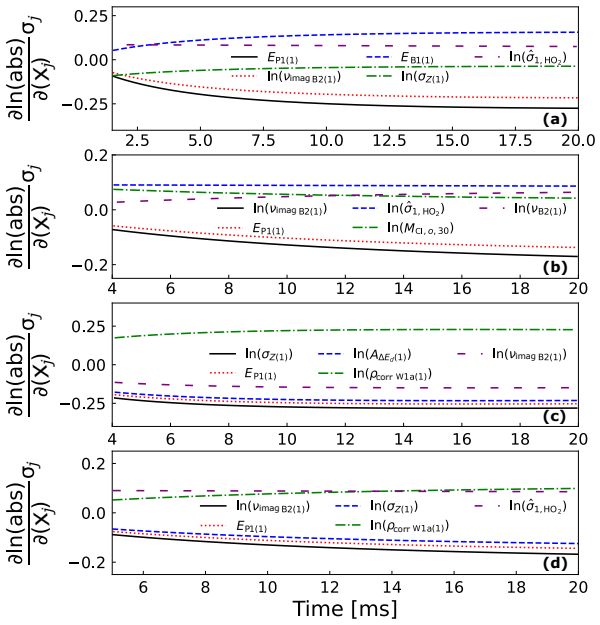


Figure 16: Uncertainty-weighted sensitivity analysis for absorbance time profiles of Kircher et al. [107]. shown in Fig. 15.

Overall, the MSI model predictions agree with the experimental determinations of k_1 within 1-17%, which is considerably lower than the differences observed for recommended rate constant expressions from atmospheric chemistry evaluations (and the *a priori* model). Likewise, in addi-

tion to yielding predictions in better agreement with the experimental data than the *a priori* model in Figs. 14 and 15, the MSI model has lower uncertainties than the *a priori* model—indicating that the low-temperature experimental data [107] effectively constrains model parameters describing R1 (and to a lesser extent R2). Uncertainty-weighted sensitivity analysis for the absorbance time profiles shown in Fig. 16 illuminates the constraints imposed on model parameters. Namely, prediction uncertainties are predominately influenced by uncertainties in parameters related to k_1 ($E_{P1(1)}$, $E_{B1(1)}$, $\nu_{imagB2(1)}$, $\rho_{corrW1a(1)}$, $\sigma_{z(1)}$, $A_{\Delta E_d(1)}$) and physical model parameters (e.g. $M_{Cl,o,e}$, $\hat{\sigma}_{1,HO_2}$, and $\hat{\sigma}_{1,H_2O_2}$), with uncertainties in the properties of the weakly bound complexes ($\rho_{corrW1a(1)}$) and collisional energy transfer ($\sigma_{z(1)}$, $A_{\Delta E_d(1)}$) notably playing a larger role at higher pressures where (partial) stabilization of the weakly bound complexes is important. In turn, slight adjustments to these parameters (e.g. a few hundredths of a kcal/mol in stationary point energies, a few per cent in imaginary frequencies, and $\sim$20% in state density correction factors to account for anharmonicity in the weakly bound complexes and $\sim$20% in estimated collisional energy transfer function parameters) are responsible for the improved agreement between the MSI predictions and the experimental data at low temperatures.

5.3.3 Role of Theory in the Present Analysis

More broadly, a key theme that emerges from the present comparison of state-of-the-art theoretical calculations and experimental measurements is that evaluating consistency, and removing apparent inconsistencies, requires consideration of the theoretical and experimental data *simultaneously*. To illustrate this, Fig. 17 compares the rate constants from the *ab initio* calculations and the present MSI model with those from data analysis procedures which do not include theory. In the two rate parameter optimization studies [60, 145] published before the theoretical calculations of Klippenstein et al. [113] (which found more channels of importance), R1 was the only HO_2 + HO_2 reaction channel considered and only rate constant determinations and macroscopic observables were considered among the targets. The derived rate constants from both studies exceed the estimated error bounds in the theoretical calculations for both the main channel (R1) and total

reaction (R1 + R2a-d) and one of them [60] exhibits a very strong temperature dependence that is also inconsistent with the theory.

For the purposes of illustration, Fig. 17 also shows results from a rate parameter optimization against the present set of rate constant determination and macroscopic observable targets (target classes II-IV only) that notably also includes the main singlet channel (R1b). This optimization was performed considering Arrhenius parameters for R1, R2a, R3, R4, R5a, and R6 where each were treated with single Arrhenius expressions (given that limited testing of expressions with more degrees of freedom unsurprisingly had convergences issues). Clearly, the results indicate that the rate constant for R2a is very poorly constrained by the data—with the optimized rate constant being drastically different from the theory. Similarly, even the resulting rate constants for the main reaction (R1) and total reaction (R1 + R2a-d) also show a stronger temperature dependence than the theoretical calculations, yet this optimization shows no noticeable improvements over the present MSI model in reproducing the raw experimental data.

That is to say, when theoretical data are not included in the optimization, the optimized solution is essentially not unique—with many possible interpretations yielding similar performance against the experimental data. However, simultaneous consideration of theoretical and experimental data within multiscale, physics-based frameworks for data analysis, as in the present MSI model, enables interpretations of the experimental data in a manner constrained by theory—thereby removing interpretation ambiguities based on the experimental data alone and ensuring that the experimental interpretations (and their extrapolations) are physically realistic.

5.4 Concluding Remarks: HO_2 + HO_2

Previous theoretical and experimental data for the reactions in the H_2O_2 decomposition mechanism (R1-R6) were analyzed using the MultiScale Informatics approach, which optimizes values and quantifies uncertainties for a set of molecular parameters (within theoretical kinetics calculations), rate parameters, and physical model parameters (within simulations of macroscopic observables) based on data from various sources and scales. Special attention was paid to HO_2 +

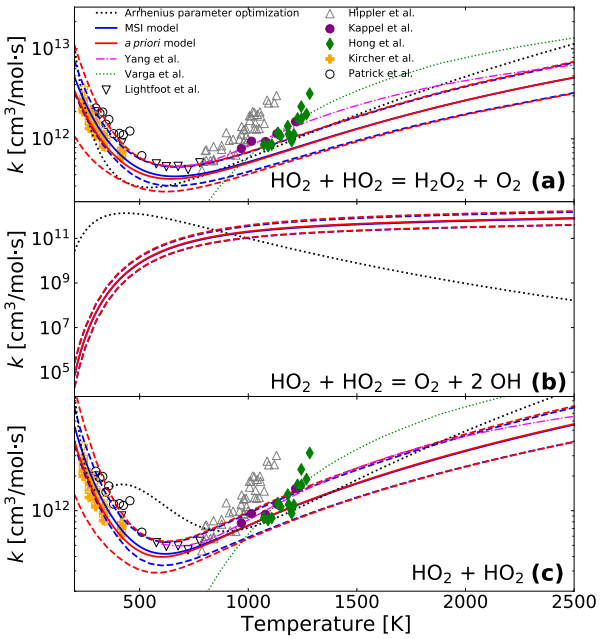


Figure 17: Rate constants at 1 atm for (a) $HO_2 + HO_2 = H_2O_2 + O_2$ (R1), $HO_2 + HO_2 = O_2 + 2OH$ (R2a), (c) $HO_2 + HO_2$ = products (R1 + R2a-d). Symbols represent experimental determinations [118, 117, 106, 105, 107, 116] as indicated in the legend. Solid symbols specifically designate the experimental determinations of Hong et al. [106], Kappel et al. [105] and Kircher et al. [107] that correspond to raw experimental data in Figs. 15, 4 and 5 (where the MSI model reproduces the raw data). Solid and dashed lines represent the values and uncertainties, respectively, for the *a priori* and MSI models. Additional lines represent data analysis procedures which do not include theory [60, 145]. (Note: given that R2 was not considered in the original experimental interpretations, the total rate constants $k_1 + k_{2a\text{-}d}$ for the experimental determinations simply correspond to the reported k_1 values).

HO_2 (R1 and R2). Overall, the MSI model constructed was found to be consistent with all *ab initio* calculations and experimental data used as targets (as well as other experiments not used as targets).

In particular, this MSI model reproduces the measured time profiles of Kappel et al. [105] and Hong et al. [106] equally as well as the original interpretations [105, 106], which suggested a sharp rise in the rate constant with increasing temperature. Interestingly, the high-temperature rate constants (and uncertainties) for $HO_2 + HO_2$ in this same MSI model are nearly identical to those from the theoretical calculations of Klippenstein et al. [113], which predicted a less pronounced temperature dependence and commensurately lower rate constants at high temperatures. While both the original and present interpretations of the data are consistent with the raw experimental

data, only the present MSI interpretation is consistent with theoretical calculations for R1-R2 as well as other theoretical and experimental data for R3-R6 and is, therefore, the most reliable.

Further analysis revealed that the difference between the original and present interpretations stems from differences in the rate constants of other reactions (R3-R6) in the H_2O_2 decomposition system. The fact that the MSI rate constants for R3-R6 are consistent with other theoretical and experimental datasets used as targets as well as experimental data not used as targets supports the notion that the MSI values are physically realistic.

Additional analysis using the MSI model was conducted to assess the influence of HOOOOH—a reaction intermediate of HO_2 + HO_2 identified by recent theoretical calculations but not considered in earlier analysis—on typical signals from low-temperature experiments. Formation of HOOOOH via R2b is found to comprise $\sim$8% of the total HO_2 + HO_2 rate constant at low temperatures at 1 atm. While HOOOOH is stable on experimental timescales at the lower experimental temperatures ($\sim$250 K), HOOOOH rapidly decomposes back to HO_2 + HO_2 at higher experimental temperatures ($\sim$400 K). Consequently, experimental determinations of the HO_2 + HO_2 rate constant would most closely correspond to k_1 + $k_{2a\text{-}d}$ at low temperatures and to k_1 at high temperatures. The HOOOOH absorption cross section is not known, but if HOOOOH were to absorb as strongly as HO_2, HOOOOH absorption would alter experimentally derived k_1 values by $\sim$12%. Altogether, while the lack of actual raw data and lack of information about HOOOOH absorption preclude definitive conclusions, the present analysis suggests effects of HOOOOH on experimental interpretations that are smaller than, but may contribute to, the $\sim$25% scatter among data from different experimental studies.

A general theme of the present work (consistent with similar analyses [69, 70, 71, 72] for other reactions) is that apparent inconsistencies among theoretically calculated and experimentally derived rate constants can often be resolved when the constraints imposed by *ab initio* electronic structure calculations and measured macroscopic observables are considered simultaneously within multiscale, physics-based frameworks for data analysis. We therefore recommend that future experimental studies report raw data (e.g. measured time profiles) and future theoretical stud-

ies provide theoretical kinetics input files (e.g. for Variflex or MESS) in addition to derived rate constants—in order to enable rigorous comparisons of theory and experiment in a manner that better reflects the information content and associated uncertainties in each.

Chapter 6: Continued Movement Towards a High-Accuracy Kinetic Database Informed by Theoretical and Experimental Data: Coupled Reaction Systems

6.1 Scientific Motivation: Coupled Reaction Systems

In the previous two studies from LaGrotta et al. [72, 146], the authors specifically focused on the smaller reaction subsystems and demonstrated how the MultiScale Informatics approach can be used to move towards a high accuracy kinetics database. However, LaGrotta et al. [72] also exemplified the difficulty in achieving data redundancy in kinetics – such that a large web of interconnected reaction systems may be required. This study begins to build a network of redundant data and will demonstrate how the MSI approach can be used to find optimized kinetic parameters and quantified uncertainties by considering multiple coupled reaction systems. Through a series of selected vignettes presented here, the key features of the MSI approach will be highlighted and results that demonstrate the advantages of implementing the MSI method will be discussed.

6.2 Implementation: Coupled Reaction Systems

The MSI approach was implemented here for a series of coupled reaction systems, which are designated into groups A-Q for the ease of discussion and organization. The list of reactions and their group affiliation are listed in Table 2. The overall kinetics model used for data analysis included the reactions listed in Table 2 along with other “secondary" reactions that are (or may be) involved in the interpretations of the raw experimental data. The nominal rate constants and uncertainty factors for these secondary reactions were taken from the Foundational Fuel Chemistry Model (FFCM) [84] for all other H_2/C_1-C_3 reactions, the USC [209] for all other H_2/C_4-C_6 re-

Table 1: List of kinetic parameters considered in the optimization.[a]

Reaction		Kinetic Parameters
A-R1	$H + O_2 = OH + O$	$A'_{(A1)}$, $n_{(A1)}$, $E_{(A1)}$
A-R2	$OH + H_2 = H_2O + H$	$A'_{(A2)}$, $n_{(A2)}$, $E_{(A2)}$
A-R3	$O + H_2 = OH + H$	$A'_{(A3)}$, $n_{(A3)}$, $E_{(A3)}$
B-R1	$HO_2 + HO_2 = H_2O_2 + O_2$	$\sigma'_{Z(B1)}$, $n_{\Delta E_d(B1)}$, $A'_{\Delta E_d(B1)}$, $\nu'_{mr(B1)}$, $\eta'_{mr(B1)}$ $\nu'_{B1(B1)}$, $E_{B1(B1)}$, $E_{P1(B1)}$, $E_{W1a(B1)}$, $\rho'_{corr\ W1b(B1)}$, $\rho'_{corr\ W1a(B1)}$, $\nu'_{W1a(B1)}$, $f'_{VRC\text{-}TST,c\ B1(B1)}$ $E_{W1b(B1)}$, $\nu'_{imag\ B2(B1)}$, $\nu'_{P1(B1)}$, $f'_{VRC\text{-}TST,c\ P1(B1)}$, $\nu'_{P2(B1)}$, $\nu'_{B8(B1)}$, $\rho'_{corr\ B8(B1)}$ $E_{P2(B1)}$, $\nu'_{W1b(B1)}$, $E_{B8(B1)}$, $E_{B2(B1)}$, $E_{P1(B1)}$
B-R2a	$HO_2 + HO_2 = O_2 + 2OH$	$\sigma'_{Z(B2)}$, $n_{\Delta E_d(B2)}$, $A'_{\Delta E_d(B2)}$
B-R2b	$HO_2 + HO_2 = HOOOOH$	$E_{B4(B2)}$, $\nu'_{W2b(B2)}$, $\nu'_{B7_{1(B2)}}$, $\nu'_{B9(B2)}$, $f'_{VRC\text{-}TST,c\ B5(B2)}$, $\nu'_{B4(B2)}$
B-R2c	$HO_2 + HO_2 = H_2O + O_3$	$f'_{VRC\text{-}TST,c\ B7(B2)}$, $\nu'_{W2a(B2)}$, $E_{W2b(B2)}$, $\nu'_{imag\ B3(B2)}$, $f'_{VRC\text{-}TST,c\ P1(B2)}$
B-R2d	$HO_2 + HO_2 = H_2O_2 + O_2(X)$	$E_{B3(B2)}$, $E_{B7(B2)}$, $E_{P4(B2)}$, $\nu'_{P3(B2)}$, $\rho'_{corr\ B9(B2)}$
B-R2e	$H_2O_2 + O_2(X) = HOOOOH$	$\nu'_{B3(B2)}$, $E_{B6(B2)}$, $\nu'_{B6(B2)}$, $\nu'_{P5(B2)}$, $E_{W3(B2)}$,
B-R2f	$H_2O_2 + O_2(X) = H_2O + O_3$	$E_{P3(B2)}$, $\nu'_{P1(B2)}$, $\rho'_{corr\ W2a(B2)}$, $\rho'_{corr\ P5(B2)}$, $\nu'_{P4(B2)}$, $\nu'_{W3(B2)}$
B-R2g	$H_2O_2 + O_2(X) = O_2 + 2OH$	$E_{P5(B2)}$, $E_{B5(B2)}$, $\rho'_{corr\ W2b(B2)}$, $\nu'_{imag\ B6(B2)}$, $E_{W2a(B2)}$,
B-R2h	$H_2O + O_3 = HOOOOH$	$E_{P1(B2)}$, $\eta'_{B7_1(B2)}$, $\eta'_{B7_2(B2)}$, $\eta'_{W3_1(B2)}$, $\eta'_{W3_2(B2)}$, $\eta'_{W3_3(B2)}$, $\eta'_{P5(B2)}$
B-R2i	$H_2O + O_3 = O_2 + 2OH$	$E_{B7_1(B2)}$, $f'_{VRC\text{-}TST,c\ B7_2(B2)}$, $\nu'_{B7_{2(B2)}}$
B-R2j	$O_2 + 2OH = HOOOOH$	$f'_{VRC\text{-}TST,c\ B4(B2)}$, $E_{B9(B2)}$, $E_{P1(B2)}$, $E_{W2a(B2)}$, $\nu'_{B5(B2)}$
B-R3	$H_2O_2 + OH = H_2O + HO_2$	$E_{B1(B3)}$, $\nu'_{B1(B3)}$, $\eta'_{B1(B3)}$, $\nu'_{imag\ B1(B3)}$
B-R4	$OH + OH = H_2O + O$	$E_{B1(B4)}$, $E_{B2(B4)}$, $\nu'_{B1,\ B2(B4)}$, $\nu'_{imag\ B1,\ B2(B4)}$
B-R5a	$OH + HO_2 = H_2O + O_2$	$f'_{VRC\text{-}TST,c\ B1(5)}$, $E_{B2(B5)}$, $\nu'_{B2(5)}$, $\eta'_{B2(B5)}$, $\nu'_{imag\ B2(B5)}$, $E_{B3(B5)}$
B-R5b	$OH + HO_2 = H_2O + O_2(X)$	$\nu'_{B3(B5)}$, $\eta'_{B3(B5)}$, $\nu'_{imag\ B3(B5)}$, $\nu'_{W1(B5)}$
B-R6	H_2O_2 (+M) = OH + OH (+M)	$A'_{(B6)}$, $n_{(B6)}$, $Ea_{(B6)}$
C-R1	$H + HO_2 = H_2 + O_2$	$\nu'_{B1(C1)}$, $\nu'_{imag\ B1(C1)}$, $E_{B1(C1)}$
C-R2	$H + HO_2 = H_2O + O_2(X)$	$A'_{(C2)}$
C-R3a	$H + HO_2 = OH + OH$	$\nu'_{W2(C3)}$, $E_{W2(C3)}$, $\nu'_{W1(C3)}$, $E_{B5(C3)}$,
C-R3b	$H + HO_2 = H_2O + O$	$E_{B1(C3)}$, $E_{W1(C3)}$, $E_{B6(C3)}$, $\nu'_{B5(C3)}$, $\eta'_{B1(C3)}$, $\nu'_{B6(C3)}$, $\nu'_{W3(C3)}$, $\nu'_{W3(C3)}$,
C-R3c	$H + HO_2 = H_2O + O(X)$	$E_{B2(C3)}$, $\eta'_{B2(C3)}$, $E_{B3(C3)}$, $\eta'_{B3(C3)}$, $E_{B4(C3)}$, $\eta'_{B4(C3)}$, $\eta'_{B7(C3)}$, $E_{B7(C3)}$
D-R1	$H + O_2(+M) = HO_2(+M)$	$A'_{(D1)}$, $n_{(D1)}$, $E_{(D1)}$
E-R1a	$NO_2 + H = NO + OH$	$\nu'_{B1(E1)}$, $E_{B1(E1)}$, $E_{B4(E1)}$, $\nu'_{W1(E1)}$, $\nu'_{B3(E1)}$ $\nu'_{P1(E1)}$, $\nu'_{B2(E1)}$, $\sigma'_{Z(E1)}$,
E-R1b	NO + OH = HONO	$E_{W3(E1)}$, $\nu'_{P2(E1)}$, $\nu'_{B4(E1)}$, $n_{\Delta E_d(E1)}$, $A'_{\Delta E_d(E1)}$, $E_{P1(E1)}$, $f'_{VRC\text{-}TST,c\ B4(E1)}$, $f'_{VRC\text{-}TST,c\ B3(E1)}$, $E_{B2(E1)}$, $E_{B3(E1)}$, $E_{W1(E1)}$, $E_{P2(E1)}$, $\nu'_{imag\ B1(E1)}$, $f'_{VRC\text{-}TST,c\ B2(B2)}$
F-R1	$CH_3 + HO_2 = CH_4 + O_2$	$E_{B1(F1)}$, $\nu'_{low\ B1(F1)}$, $(Q_{anh,c,CH_3(F1)})$
F-R2	$CH_3 + HO_2 = CH_3O + OH$	$(Q_{anh,c,CH_3(F2)})$, $f'_{VRC-TST(F2)}$
F-R3	$OH + CH_4 = CH_3 + H_2O$	$A'_{(F3)}$, $n_{(F3)}$, $Ea_{(F3)}$

[a]Note that ' indicates ln() of that quantity.

actions, and a H_2 oxidation model from Konnov [139] for all other electronically excited species kinetics.

Physical model parameters from experiments included as targets were also among the active model parameters. These parameters included initial temperature, T_e, initial pressures, P_e

Table 1: Continued list of kinetic parameters considered in the optimization. [a]

Reaction		Kinetic Parameters
G-R1a	$C_2H_6 = CH_3 + CH_3$	$\nu'_{B1(G1)}$, $E_{B1(G1)}$, $\nu'_{W1(G1)}$, $\nu'_{B3(G1)}$, $\nu'_{B2(G1)}$, $f'_{VRC\text{-}TST,c\ B1(G1)}$,
G-R1b	$C_2H_6 = C_2H_5 + H$	$\sigma'_{Z(G1)}$, $n_{\Delta E_d(G1)}$, $A'_{\Delta E_d(G1)}$,
G-R1c	$C_2H_6 = C_2H_4 + H_2$	$\nu'_{P1(G1)}$, $E_{P3(G1)}$, $\eta'_{W1(G1)}$, $\nu'_{P2(G1)}$, $\eta'_{W1(G1)}$, $E_{W1(G1)}$
G-R1d	$CH_3 + CH_3 = C_2H_4 + H_2$	$\nu'_{imag\ B3(G1)}$, $\eta'_{B2(G1)}$, $E_{B2(G1)}$, $E_{P1(G1)}$
G-R1e	$CH_3 + CH_3 = C_2H_6$	$E_{B3(G1)}$, $\nu'_{P3(G1)}$, $E_{P2(G1)}$
G-R2	$C_2H_5 + H = C_2H_4 + H_2$	$\nu'_{B1(G2)}$, $E_{B1(G2)}$, $\nu'_{P1(G2)}$
H-R1	$CH_3COCH_3 + OH = CH_2CO + CH_3 + H_2O$	$A'_{(H1)}$
I-R1a	$CH_3 + OH = CH_3O + H$	$E_{B1(I1)}$, $E_{B4(I1)}$, $\nu'_{W1(I1)}$, $\sigma'_{Z(I1)}$, $n_{\Delta E_d(I1)}$, $A'_{\Delta E_d(I1)}$, $E_{P7(I1)}$, $E_{R(I1)}$,
I-R1b	$CH_3 + OH = CH_3OH$	$\nu'_{B8(I1)}$, $\nu'_{B4(I1)}$, $E_{B3(I1)}$, $E_{B7(I1)}$, $E_{B2(I1)}$, $\nu'_{P3(I1)}$, $\nu'_{B3(I1)}$, $E_{P1(I1)}$, $E_{P2(I1)}$,
I-R1c	$CH_3OH = CH_2 + H_2O$	$E_{B6(I1)}$, $\nu'_{P7(I1)}$, $\nu'_{P5(I1)}$, $\nu'_{B5(I1)}$, $E_{P6(I1)}$, $\nu'_{P7(I1)}$, $\nu'_{P1(I1)}$, $\nu'_{R(I1)}$,
I-R1d	$CH_3OH = CH_3 + OH$	$\nu'_{P4(I1)}$, $E_{B5(I1)}$, $E_{B8(I1)}$, $\nu'_{P6(I1)}$, ρ_{CF}, $E_{P5(I1)}$, $f'_{VRC\text{-}TST,c\ B7(I1)}$
I-R1e	$CH_4 + OX = CH_2OH + H$	$\nu'_{P2(I1)}$, $E_{P3(I1)}$, $E_{P4(I1)}$, $f'_{VRC\text{-}TST,c\ B2(I1)}$, $f'_{VRC\text{-}TST,c\ B1(I1)}$, $f'_{VRC\text{-}TST,c\ B6(I1)}$
I-R2	$CH_4 + OX = CH_3 + OH$	$\nu'_{B1(I2)}$, $E_{B1(I2)}$, $\nu'_{P7(I2)}$, $\nu'_{imag\ B1(I2)}$, $\nu'_{R(I2)}$
J-R1	$CH_2 + OH = CH_2O + H$	$A'_{(J1)}$
J-R2	$CH_2O + H\ (+M) = CH_3O\ (+M)$	$A'_{(J2)}$
J-R3	$CH_2O + H\ (+M) = CH_2OH\ (+M)$	$A'_{(J3)}$
K-R1a	$CH_2O = CO + H_2$	$\nu'_{W2(K1)}$, $\nu'_{B1(K1)}$, $E_{B4(K1)}$, $E_{B1(K1)}$, $\nu'_{B2(K1)}$, $\nu'_{imag\ B15(K1)}$, $E_{B6t(K1)}$, $E_{W5c(K1)}$,
K-R1b	$H + CO = CO + H_2$	$\nu'_{B13(K1)}$, $E_{B10(K1)}$, $E_{B11(K1)}$, $\nu'_{B11(K1)}$, $E_{B2(K1)}$, $E_{B15(K1)}$, $\nu'_{B15(K1)}$, $E_{W1(K1)}$, $E_{W0(K1)}$,
K-R1c	$CH_2O = H + HCO$	$\nu'_{B14(K1)}$, $\nu'_{W0(K1)}$, $E_{W4(K1)}$, $\nu'_{B4(K1)}$, $E_{W5t(K1)}$, $\nu'_{P2(K1)}$, $E_{B14(K1)}$, $\sigma'_{Z(K1)}$,
		$A'_{\Delta E_d(K2)}$, $E_{P2(K2)}$, $\nu'_{W3(K1)}$, $\nu'_{B12(K1)}$, $\nu'_{B7(K1)}$, $E_{B8(K1)}$, $E_{B9(K1)}$,
		$E_{B3(K1)}$, $E_{W2(K1)}$, $\nu'_{W1(K1)}$, $\nu'_{B8(K1)}$, $\nu'_{W4(K1)}$, $\nu'_{B9(K1)}$, $\nu'_{B3(K1)}$, $f'_{VRC\text{-}TST,c\ B6c(K1)}$ $E_{B6c(K1)}$,
		$E_{B7(K1)}$, $\nu'_{W5c(K1)}$, $\nu'_{W4(K1)}$, $\nu'_{imag\ B13(K1)}$, $\nu'_{imag\ B12(K1)}$, $\nu'_{B5(K1)}$, $\nu'_{W5t(K1)}$, $E_{P1(K1)}$
		$E_{W3(K1)}$, $\nu'_{P1(K1)}$, $E_{B5(K1)}$, $E_{B12(K1)}$, $\nu'_{B10(K1)}$, $\nu'_{B6t(K1)}$, $\nu'_{B6c(K1)}$, $E_{B13(K1)}$,
		$f'_{VRC\text{-}TST,c\ B0(K1)}$, $f'_{VRC\text{-}TST,c\ B1(K1)}$,$n_{\Delta E_d(K1)}$, $f'_{VRC\text{-}TST,c\ B4(K1)}$
		$\nu'_{imag\ B6t(K1)}$, $f'_{VRC\text{-}TST,c\ B5(K1)}$, $f'_{VRC\text{-}TST,c\ B2(K1)}$, $f'_{VRC\text{-}TST,c\ B3(K1)}$, $\nu'_{imag\ B14(K1)}$
K-R2	$CH_2O + O_2 = HCO + HO_2$	$\nu'_{B1(K2)}$, $\eta'_{B1(K2)}$, $\nu'_{imag\ B1(K2)}$, $\nu'_{P1(K2)}$, $E_{B1(K2)}$, $E_{P1(K2)}$
L-R1	$HCO = H + CO$	$\nu'_{B1(L1)}$, $E_{B1(L1)}$, $\nu'_{W1(L1)}$, $\nu'_{imag\ B1(L1)}$, $E_{W1(L1)}$, $E_{P0(L1)}$, $\nu'_{P0(L1)}$
L-R2	$HCO + O_2 = HO_2 + CO$	$E_{P1(L2)}$, $\nu'_{P1(L2)}$, $E_{W0(L2)}$, $\nu'_{W0(L2)}$, $\nu'_{W0(L2)}$, $f'_{VRC\text{-}TST,c\ B1(L2)}$, $\nu'_{B1(L2)}$, $E_{B1(L2)}$
M-R1	$H + CH_2O = HCO + H_2$	$\sigma'_{Z(M1)}$, $n_{\Delta E_d(M1)}$, $A'_{\Delta E_d(M1)}$, $\nu'_{B1(M1)}$, $E_{B1(M1)}$, $E_{P1(M1)}$, $E_{P2(M1)}$
		$\nu'_{W1(M1)}$, $\nu'_{P2(M1)}$, $E_{B2(M1)}$, $\nu'_{imag\ B2(M1)}$, $\nu'_{B2(M1)}$, $\nu'_{P1(M1)}$
M-R2	$H + HCO = H_2 + CO$	$A'_{(R2)}$
N-R1	$C_2H_5I\ (+M) = C_2H_5 + I\ (+M)$	$A'_{(N1)}$
N-R2	$C_2H_5I\ (+M) = C_2H_4 + HI\ (+M)$	$A'_{(N2)}$
O-R1	$OH + CH_2O = HCO + H_2O$	$\nu'_{B1(O1)}$, $E_{B1(O1)}$, $\sigma'_{Z(O1)}$, $n_{\Delta E_d(O1)}$, $A'_{\Delta E_d(O1)}$,
		$\nu'_{P1(O1)}$, $\nu'_{B0(O1)}$, $\nu'_{P0(O1)}$, $E_{B0(O1)}$, $E_{P1(O1)}$
P-R1a	$CO + OH = CO_2 + H$	$\nu'_{W2(P1)}$, $E_{W2(P1)}$, $\nu'_{W1(P1)}$, $\nu'_{P2(P1)}$, $\nu'_{B1(P1)}$, $E_{B1(P1)}$, $\nu'_{B3(P1)}$,
P-R1b	$CO + OH = HOCO$	$\sigma'_{Z(P1)}$, $n_{\Delta E_d(P1)}$, $A'_{\Delta E_d(P1)}$, $E_{P2(P1)}$, $\nu'_{B2(P1)}$, $\nu'_{imag\ B3(P1)}$, $\nu'_{W3(P1)}$,
P-R1c	$CO_2 + H = HOCO$	$E_{W1(P1)}$, $E_{B3(P1)}$, $\nu'_{imag\ B2(P1)}$, $E_{W3(P1)}$, $\nu'_{P1(P1)}$, $E_{B2(P1)}$, $E_{P1(P1)}$
Q-R1	$CH_3 + O_2 = CH_3O + O$	$A'_{(Q1)}$
Q-R2	$CH_3 + O_2 = CH_2O + OH$	$A'_{(Q2)}$
R-R1	$2H + M = H_2 + M$	$A'_{(R1)}$

[a] Note that ' indicates ln() of that quantity.

Table 2: List of physical model parameters considered in the optimization.[a]

Macroscopic Observables		Physical Model Parameters	
E1-E2	Shock-heated $H_2/O_2/H/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{H,o,e}$	$e = 1...2$
E3-E5	Shock-heated $H_2/O_2/H/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{H,o,e}$	$e = 3...5$
E6	Shock-heated $H_2/O_2/H/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{H,o,e}$	$e = 6$
E7-E9	Flash-photolyzed $H_2O/O_2/OH/H/Ar$	$T'_e, P'_e, M'_{H_2O,o,e}, M'_{O_2,o,e}, M'_{OH,o,e}, M'_{H,o,e}$	$e = 7...9$
E10-E17	Shock-heated $CH_2O/O_2/H/Ar$	$T'_e, P'_e, M'_{CH_2O,o,e}, M'_{O_2,o,e}, M'_{H,o,e}$	$e = 10...17$
E18-E22	Flash-photolyzed $H/H_2O/Ar$	$T'_e, P'_e, M'_{H,o,e}, M'_{H_2O,o,e}$	$e = 18...22$
E23	Shock-heated $C_4H_{10}O_2/H_2/He$	$T'_e, P'_e, M'_{C_4H_{10}O_2,o,e}, M'_{H_2O,o,e}$	$e = 23$
E24-E30	Flash-photolyzed $O/H_2/He$	$T'_e, P'_e, M'_{O,o,e}, M'_{H_2,o,e}$	$e = 24...30$
E31-E35	Shock-heated $O/H_2/He$	$T'_e, P'_e, M'_{C_4H_{10}O_2,o,e}, M'_{H_2,o,e}, M'_{H_2O,o,e}, M'_{H,o,e}$	$e = 31...35$
E36-E39	Shock-heated $H_2O_2/H_2O/O_2/Ar$	$T'_e, P'_e, M'_{H_2O_2,o,e}, M'_{H_2O,o,e}, M'_{O_2,o,e}$	$e = 36...39$
E40	Shock-heated $H_2O/O_2/H/Ar$	$T'_e, P'_e, M'_{H,o,e}, M'_{H_2O,o,e}, M'_{O_2,o,e}$	$e = 40$
E41-E45	Shock-heated H_2O_2/Ar	$T'_e, P'_e, M'_{H_2O_2,o,e}, \sigma'_{1,H_2O_2}, \sigma'_{2,H_2O_2}, \sigma'_{1,HO_2}, \sigma'_{2,HO_2}$	$e = 41...45$
E46-E50	Shock-heated $H_2O_2/H_2O/O_2/Ar$	$T'_e, P'_e, M'_{H_2O_2,o,e}, M'_{H_2O,o,e}, M'_{O_2,o,e}$, $\hat{\sigma}'_{1,H_2O_2}, \hat{\sigma}'_{2,H_2O_2}, \hat{\sigma}'_{1,HO_2}, \hat{\sigma}'_{2,HO_2}$	$e = 46...50$
E51-E59	Flash-photolyzed $Cl/Cl_2/CH_3OH/O_2/Ar$	$T'_e, P'_e, M'_{Cl,o,e}, M'_{Cl_2,o,e}, M'_{CH_3OH,o,e}, M'_{O_2,o,e}$, $\hat{\sigma}'_{1,H_2O_2}, \hat{\sigma}'_{2,H_2O_2}, \hat{\sigma}'_{1,HO_2}, \hat{\sigma}'_{2,HO_2}$	$e = 51...59$
E60-E61	Shock-heated $H_2/O_2/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}$	$e = 60...61$
E62	Shock-heated $H_2/O_2/NO/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{NO,o,e}$	$e = 62$
E63-E79	Shock-heated $H_2/O_2/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, \hat{\sigma}'_{1,H}$	$e = 63...79$
E80	Shock-heated $H_2/O_2/H/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{H,o,e}$	$e = 80$
E81	Flow-reactor $H_2/O_2/NO/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{NO,o,e}$	$e = 81$
E82-E84	Shock-heated $H_2/O_2/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}$	$e = 82...84$
E85	Flow-reactor $H_2/O_2/NO/Ar$	$T'_e, P'_e, M'_{H_2,o,e}, M'_{O_2,o,e}, M'_{NO,o,e}$	$e = 85$
E86	Shock-heated $C_2H_5I/NO_2/Ar$	$T'_e, P'_e, M'_{C_2H_5I,o,e}, M'_{NO_2,o,e}$	$e = 86$
E87	Shock-heated $H_2O_2/H_2O/O_2/CH_4/H/Ar$	$T'_e, P'_e, M'_{H_2O_2,o,e}, M'_{HO_2,o,e}, M'_{H_2O,o,e}, M'_{O_2,o,e}$ $M'_{H,o,e}, \hat{\sigma}^{\dagger\prime}{}'_{1,H_2O_2}, \hat{\sigma}^{\dagger\prime}{}'_{1,HO_2}$	$e = 87$
E88	Shock-heated C_2H_6/Ar	$T'_e, P'_e, M'_{C_2H_6,o,e}$	$e = 88$
E89-E106	Shock-heated CH_3/Ar	$T'_e, P'_e, M'_{CH_3,o,e}, \sigma^{\dagger\dagger\prime}{}'_{1,1,CH_3}, \sigma^{\dagger\dagger\prime}{}'_{2,CH_3}$	$e = 89...106$
E107-E115	Shock-heated CH_3/Ar	$T'_e, P'_e, M'_{CH_3,o,e}$	$e = 107...115$
E116	Shock-heated C_2H_6N/Ar	$T'_e, P'_e, M'_{C_2H_6N_2,o,e}$	$e = 116$
E117	Shock-heated $C_2H_6/CH_4/Ar$	$T'_e, P'_e, M'_{C_2H_6,o,e}, M'_{CH_4,o,e}$	$e = 117$
E118	Shock-heated CH_3I/Ar	$T'_e, P'_e, M'_{CH_3I,o,e}$	$e = 118$

[a]Note that ' indicates ln() of that quantity.

and initial species mole fractions $M_{i,o,e}$, coefficients for temperature-dependent cross sections of necessary species, and time shift $t_{i,o,e}$ (see Table 2). Rate constants and brute-force sensitivity coefficients were calculated using an automated wrapper [86] for Variflex [82] and MESS [81]. Macroscopic observables and sensitivity coefficients were calculated via homogeneous, isochoric or isobaric models in Cantera [83]. Much like our previous works, active parameters to represent the reactions of interest were either molecular parameters (if the authors had access to a validated MESS or Variflex file) or if no theoretical kinetics calculations were readily available for a reaction,

Table 2: Continued list of physical model parameters used in the optimization.[a]

Macroscopic Observables		Physical Model Parameters	
E119	Shock-heated $CH_3COCH_3/C_4H_{10}O_2/H_2O/Ar$	$T'_e, P'_e, M'_{CH_3COCH_3,o,e}, M'_{C_4H_{10}O_2,o,e}, M'_{H_2O,o,e}$	$e = 119$
E120-E122	Shock-heated $CH_3OH/O_2/Ar$	$T'_e, P'_e, M'_{CH_3OH,o,e}, M'_{O_2,o,e}$	$e = 120...122$
E123-E131	Shock-heated CH_3OH/Ar	$T'_e, P'_e, M'_{CH_3OH,o,e}$	$e = 123...131$
E132-E134	Shock-heated $C_4H_{10}O_2/C_2H_6N_2/H/Ar$	$T'_e, P'_e, M'_{C_4H_{10}O_2,o,e}, M'_{C_2H_6N_2,o,e}, M'_{H,o,e}$	$e = 132...134$
E135-E136	Shock-heated CH_2O/Ar	$T'_e, P'_e, M'_{CH_2O,o,e}$	$e = 135...136$
E137-E146	Shock-heated CH_2O/Ar	$T'_e, P'_e, M'_{CH_2O,o,e}$	$e = 137...146$
E147-E148	Shock-heated $CH_2O/H/Ar$	$T'_e, P'_e, M'_{CH_2O,o,e}$	$e = 147...148$
E149	Shock-heated $CH_2O/O_2/H/Ar$	$T'_e, P'_e, M'_{CH_2O,o,e}, M'_{O_2,o,e}$	$e = 149$
E150-E152	Shock-heated $CH_2O/O_2/H/Ar$	$T'_e, P'_e, M'_{CH_2O,o,e}, M'_{O_2,o,e}, M'_{H,o,e}$	$e = 150...152$
E153-E154	Shock-heated OCHCHO/Ar	$T'_e, P'_e, M'_{OCHCHO,o,e}$	$e = 153...154$
E155-E157	Shock-heated OCHCHO/Ar	$T'_e, P'_e, M'_{OCHCHO,o,e}$	$e = 155...157$
E158-E161	Shock-heated $CH_2O/C_2H_5I/Ar$	$T'_e, P'_e, M'_{CH_2O,o,e}, M'_{C_2H_5I,o,e}$	$e = 158...161$
E162-E164	Shock-heated $CH_2O/C_2H_5I/Ar$	$T'_e, P'_e, M'_{CH_2O,o,e}, M'_{C_2H_5I,o,e}$, $\hat{\sigma}'_{1,CO_2}, \hat{\sigma}'_{2,CO_2}, \hat{\sigma}'_{1,CH_2O_2}, \hat{\sigma}'_{2,CH_2O}$	$e = 162...164$
E165-E167	Shock-heated CH_2O/Ar	$T'_e, P'_e, M'_{CH_2O,o,e}$	$e = 165...167$
E168-E169	Shock-heated C_2H_5I/Ar	$T'_e, P'_e, M'_{C_2H5_1,o,e}$	$e = 168...169$
E170-E172	Shock-heated C_2H_5I/Ar	$T'_e, P'_e, M'_{C_2H5_1,o,e}$	$e = 170...172$
E173	Shock-heated C_2H_5I/Ar	$T'_e, P'_e, M'_{C_2H5_1,o,e}$	$e = 173$
E174	Shock-heated $C_4H_{10}O_2/CH_2O/Ar$	$T'_e, P'_e, M'_{C_4H_{10}O_2,o,e}, M'_{CH_2O,o,e}$	$e = 174$
E175-E176	Shock-heated $HNO_3/CO/Ar$	$T'_e, P'_e, M'_{HNO_3,o,e}, M'_{CO,o,e}$	$e = 175...176$
E177-E179	Shock-heated $CH_3I/O_2/Ar$	$T'_e, P'_e, M'_{CH_3I,o,e}, M'_{O_2,o,e}$	$e = 177...179$
E180	Shock-heated $CH_3I/O_2/Ar$	$T'_e, P'_e, M'_{CH_3I,o,e}, M'_{O_2,o,e}$	$e = 180$
E181-E182	Shock-heated $CH_3I/O_2/Ar$	$T'_e, P'_e, M'_{CH_3I,o,e}, M'_{O_2,o,e}$	$e = 181...182$
E183	Shock-heated $CH_3I/O_2/Ar$	$T'_e, P'_e, M'_{CH_3I,o,e}, M'_{O_2,o,e}$	$e = 183$

Note that ' indicates ln() of that quantity.

Arrhenius parameters were used. If the authors believed the reaction was well constrained across the full temperature range, the pre-exponential factor (A), temperature exponent (n), and activation energy (Ea) were used as active parameters; alternatively only the pre-exponential factor (A) was used.

The active parameters used to represent the reactions of interest are listed next to their corresponding reactions in Table 2. To clarify the labeling schema, for example B-R2a - B-R2j are all reactions on the same potential energy surface (PES) and share the same active molecular parameters. The molecular parameters included in the study consisted of energies of each stationary point (E), scaling factors for all harmonic frequencies of each stationary point (ν), scaling factors for each hindered rotor potential (η), scaling factors for the imaginary frequency used in the Eckart tunneling correction for each tight transition state (ν_{imag}), state density correction factors to account

Table 3: List of kinetic targets considered in the optimization.[a]

I. *Ab initio* calculations	Source
$\sigma'_{Z(B1)}(1.1)$, $n_{\Delta E_d(B1)}(1.1)$, $A'_{\Delta E_d(B1)}(0.2)$, $\nu'_{(B1)}(0.045)$, $\eta'_{mr(B1)}(0.07)$, $\nu'_{B1(B1)}(0.045)$, $E_{B1(B1)}(0.2$ kcal/mol), $E_{P1(B1)}(0.2$ kcal/mol), $E_{W1a(B1)}(0.2$ kcal/mol), $\rho'_{\text{corr W1b(B1)}}(1.60)$, $\rho'_{\text{corr W1a(B1)}}(1.60)$, $\nu'_{W1a(B1)}(0.135)$, $E_{W1b(B1)}(0.2$ kcal/mol), $\nu'_{\text{imag B2(B1)}}(0.18)$, $\nu'_{P1(B1)}(0.045)$, $f'_{\text{VRC-TST,c P1(B1)}}(0.69)$, $\nu'_{P2(B1)}(0.045)$, $\nu'_{B8(B1)}(0.135)$, $\rho'_{\text{corr B8(B1)}}(1.60)$, $E_{P2(B1)}(0.2$ kcal/mol), $\nu'_{W1b(B1)}(0.135)$, $E_{B8(B1)}(0.2$ kcal/mol), $E_{B2(B1)}$, $E_{P1(B1)}(0.2$ kcal/mol), $f'_{\text{VRC-TST,c B1(B1)}}(0.69)$	[113]
$\sigma'_{Z(B2)}(1.1)$, $n_{\Delta E_d(B2)}(1.1)$, $A'_{\Delta E_d(B2)}(0.2)$, $E_{B9(B2)}(1.0$ kcal/mol), $E_{P1(B2)}(0.2$ kcal/mol), $E_{B4(B2)}(0.2)$, $\nu'_{W2b(B2)}(0.135)$, $\nu'_{B7_{1(B2)}}(0.075)$, $\nu'_{B9(B2)}(0.075)$, $f'_{\text{VRC-TST,c B5(B2)}}(0.69)$, $\nu'_{B4(B2)}(0.045)$ $f'_{\text{VRC-TST,c B7(B2)}}(0.69)$, $\nu'_{W2a(B2)}(0.135)$, $E_{W2b(B2)}(1.0$ kcal/mol), $\nu'_{\text{imag B3(B2)}}(0.18)$, $f'_{\text{VRC-TST,c P1(B2)}}$ $E_{B3(B2)}(2.0$ kcal/mol), $E_{B7(B2)}(0.4$ kcal/mol), $E_{P4(B2)}(2.0$ kcal/mol), $\nu'_{P3(B2)}(0.045)$, $\rho'_{\text{corr B9(B2)}}$ $\nu'_{B3(B2)}(0.045)$, $E_{B6(B2)}(2.0$ kcal/mol), $\nu'_{B6(B2)}(0.045)$, $\nu'_{P5(B2)}(0.045)$, $E_{W3(B2)}(0.4$ kcal/mol), $E_{P3(B2)}(0.2$ kcal/mol), $\nu'_{P1(B2)}(0.045)$, $\rho'_{\text{corr W2a(B2)}}(1.60)$, $\rho'_{\text{corr P5(B2)}}(0.69)$, $\nu'_{P4(B2)}(0.045)$, $\nu'_{W3(B2)}(0.045)$, $\nu'_{B5(B2)}(0.075)$ $E_{P5(B2)}(0.4$ kcal/mol), $E_{B5(B2)}$, $\rho'_{\text{corr W2b(B2)}}(1.60)$, $\nu'_{\text{imag B6(B2)}}(0.18)$, $E_{W2a(B2)}(1.0$ kcal/mol), $E_{W2a(B2)}(1.0$ kcal/mol), $E_{P1(B2)}(0.2$ kcal/mol), $\eta'_{B7_1(B2)}(0.07)$, $\eta'_{B7_2(B2)}(0.07)$, $\eta'_{W3_1(B2)}(0.07)$, $\eta'_{W3_2(B2)}(0.07)$, $\eta'_{W3_3(B2)}(0.07)$, $\eta'_{P5(B2)}(0.07)$ $E_{B7_1(B2)}(0.4$ kcal/mol), $f'_{\text{VRC-TST,c B7}_2\text{(B2)}}(0.69)$, $\nu'_{B7_{2(B2)}}$, $f'_{\text{VRC-TST,c B4(B2)}}(0.69)$	[113]
$E_{B1(B3)}(3$ kcal/mol), $\nu'_{B1(B3)}(0.1)$, $\eta'_{B1(B3)}(0.1)$, $\nu'_{\text{imag B1(B3)}}(0.18)$ $E_{B1(B4)}(1.4$ kcal/mol), $E_{B2(B4)}(1.4$ kcal/mol), $\nu'_{B1,\,B2(B4)}(0.1)$, $\nu'_{\text{imag B1,B2(B4)}}(0.18)$ $f'_{\text{VRC-TST,c B1(B5)}}(0.7)$, $E_{B2(B5)}(2.0$ kcal/mol), $\nu'_{B2(B5)}(0.1)$, $\eta'_{B2(B5)}(0.41)$, $\nu'_{\text{imag B2(B5)}}(0.18)$, $\nu'_{B2(B5)}(0.1)$, $\eta'_{B3(B5)}(0.41)$, $\nu'_{\text{imag B3(B5)}}(0.18)$, $\nu'_{W1(B5)}(0.1)$, $E_{B3(B5)}(2.0$ kcal/mol)	Harding and Klippenstein Unpublished
$\nu'_{B1(C1)}(0.1)$, $\nu'_{\text{imag B1(C1)}}(0.18)$, $E_{B1(C1)}(1.1$ kcal/mol) $\nu'_{W2(C3)}(0.1)$, $E_{W2(C3)}(2.0$ kcal/mol), $\nu'_{W1(C3)}(0.1)$, $E_{B5(C3)}(2.0$ kcal/mol), $E_{B5(C3)}(2.0$ kcal/mol), $E_{B6(C3)}(2.0$ kcal/mol) $\nu'_{B5(C3)}(0.1)$, $\nu'_{B6(C3)}(0.1)$, $\nu'_{W3(C3)}(0.1)$, $\eta'_{B1(C3)}(0.7)$, $E_{B2(C3)}(2.0$ kcal/mol), $\eta'_{B2(C3)}(0.7)$, $E_{W2(C3)}(2.0$ kcal/mol), $E_{B3(C3)}(2.0$ kcal/mol), $\eta'_{B3(C3)}(0.7)$, $E_{B4(C3)}(2.0$ kcal/mol), $\eta'_{B4(C3)}(0.7)$,$\eta'_{B7(C3)}(0.7)$, $E_{B7(C3)}(2.0$ kcal/mol, $E_{B1(C3)}(2.0$ kcal/mol),	Klippenstein Unpublished
$\nu'_{B1(E1)}(0.1)$, $E_{B1(E1)}(1.1$ kcal/mol), $E_{B4(E1)}(1.1$ kcal/mol), $\nu'_{W1(E1)}(0.1)$, $\nu'_{B3(E1)}(0.1)$ $\nu'_{P1(E1)}(0.1)$, $\nu'_{B2(E1)}(0.1)$, $\sigma'_{Z(E1)}(0.26)$, $E_{W3(E1)}(1.1$ kcal/mol), $\nu'_{P2(E1)}(0.1)$, $\nu'_{B4(E1)}(0.1)$, $n_{\Delta E_d(E1)}(0.2)$, $A'_{\Delta E_d(E1)}(1.1)$, $E_{P1(E1)}(1.1$ kcal/mol), $f'_{\text{VRC-TST,c B4(E1)}}(0.69)$, $f'_{\text{VRC-TST,c B3(E1)}}(0.69)$, $E_{B2(E1)}(1.1$ kcal/mol), $E_{B3(E1)}(1.1$ kcal/mol), $E_{W1(E1)}(1.1$ kcal/mol), $E_{P2(E1)}(1.4$ kcal/mol), $\nu'_{\text{imag B1(E1)}}(0.1)$, $f'_{\text{VRC-TST,c B2(B2)}}(0.69)$	[147]
$E_{B1(F1)}(1$ kcal/mol), $\nu'_{\text{low B1(F1)}}(0.1)$, $(Q_{\text{anh,c,CH}_3\text{(F1)}})(1)$, $(Q_{\text{anh,c,CH}_3\text{(F2)}})(1)$, $f'_{\text{VRC}-\text{TST(F2)}}(0.262)$	[96, 97]
$\nu'_{B1(G1)}(0.1)$, $E_{B1(G1)}(2.0$ kcal/mol), $\nu'_{W1(G1)}(0.1)$, $\nu'_{B3(G1)}(0.1)$, $\nu'_{B2(G1)}(0.1)$, $f'_{\text{VRC-TST,c B1(G1)}}(0.69)$, $\sigma'_{Z(G1)}(0.262)$, $n_{\Delta E_d(G1)}(0.2)$, $A'_{\Delta E_d(G1)}(1.1)$, $\nu'_{P1(G1)}(0.1)$, $E_{P3(G1)}(2.0$ kcal/mol), $\eta'_{W1(G1)}(0.07)$, $\nu'_{P2(G1)}(0.1)$, $\eta'_{W1(G1)}(0.07)$, $E_{W1(G1)}(2.0$ kcal/mol), $\nu'_{\text{imag B3(G1)}}(0.18)$, $\eta'_{B2(G1)}(0.07)$, $E_{B3(G1)}(2.0$ kcal/mol), $\nu'_{P3(G1)}(0.1)$, $E_{P2(G1)}(2.0$ kcal/mol), $\nu'_{B1(G2)}(0.1)$, $E_{B1(G2)}(2.0$ kcal/mol), $\nu'_{P1(G2)}(0.1)$, $E_{B2(G1)}(2.0$ kcal/mol), $E_{P1(G1)}(2.0$ kcal/mol)	[148]
$E_{B1(I1)}(2.0$ kcal/mol), $E_{B4(I1)}(2.0$ kcal/mol), $\nu'_{W1(I1)}(0.1)$, $\sigma'_{Z(I1)}(0.262)$, $n_{\Delta E_d(I1)}(0.2)$, $A'_{\Delta E_d(I1)}(1.1)$, $E_{P7(I1)}(2.0$ kcal/mol), $E_{R(I1)}(2.0$ kcal/mol), $f'_{\text{VRC-TST,c B6(I1)}}(0.69)$ $\nu'_{B8(I1)}(0.1)$, $\nu'_{B4(I1)}(0.1)$, $E_{B3(I1)}(2.0$ kcal/mol), $E_{B7(I1)}(2.0$ kcal/mol), $E_{B2(I1)}(2.0$ kcal/mol), $\nu'_{P3(I1)}(0.1)$, $\nu'_{B3(I1)}(0.1)$, $E_{P1(I1)}(2.0$ kcal/mol), $E_{P2(I1)}(2.0$ kcal/mol), $E_{B6(I1)}(2.0$ kcal/mol), $\nu'_{P7(I1)}(0.1)$, $\nu'_{P5(I1)}(0.1)$, $\nu'_{B5(I1)}(0.1)$, $E_{P6(I1)}(2.0$ kcal/mol), $\nu'_{P7(I1)}(0.1)$, $\nu'_{P1(I1)}(0.1)$, $\nu'_{R(I1)}(0.1)$, $\nu'_{P4(I1)}(0.1)$, $E_{B5(I1)}(2.0$ kcal/mol), $E_{B8(I1)}(2.0$ kcal/mol), $\nu'_{P6(I1)}(0.1)$, $E_{P5(I1)}(2.0$ kcal/mol), $f'_{\text{VRC-TST,c B7(I1)}}(0.69)$, $f'_{\text{VRC-TST,c B1(I1)}}(0.69)$ $\nu'_{B1(I2)}(0.1)$, $E_{B1(I2)}(2.0$ kcal/mol), $\nu'_{P7(I2)}(0.1)$, $\nu'_{\text{imag B1(I2)}}(0.18)$, $\nu'_{R(I2)}(0.1)$,$\nu'_{P2(I1)}(0.1)$, $E_{P3(I1)}(2.0$ kcal/mol), $E_{P4(I1)}(2.0$ kcal/mol), $f'_{\text{VRC-TST,c B2(I1)}}(0.69)$,	[149]
$\nu'_{W2(K1)}(0.1)$, $\nu'_{B1(K1)}(0.1)$, $E_{B4(K1)}(0.2$ kcal/mol), $E_{B1(K1)}(0.2$ kcal/mol), $\nu'_{B2(K1)}(0.1)$, $\nu'_{\text{imag B15(K1)}}(0.18)$, $E_{B6t(K1)}(0.2$ kcal/mol), $E_{W5c(K1)}(0.2$ kcal/mol), $\nu'_{B13(K1)}(0.1)$, $E_{B10(K1)}(0.2$ kcal/mol), $E_{B11(K1)}(0.2$ kcal/mol), $\nu'_{B11(K1)}(0.1)$, $f'_{\text{VRC-TST,c B0(K1)}}(0.69)$, $f'_{\text{VRC-TST,c B1(K1)}}(0.69)$ $E_{B2(K1)}(0.2$ kcal/mol), $E_{B15(K1)}(0.2$ kcal/mol), $\nu'_{B15(K1)}(0.1)$, $E_{W1(K1)}(0.2$ kcal/mol), $E_{W0(K1)}(0.2$ kcal/mol), $f'_{\text{VRC-TST,c B2(K1)}}(0.69)$, $f'_{\text{VRC-TST,c B3(K1)}}(0.69)$ $\nu'_{B14(K1)}(0.1)$, $\nu'_{W0(K1)}(0.1)$, $E_{W4(K1)}(0.2$ kcal/mol), $\nu'_{B4(K1)}(0.1)$, $E_{W5t(K1)}(0.2$ kcal/mol), $\nu'_{P2(K1)}(0.1)$, $E_{B14(K1)}(0.2$ kcal/mol), $\sigma'_{Z(K1)}(1.1)$, $n_{\Delta E_d(K1)}(1.1)$, $f'_{\text{VRC-TST,c B4(K1)}}(0.69)$ $A'_{\Delta E_d(K2)}(0.2)$, $E_{P2(K2)}(0.2$ kcal/mol), $\nu'_{W3(K1)}(0.1)$, $\nu'_{B12(K1)}(0.1)$, $\nu'_{B7(K1)}(0.1)$, $E_{B8(K1)}(0.2$ kcal/mol), $E_{B9(K1)}(0.2$ kcal/mol), $\nu'_{\text{imag B6t(K1)}}(0.18)$, $f'_{\text{VRC-TST,c B5(K1)}}(0.69)$ $E_{B3(K1)}(0.2$ kcal/mol), $E_{W2(K1)}(0.2$ kcal/mol), $\nu'_{W1(K1)}$, $\nu'_{B8(K1)}(0.1)$, $\nu'_{W4(K1)}(0.1)$, $\nu'_{B9(K1)}(0.1)$, $\nu'_{B3(K1)}(0.1)$, $f'_{\text{VRC-TST,c B6c(K1)}}(0.69)$, $E_{B6c(K1)}(0.2$ kcal/mol), $\nu'_{\text{imag B14(K1)}}(0.18)$ $E_{B7(K1)}(0.2$ kcal/mol), $\nu'_{W5c(K1)}(0.1)$, $\nu'_{W4(K1)}(0.1)$, $\nu'_{\text{imag B13(K1)}}(0.1)$, $\nu'_{\text{imag B12(K1)}}(0.18)$, $\nu'_{B5(K1)}(0.1)$, $\nu'_{W5t(K1)}(0.1)$, $E_{B13(K1)}(0.2$ kcal/mol), $E_{W3(K1)}(0.2$ kcal/mol), $\nu'_{P1(K1)}(0.1)$, $E_{B5(K1)}(0.2$ kcal/mol), $E_{B12(K1)}(0.2$ kcal/mol), $\nu'_{B10(K1)}(0.1)$, $\nu'_{B6t(K1)}(0.1)$, $\nu'_{B6c(K1)}(0.1)$, $E_{P1(K1)}(0.2$ kcal/mol) $\nu'_{B1(K2)}(0.1)$, $\eta'_{B1(K2)}(0.07)$, $\nu'_{\text{imag B1(K2)}}(0.18)$, $\nu'_{P1(K2)}(0.1)$, $E_{B1(K2)}(3.0$ kcal/mol), $E_{P1(K2)}(3.0$ kcal/mol)	Klippenstein Unpublished
$\nu'_{B1(L1)}(0.1)$, $E_{B1(L1)}(0.5$ kcal/mol), $\nu'_{W1(L1)}(0.1)$, $\nu'_{\text{imag B1(L1)}}(0.18)$, $E_{W1(L1)}(1.1$ kcal/mol), $E_{P0(L1)}(0.5$ kcal/mol), $\nu'_{P0(L1)}(0.1)$ $E_{P1(L2)}(3.0$ kcal/mol) , $\nu'_{P1(L2)}(0.26)$, $E_{W0(L2)}(3.0$ kcal/mol), $\nu'_{W0(L2)}(0.26)$, $\nu'_{W0(L2)}(0.26)$ $f'_{\text{VRC-TST,c B1(L2)}}(0.7)$, $\nu'_{B1(L2)}(0.26)$, $E_{B1(L2)}(3.0$ kcal/mol)	[47] Klippenstein Unpublished
$\sigma'_{Z(M1)}(0.26)$, $n_{\Delta E_d(M1)}(0.2)$, $A'_{\Delta E_d(M1)}(1.1)$, $\nu'_{B1(M1)}(0.1)$, $E_{B1(M1)}(0.5$ kcal/mol), $E_{P1(M1)}(0.5$ kcal/mol), $E_{P2(M1)}(0.5$ kcal/mol) $\nu'_{W1(M1)}(0.1)$, $\nu'_{P2(M1)}(0.1)$, $E_{B2(M1)}(0.5$ kcal/mol), $\nu'_{\text{imag B2(M1)}}(0.18)$, $\nu'_{B2(M1)}(0.1)$, $\nu'_{P1(M1)}(0.1)$	Klippenstein Unpublished
$\nu'_{B1(O1)}(0.1)$, $E_{B1(O1)}(0.5$ kcal/mol), $\sigma'_{Z(O1)}(0.26)$, $n_{\Delta E_d(O1)}(0.2)$, $A'_{\Delta E_d(O1)}(1.1)$, $\nu'_{P1(O1)}(0.1)$, $\nu'_{B0(O1)}(0.1)$, $\nu'_{P0(O1)}(0.1)$, $E_{B0(O1)}(0.5$ kcal/mol), $E_{P1(O1)}(0.5$ kcal/mol)	Klippenstein Unpublished
$\nu'_{W2(P1)}(0.1)$, $E_{W2(P1)}(2.0$ kcal/mol), $\nu'_{W1(P1)}(0.1)$, $\nu'_{P2(P1)}(0.1)$, $\nu'_{B1(P1)}(0.1)$, $E_{B1(P1)}(2.0$ kcal/mol), $\nu'_{B3(P1)}(0.1)$, $\sigma'_{Z(P1)}(0.26)$, $n_{\Delta E_d(P1)}(0.1)$, $A'_{\Delta E_d(P1)}(1.1)$, $E_{P2(P1)}(2.0$ kcal/mol), $\nu'_{B2(P1)})(0.1)$, $\nu'_{\text{imag B3(P1)}}(0.18)$, $\nu'_{W3(P1)}(0.1)$, $E_{W1(P1)}(2.0$ kcal/mol), $E_{B3(P1)}(2.0$ kcal/mol), $\nu'_{\text{imag B2(P1)}}(0.18)$, $E_{W3(P1)}(2.0$ kcal/mol), $\nu'_{P1(P1)}(0.1)$, $E_{B2(P1)}(2.0$ kcal/mol), $E_{P1(P1)}(2.0$ kcal/mol)	Klippenstein Unpublished

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

for any significant uncertainties in the treatment of anharmonicity of some stationary points (ρ_{corr}), VRC-TST correction factors to account for uncertainties in the treatment of barrierless transition

Table 3: List of kinetic targets considered in the optimization continued.[a]

II. Rate constant measurements	
k'_{A1}, k'_{A2}, k'_{A3}	[150, 151, 136, 152]
k'_{B3}, k'_{B4}, $(k'_{B5a} + k'_{B5b})$	[69]
$(k'_{C1} + k'_{C2} + k'_{C3a} + k'_{C3b} + k'_{C3c})$, $(k'_{C3a}/(k'_{C1} + k'_{C2} + k'_{C3a} + k'_{C3b} + k'_{C3c}))$, $((k'_{C3b} + k'_{C3c})/(k'_{C1} + k'_{C2} + k'_{C3a} + k'_{C3b} + k'_{C3c}))$,	[153, 154]
k'_{E1}	[155, 156]
k'_{F1}, k'_{F3}	[96, 100, 101]
k'_{G1a}	[157, 158]
$k'_{I1e}/(k'_{IR2} + k'_{I1e} + k'_{I1d})$, $k'_{IR2} + k'_{I1e} + k'_{I1d}$,	[136]
k'_{L1}, k'_{L2},	[159, 160, 161, 162]
k'_{M1}	[163, 164]
k'_{O1}	[165]

[a]Note that $'$ indicates ln() of the quantity.
[b]Uncertainties listed in () are intended to reflect two standard deviations.

states ($f_{\text{VRC-TST}}$), and scaling factors for the parameters describing a temperature-dependent average energy per down collision, $\Delta E_d = AT^n$, in the exponential-down model ($A_{\Delta E_d}$ and $n_{\Delta E_d}$) and the cross section of the well in the Lennard-Jones collision model (σ_Z) to account for uncertainties in the treatment of collisional energy transfer and if there is more than one exponential the Fraction (β) to define the fraction of each. Additionally, for reactions F-R1 a scaling factor for the lowest four harmonic frequencies in the transition state (ν_{low}), and $Q_{\text{anh,c,CH}_3}$ a scaling factor for aharmonic correction were assigned as active parameters. Additionally, for F-R2 VRC-TST correction factors to account for uncertainties in the treatment of barrierless transition states ($f_{\text{VRC-TST}}$), as well as $Q_{\text{anh,c,CH}_3}$ were included as active parameters.

The theoretical kinetics calculations of Klippenstein et al. [113] and Harding and Klippenstein (Unpublished) were used for the active kinetic treatment of $HO_2 + HO_2$ reactions (B-R1 and B-R2) and other H_2O_2 decomposition system reactions (B-R3 - B-R5), respectively. Further details regarding these calculations can be found in Ref. [146]. Calculations of Klippenstein (Unpublished) were used for for the active kinetic treatment of C-R1, calculations of Chen et al. [147] were used for for the active kinetic treatment of E-R1, calculations of Srinivasan et al. [96] and Jasper et al. [97] were used for the active kinetic parameter treatment of F-R1 and F-R2 respectively, (more information regarding these calculations can be found in Ref. [72]), calculations of Klippenstein et al [148] were used for the active kinetic parameter treatment of G-R1 and G-R2, calculations of Jasper et al [149] were used for the active kinetic parameter treatment of I-R1 and I-R2, calculations of Labbe et al. [47] were used for the active kinetic treatment of L-R1, and calculations of

Table 4: List of physical targets considered in the optimization.[a,b]

III. Global exp.	IV. Exp. conditions		
$M'_{H_2O,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05), M'_{O_2,o,e}(0.05), M'_{H,o,e}(0.1)$	$e = 1...2$	[166]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{O_2,o,e}(0.05), M'_{H_2,o,e}(0.05), M'_{H,o,e}(0.1)$	$e = 3...5$	[167]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{O_2,o,e}(0.05), M'_{H_2,o,e}(0.05), M'_{H,o,e}(0.05), M'_{H,o,e}(0.1)$	$e = 6$	[168]
$M'_{H,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{O_2,o,e}(0.05), M'_{H_2O,o,e}(0.05),$ $M'_{OH,o,e}(0.05), M'_{H,o,e}(0.1)$	$e = 7...9$	[169]
$M'_{OH,e}(t)(0.05)$ $M'_{CO,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_2O,o,e}(0.05), M'_{O_2,o,e}(0.1), M'_{H,o,e}(0.1)$	$e = 10...17$	[170]
$M'_{H,e}(t)(0.1),$	$T'_e(0.01), P'_e(0.02),$ $M'_{H,o,e}(0.05), M'_{H_2O,o,e}(0.05)$	$e = 18...22$	[171]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_4H_{10}O_2,o,e}(0.05), M'_{H_2,o,e}(0.05)$	$e = 23$	[172]
$M'_{O,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{O,o,e}(0.05), M'_{H_2,o,e}(0.05)$	$e = 24...30$	[173]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_4H_{10}O_2,o,e}(0.05), M'_{H_2,o,e}(0.05)$ $M'_{H_2,o,e}(0.05), M'_{H_2O,o,e}(0.05)M'_{H,o,e}(2.3)$	$e = 31...35$	[174]
$M'_{OH,e}(t)(0.05),$ $M'_{H_2O,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2O_2,o,e}(0.05), M'_{H_2O,o,e}(0.05), M'_{O_2,o,e}(0.05)$	$e = 36...39$	[102, 103]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2O,o,e}(0.1), M'_{O_2,o,e}(0.01), M'_{H,o,i}(2.3)$	$e = 40$	[104]
$abs'_e(t)(0.1)$	$T'_e(0.02), P'_e(0.04),$ $M'_{H_2O_2,o,e}(0.1)$ $\sigma'_{1,H_2O_2}(0.7), \sigma'_{2,H_2O_2}(0.3)$ $\sigma'_{a,HO_2}(0.7), \sigma'_{2,HO_2}(0.3)$	$e = 41...45$	[105]
$abs'_e(t)(0.1),$ $M'_{H_2O,e}(t)(0.05),$ $M'_{OH,e}(t)(0.05);$	$T'_e(0.02), P'_e(0.04),$ $M'_{H_2O_2,o,e}(0.05); M'_{H_2O,o,e}(0.2); M'_{O_2,o,e}(0.2),$ $\hat{\sigma}'_{1,H_2O_2}(0.7), \hat{\sigma}'_{2,H_2O_2}(0.7)$ $\hat{\sigma}'_{1,HO_2}(0.7), \hat{\sigma}'_{2,HO_2}(0.7)$	$e = 46...50$	[106]
$abs'_e(t)(0.1)$	$T'_e(0.02), P'_e(0.04),$ $M'_{Cl,o,e}(0.1), M'_{Cl_2,o,e}(0.05), M'_{CH_3OH,o,e}(0.05), M'_{O_2,o,e}(0.05),$ $\hat{\sigma}'_{1,H_2O_2}(0.7), \hat{\sigma}'_{2,H_2O_2}(0.7)$ $\hat{\sigma}'_{1,HO_2}(0.7), \hat{\sigma}'_{2,HO_2}(0.7)$	$e = 51...59$	[107]
$M'_{O,e}(t)(0.12)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05), M'_{O_2,o,i}(0.05)$	$e = 60...61$	[175]

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

Klippenstein (Unpublished) were used for the active kinetic parameters of M-R1 and O-R1.

It should also be noted the *a priori* fit for D-R1 was determined by using the high pressure limit

Table 4: List of physical targets considered in the optimization continued. [a,b]

III. Global exp.	IV. Exp. conditions		
$M'_{NO_2,e}(t)(0.15)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05); M'_{O_2,o,e}(0.05); M'_{NO,o,e}(0.05),$	$e = 62$	[176]
$M'_{abs,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05), M'_{O_2,o,e}(0.05),$	$e = 63...79$	[177]
$M'_{OH,e}(t)(0.15)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05); M'_{O_2,o,e}(0.05); M'_{H,o,e}(2.3),$	$e = 80$	[178]
$M'_{O_2,e}(t)(0.12), M'_{H_2O,e}(t)(0.12)$ $M'_{H_2,e}(t)(0.12), M'_{NO_2,e}(t)(0.12)$ $M'_{NO,e}(t)(0.12)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05); M'_{O_2,o,e}(0.05)$ $M'_{NO,o,i}(0.05), t'(0.0258)$	$e = 81$	[179]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05), M'_{O_2,o,e}(0.05)(0.05)$	$e = 82...84$	[180]
$M'_{O_2,e}(t)(0.1), M'_{H_2,e}(t)(0.1),$ $M'_{NO_2,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{H_2,o,e}(0.05), M'_{O_2,o,i}(0.05), t'(0.5)$	$e = 85$	[181]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2HI,o,e}(0.05), M'_{NO_2,o,e}(0.05)$	$e = 86$	[182]
$M'_{abs,e}(t)(0.1), M'_{H_2O,e}(t)(0.1),$ $M'_{OH,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{O_2,o,e}(0.05), M'_{H_2O_2,o,i}(0.5), M'_{H,o,i}(0.5)$ $M'_{H_2O,o,e}(0.05), M'_{O_2,o,i}(0.5), M'_{CH_4,o,i}(0.5)$ $\hat{\sigma}^{\dagger\prime}{}_{1,H_2O_2}, \hat{\sigma}^{\dagger\prime}{}_{1,HO_2}$	$e = 87$	[108]
$M'_{CH_3,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_6,o,e}(0.05)$	$e = 88$	[183]
$M'_{abs,e}(t)(0.12)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3,o,e}(0.05)\sigma^{\dagger}\dagger'_{1,1,CH_3}, \sigma^{\dagger}\dagger'_{2,CH_3}$	$e = 89...106$	[184]
$M'_{CH_3,e}(t)(0.08)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_6,o,e}(0.05)$	$e = 107...115$	[157]
$M'_{CH_3,e}(t)(0.08)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_6N,o,e}(0.05)$	$e = 116$	[185]
$M'_{CH_3,e}(t)(0.08)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_6,o,e}(0.05), M'_{CH_4,o,e}(0.05)$	$e = 117$	[185]
$M'_{CH_3,e}(t)(0.11)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3I,o,e}(0.05)$	$e = 118$	[185]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3COCH_3,o,e}(0.05), M'_{C_4H_{10}O_2,o,e}(0.05), M'_{H_2O,o,e}(0.05)$	$e = 119$	[186]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3OH,o,e}(0.05)$	$e = 120...122$	[187]

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

from Harding et al. [210], and the low pressure limit and F_{cent} values were fit from calculations presented in Klippenstein et al. [211].

Table 4: List of physical targets considered in the optimization continued.[a,b]

III. Global exp.	IV. Exp. conditions		
$M'_{\mathrm{H},e}(t)(0.10)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_3OH},o,e}(0.05)$	$e = 123...131$	[188]
$M'_{\mathrm{OH},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{C_4H_{10}O_2},o,e}(0.05), M'_{\mathrm{C_2H_6N_2},o,e}(0.05)$	$e = 132...134$	[189]
$M'_{\mathrm{CO},e}(t)(0.07)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_2O},o,e}(0.05)$	$e = 135...136$	[190]
$M'_{\mathrm{CH_2O},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_2O},o,e}(0.05)$	$e = 137...146$	[191]
$M'_{\mathrm{CH_2O},e}(t)(0.07)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_2O},o,e}(0.05), M'_{\mathrm{H},o,e}(0.05)$	$e = 147...148$	[192]
$M'_{\mathrm{H},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_2O},o,e}(0.05)$	$e = 149$	[193]
$M'_{\mathrm{OH},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_2O},o,e}(0.05), M'_{\mathrm{O_2},o,e}(0.05), M'_{\mathrm{H},o,e}(0.05)$	$e = 150...152$	[194]
$M'_{\mathrm{HCO},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{OCHCHO},o,e}(0.05)$	$e = 153...154$	[195]
$M'_{\mathrm{HCO},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{OCHCHO},o,e}(0.05)$	$e = 155...157$	[196]
$M'_{\mathrm{CO},e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{\mathrm{CH_2O},o,e}(0.05), M'_{\mathrm{C_2H_5I},o,e}(0.05)$	$e = 158...161$	[197]

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

Target Class (I): *ab initio* calculations

Molecular properties from *ab initio* calculations were used as targets for the molecular parameters of B-R1, B-R2, B-R3, B-R4, B-R5, C-R1, E-R1, F-R1, F-R2, G-R1, G-R2, I-R1, I-R2, K-R1, K-R2, L-R1, M-R1, and O-R1. Uncertainties listed in Table 3 reflect the level of electronic structure theory and the level of detail in the treatment. Specifically, barrier height uncertainty estimates for the various methods and basis set combinations were based on values presented in a publication from Klippenstein and Cavallotti [80].

Table 4: List of physical targets considered in the optimization continued.[a,b]

III. Global exp.	IV. Exp. conditions		
$M'_{CH_2O,e}(t)(0.1)$ $M'_{abs,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_2O,o,e}(0.05), M'_{C_2H_5I,o,e}(0.05)$	$e = 162...164$	[198]
$M'_{CH_2O,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_2O,o,e}(0.05)$	$e = 165...167$	[199]
$M'_{H,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_5I,o,e}(0.05)$	$e = 168...169$	[200]
$M'_{H,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_5I,o,e}(0.05)$	$e = 170...172$	[201]
$M'_{H,e}(t)(0.1)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_2H_5I,o,e}(0.05)$	$e = 173$	[202]
$M'_{OH,e}(t)(0.07)$	$T'_e(0.01), P'_e(0.02),$ $M'_{C_4H_{10}O_2,o,e}(0.05), M'_{CH_2O,o,e}(0.05)$	$e = 174$	[203]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CO,o,e}(0.05), M'_{HNO_3,o,e}(0.05)$	$e = 175...176$	[204]
$M'_{O,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3I,o,e}(0.05), M'_{O_2,o,e}(0.05)$	$e = 177...179$	[205]
$M'_{O,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3I,o,e}(0.05), M'_{O_2,o,e}(0.05)$	$e = 180$	[206]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3I,o,e}(0.05), M'_{O_2,o,e}(0.05)$	$e = 181...182$	[207]
$M'_{OH,e}(t)(0.05)$	$T'_e(0.01), P'_e(0.02),$ $M'_{CH_3I,o,e}(0.05), M'_{O_2,o,e}(0.05)$	$e = 181...182$	[208]

[a] Note that $'$ indicates ln() of the quantity.
[b] Uncertainties listed in () are intended to reflect two standard deviations.

Target Class (II): Rate constant determinations

Rate constant determinations (from studies other than those treated as macroscopic observables) were included as targets for, k_{A1}, k_{A2}, k_{A3}, k_{B3}, k_{B4}, $(k_{B5a} + k_{B5b})$, $(k_{C1} + k_{C2} + k_{C3a} + k_{C3b} + k_{C3c})$, $(k_{C3a}/(k_{C1} + k_{C2} + k_{C3a} + k_{C3b} + k_{C3c}))$, $((k_{C3b} + k_{C3c})/(k_{C1} + k_{C2} + k_{C3a} + k_{C3b} + k_{C3c}))$, k_{E1}, k_{F1}, k_{F3}, k_{G1a}, $k_{I1e}/(k_{I2} + k_{I1e} + k_{I1d})$, $k_{I2} + k_{I1e} + k_{I1d}$, k_{L1}, k_{L2}, k_{M1}, k_{O1}. Given that uncertainties in rate constant determinations are often higher than typically estimated and prone to systematic errors, [69, 71, 44] conservative uncertainties of a factor of two (or the scatter in the data if lager) were applied.

Target Class (III) and (IV): Macroscopic observables and experimental conditions

The raw data for the macroscopic observables from the experiments listed in Table 2 were included as targets using uncertainties as reported or estimated using typical values [166, 167, 169, 170, 171, 172, 173, 174, 103, 102, 104, 105, 106, 107, 175, 178, 179, 180, 177, 181, 182, 108, 183, 184, 185, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208]. Macroscopic observables used as targets included time-resolved H_2O, H, O, OH, CO, O_2, H_2, NO_2, NO, CH_3, CH_2O and HCO mole fractions, time-resolved O, H, OH, CO, HCO and CH_2O concentrations and absorbance at 215 nm, 227 nm, 193 nm, and 174 nm. Reported values for the experimental conditions were used as targets for physical model parameters with uncertainties as reported or estimated using typical values [166, 167, 169, 170, 171, 172, 173, 174, 103, 102, 104, 105, 106, 107, 175, 178, 179, 180, 177, 181, 182, 108, 183, 184, 185, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208].

Determinations from Pirraglia et al. [169], Sutherland et al. [173] and Michael et al. [171] were of interest in order to constrain group A reactions across a wider temperature range and determinations from Michael et al. [177] were of interest to constrain D-R1 across a wider pressure range at low temperatures. However, the raw data used for the experimental determinations are not reported in the original papers. To assess the potential influence of including this data as well as provide better model constraints, the present analysis makes use of a procedure to (approximately) regenerate the raw data from experimental rate constant determinations from the information reported. A detailed description of this method can be found in LaGrotta et al. [146] for data generated from a work by Kircher et al. [107]– the same general approach was extended here.

6.3 Results and discussion

Overall, predictions made by the optimized model, which included optimized molecular parameters, rate parameters and physical model parameters are in reasonable consistency with target

data included from targets from *ab initio* calculations, rate constant determinations, macroscopic observables, and reported experimental conditions. Although this study consisted of a variety of target data and active parameters, the authors chose a few vignettes from their results they believe best highlighted the advantages of implementing MSI. Some of these results will explore the disagreement between MSI rate constant determinations and those previously reported in the literature– specifically focusing on discrepancies that can be attributed to coupled reaction systems.

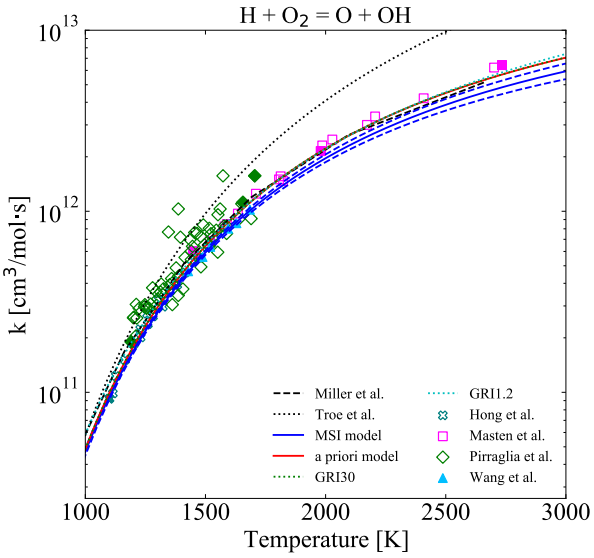


Figure 1: Rate constants for A-R1 at 1 atm. Open symbols represent experimental determinations [166, 167, 169, 170] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from Troe et al., GRI30, GRI1.12 and Miller et al. [212, 213, 214, 215], the *a priori* model, and the MSI model.

The analysis begins with group A reactions, which are fundamental building blocks of many combustion systems. Generally, there is good agreement between k_{A1}, k_{A2}, and k_{A3} and rate constant determinations from the literature as seen in Fig. 1, 3 and 4. However, there is a noticeable difference between Masten et al. [167] rate constant determinations for k_{A1} and the MSI model at higher temperatures. Uncertainty weighted sensitivity analysis of the time profiles for OH and H (Fig. 6) helps reveal the prospective source of this discrepancy. Uncertainty weighted sensitivity coefficients, $\frac{\partial f_i}{\partial x_j}\sigma_j$, indicate the parameters whose prior uncertainties (σ_j) influence uncertainties

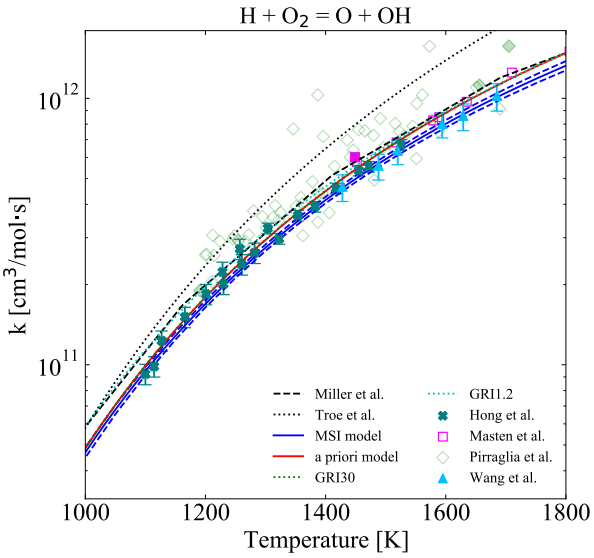


Figure 2: The same data and simulations are presented as Fig. 1, with an adjusted temperature range. Error bars are also included for Hong et al. [166] and Wang et al. [170] data.

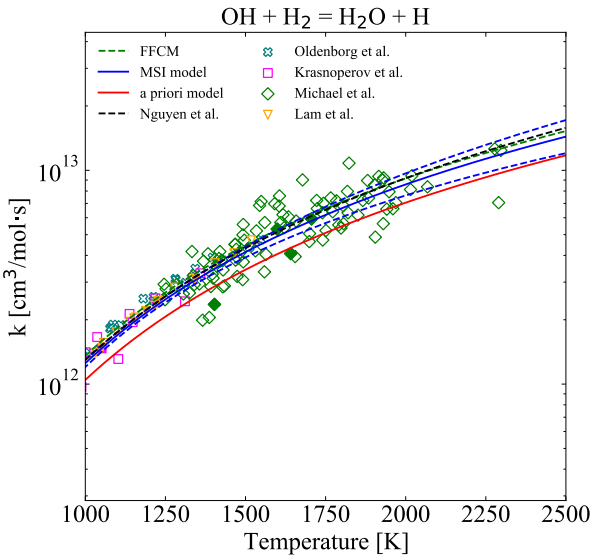


Figure 3: Rate constants for A-R2 at 1 atm. Open symbols represent experimental determinations [172, 171, 174, 151] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from Nguyen et al. and FFCM1 [216, 84], the *a priori* model, and the MSI model.

in the model predictions, and likewise which parameters can be constrained by inclusion of data for those observables. Fig. 6 reveals that as temperature increases (panel c being at the highest temperature), predictions of the observables have strong competing influence of parameters related

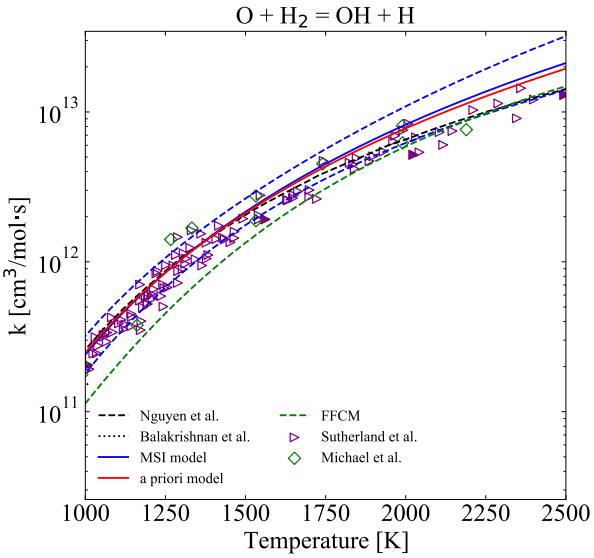


Figure 4: Rate constants for A-R3 at 1 atm. Open symbols represent experimental determinations [173, 171] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from Nguyen et al. and Balakrishnan et al., FFCM [217, 218, 84], the *a priori* model, and the MSI model.

to k_{C1}. These experiments were used to determine k_{A1} however, after including constraints from *ab initio* calculations (Klippenstein unpublished), it can be noted that these higher temperature species profiles from Masten et al. can not exclusively be used to inform k_{A1}. In fact the results of the UWSA imply that the value for k_{C1} must be considered simultaneously. Masten et al. [167] reported that their determinations were fairly independent of k_{C1} and assumed a constant value for k_{C1} across the full experimental temperature range. This value was approximately 1/3 that of the MSI value for k_{C1} at the high temperatures. The MSI model and Masten et al. determinations are able to reproduce the OH and H species profiles in Fig. 5 well within uncertainties. However, given that UWSA reveals there is clearly an influence of secondary chemistry (parameters used to represent k_{C1}) as well as other physical model parameters, it is reasonable to conclude that the MSI interpretation of the data is more accurate given the updated secondary chemistry as well as the inclusion of physical model parameters.

Wang et al. [170] also derived rate constants for k_{A1} using OH time-histories from shock heated mixtures of $C_3H_6O_3/O_2/Ar$. Wang et al. note how only lower $C_3H_6O_3$ concentration mixtures are

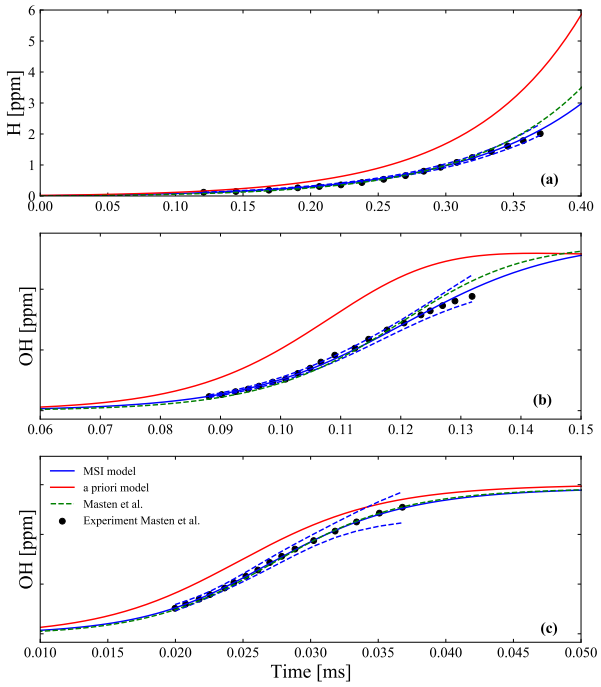


Figure 5: Time profiles of H and OH in shock-heated $O_2/H_2/H/Ar$ near 1700 K 0.794 atm (a), 1980 K 0.675 atm (b) and 2898 K 1.98 atm (c). Symbols represent experimental data from Masten et al. [167], lines represent model predictions using the *a priori* model, MSI model, and the original Masten et al. model [167] with group A reactions derived from experimental data shown. OH profiles are presented in arbitrary units because a multiplier is used to adjust for OH measurement calibration error reported in [167].

able to be used for rate constant determinations for k_{A1}, due to the strong influence of secondary chemistry as $C_3H_6O_3$ concentrations increase. The high sensitivity to secondary chemistry is confirmed with an UWSA as well as traditional sensitivity analysis. Panel (a) of of Fig. 10 and panel (a) of Fig. 9 show the uncertainty-weighted sensitivity analysis and traditional sensitivity analyses to the OH profile in Fig. 8 panel (a). This OH profile was generated using a lower initial $C_3H_6O_3$ concentration and the resulting sensitivity analyses show that although parameters related to k_{A1} are certainly the most sensitive, there is still potentially non-negligible influence from secondary chemistry. Panel (b) of Fig. 10 and panel (b) of Fig. 9 show the uncertainty-weighted sensitivity as well as traditional sensitivity respectively, to the OH profile in Fig. 8 panel (c). This OH (Fig. 8 panel (c)) profile was generated using a higher initial $C_3H_6O_3$ concentration; the resulting sensitivity analysis show the resulting OH profile is sensitive to a variety of parameters related to other

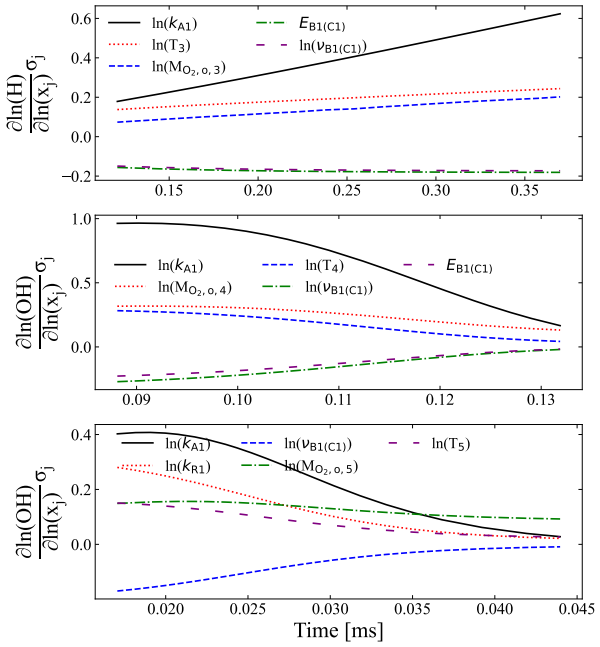


Figure 6: Uncertainty-weighted sensitivity analysis for H and OH of Masten et al. [167] shown in Fig. 5.

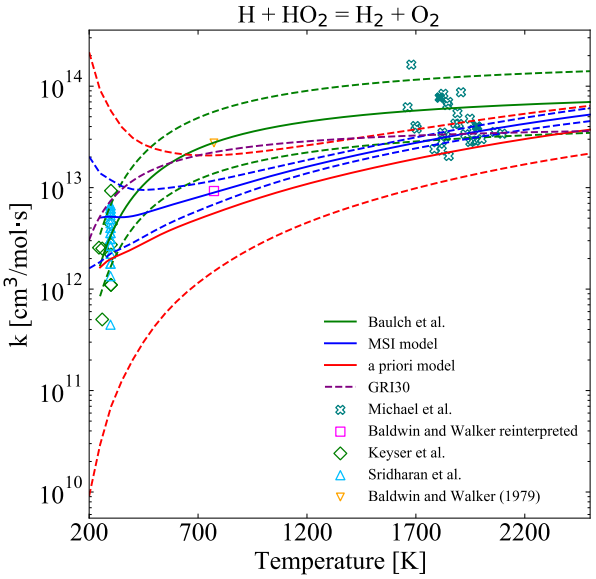


Figure 7: Rate constants for C-R1 at 1 atm. Open symbols represent experimental determinations [153, 154, 175, 219] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from Baulch et al. and Masten et al. [44, 167], the *a priori* model, and the MSI model.

reactions, which is why Wang et al. choose to exclude these profiles when making their k_{A1} rate constant determinations.

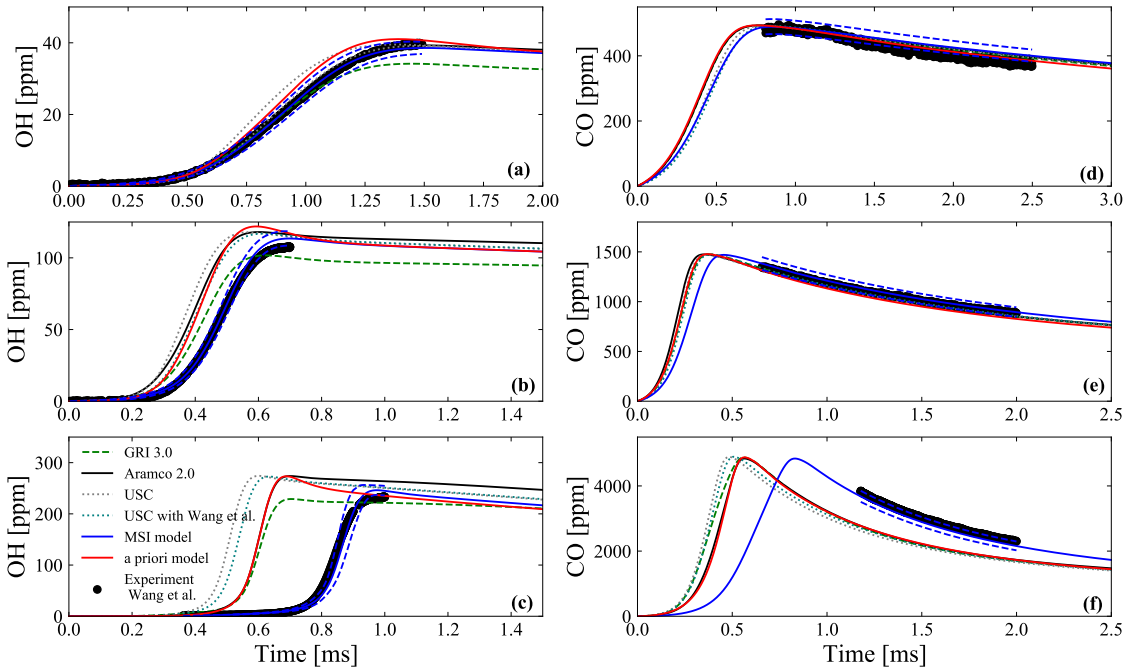


Figure 8: Time profiles of OH and CO in shock-heated $C_3H_6O_3/O_2/H/Ar$ near 1520 K 1.55 atm (a,d), 1524 K 1.55 atm (b,e) and 1364 K 1.66 atm (c,f). Symbols represent experimental data from Masten et al. [170], lines represent model predictions using the *a priori* model, MSI model, and GRI30 model [213], USC model [209], Aramco 2.0 model [220], and USC model with updated Wang et al. rate constants [170]

The MSI model agrees with the Wang et al. k_{A1} determinations within uncertainties (Fig. 1) and unsurprisingly, is able to reproduce OH time histories at the low concentration conditions used to derive these rate constants (Fig. 8 panel (a)). However, closer inspection of Fig. 8 panel (a) reveals the MSI model actually captures the full curvature of the raw data better than that of the Wang et al. model. This is likely due to the influence of secondary chemistry unaccounted for by Wang et al. even at lower concentration conditions.

The MSI model is also able to reproduce the experimentally determined OH profiles generated with higher concentration mixtures of $C_3H_6O_3$ as seen in Fig. 8 panel (b and c). Closer examination of Fig. 8 panel (b and c) reveals that other models (simulated without impurities) ignite far too quickly thus indicating that the discrepancies between these model predictions and experimental determinations can not be attributed to impurities or vibrational relaxation times (which occur on a <10 μs time scale). Rather the discrepancies are due to the rate constant values of secondary

chemistry. The secondary chemistry of note influencing these OH profiles are parameters related to k_{L1}, k_{M1}, and k_{L2} as seen in Fig. 10 (b). Given that these profiles were informative of a variety of other rate constants, which are active parameters in our model structure, MSI is able to appropriately include these species profiles as target data, to help create redundant overlapping reaction systems. Additionally, because MSI is able to correctly classify the information content of raw experimental data sets, the model is able to reproduce the higher concentration OH time histories, Fig. 8 panel (b,c).

Another important piece of information that can be gleamed from group A analysis (A-R1, A-R2 and A-R3) is the importance of these branching reactions, such that even small uncertainties in any of the rate constants can become the largest uncertainties in combustion model predictions. The MSI model determinations for k_{A1}, k_{A2}, and k_{A3} show minimal deviation from the prior model, generally less than 12%. However, these small adjustments in the rate constants contribute to major differences between the *a priori* model and MSI model when making combustion model predictions (Fig. 5, Fig. 8, and Fig. 12). Demonstrating how for these reactions extremely accurate rate constant determinations are necessary to make accurate combustion predictions.

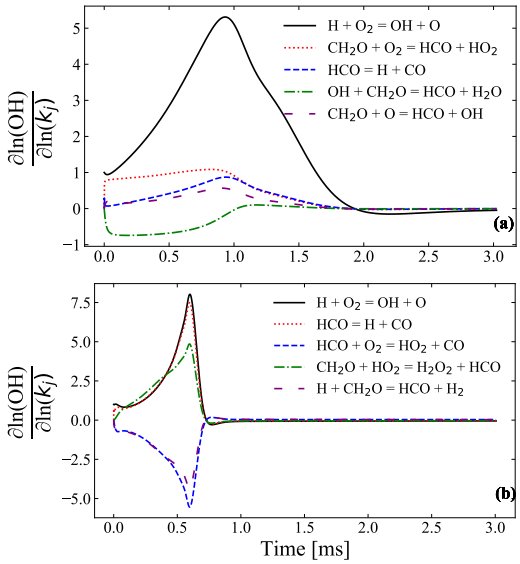


Figure 9: Panel a: Traditional sensitivity analysis for OH of Wang et al. [170] shown in Fig. 8 (a). Panel b: Traditional sensitivity analysis for OH of Wang et al. [170] shown in Fig. 8 (c).

Group C rate constant determinations generally agree fairly well with previous determinations

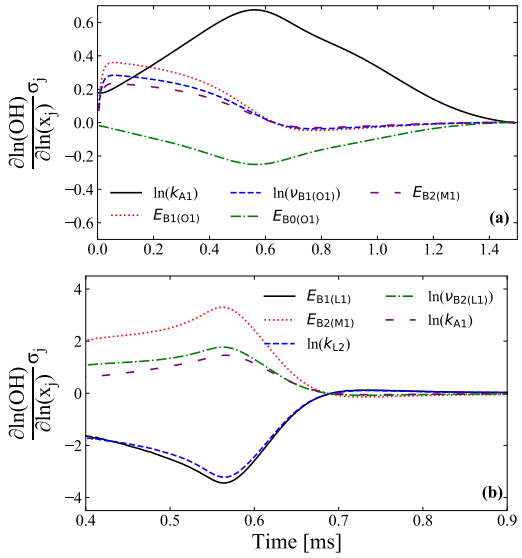


Figure 10: Panel a: Uncertainty-weighted sensitivity analysis for OH of Wang et al. [170] shown in Fig. 8 (a). Panel b: Uncertainty-weighted sensitivity analysis for OH of Wang et al. [170] shown in Fig. 8 (c).

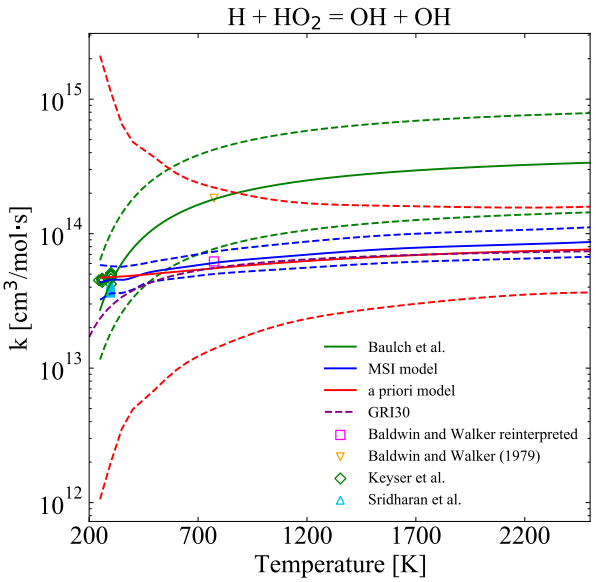


Figure 11: Rate constants for C-R3a at 1 atm. Open symbols represent experimental determinations [153, 154, 219] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from Baulch et al. [44], *a priori* model, and the MSI model.

from literature (Fig. 14 and Fig. 11), with the exception of the Baldwin and Walker determinations [219], which are older and generally considered less reliable. However, some determinations from

Michael et al. [175] fall outside the uncertainties of the prior model as well as the MSI predictions (Fig. 14). Based on the UWSA in Fig. 15, the two O atom time histories used as macroscopic target data were fairly informative of not only parameters representing k_{C1} but also physical model parameters $T_{i,o,e}$ and $M_{\mathrm{H_2},o,e}$. In fact, although the MSI model is able to reproduce the species profiles presented by Michael et al. [175] in Fig. 13, the species profiles in Fig. 13 (b) (corresponding to the filled marker in the Fig. 14 outside of the prior and posterior uncertainties) is only able to be reproduced with an adjustment to the initial temperature. Additionally, although the MSI model and Baulch et al. [44] model (determined with an Arrhenius parameter optimization) both seem to encompass a majority of the experimental determinations, the Baulch et al. [44] model has a temperature dependence inconsistent with theory for k_{C1} and k_{C3a}. Both these results help demonstrate how including *ab initio* calculations and physical model parameters in the model structure (with appropriate uncertainties) protects against unrealistic adjustments to rate constant determinations by capturing the full information content of a data set as well as insuring the model maintains a physically realistic temperature dependence.

The results from the group C rate constants also provide some additional insights into the results seen for group A rate constants. Specifically, recounting the history for the rate constant determinations for C-R3a and C-R1 helps explain the results presented here. Baldwin and Walker's 1979 determinations for C-R3a and C-R1 hinged on rate constant determinations for the reactions of A-R1 and B-R2 (both of which have been revised significantly since 1979) [219]. (Reinterpreted values for the Baldwin and Walker method are also plotted, using the rate constant values determined by MSI). Baulch et al. based their recommendations for C-R3a and C-R1 around Baldwin and Walker's original determination [44]. The GRI mechanism [213] used the Baulch et al. determination as priors for their optimization; the optimization determined lower values for both rate constants but maintained the Baulch et al. temperature dependence (Fig. 14 and Fig 11). Then, Hong et al. [166] used the GRI mechanism as the secondary chemistry when making rate constant determinations for A-R1. Examining Fig. 2 it can be noted that at lower temperatures the MSI model is scraping the edge of the Hong et al. error bars, while at higher temperatures there is much

better agreement.

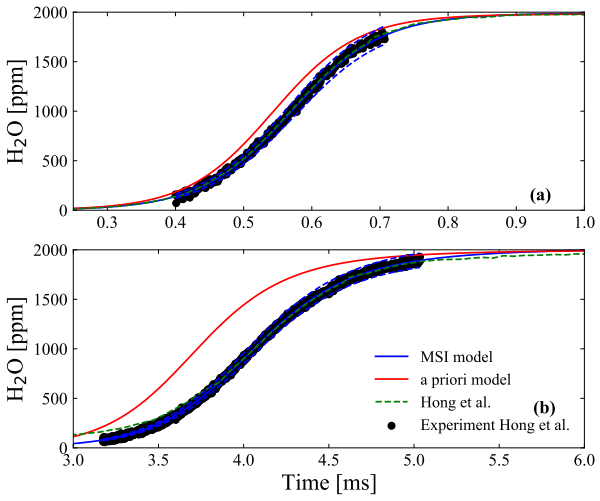


Figure 12: Time profiles of H_2O in shock-heated $O_2/H_2/Ar$ near 1472 K 1.83 atm (a), and 1100 K 1.95 atm (b). Symbols represent experimental data from Hong et al. [166], lines represent model predictions using the *a priori* model, MSI model, and Hong et al. model [166].

The raw data used to determine A-R1 can be seen in Fig. 12, where panel b is at the lowest temperature from the data set (1100 K). In their work Hong et al. [166] present a sensitivity analysis for the maximum slope of the H_2O profile (the quantity from which they derive A-R1). This sensitivity analysis indicates that in the low temperature determinations errors are introduced by interfering reactions, namely C-R3a and C-R1. This explains the discrepancy for A-R1 between the MSI model and the Hong et al. determinations at low temperatures. This decades long propagation of structural uncertainties could have been avoided if rate constant determinations were always made with raw data and the appropriate coupled data sets.

Analysis of group D and E can begin by examining Fig. 16 and noting the discrepancy between the MSI model and the Davidson et al. [178] determinations, despite the ability of the MSI model to predict the OH time history presented by Davidson et al. within uncertainties (Fig. 17). Davidson et al. used shock heated $H_2/O_2/Ar/N_2$ mixtures to generate OH profiles and determine k_{D1}. Sensitivity analyses presented by Davidson et al. and reproduced here (Fig. 18) for the OH time history indicate that the primary rate constant of interest k_{D1} competes strongly with k_{A1}; meaning that in order to make an accurate determination for k_{D1} the determination from k_{A1} must

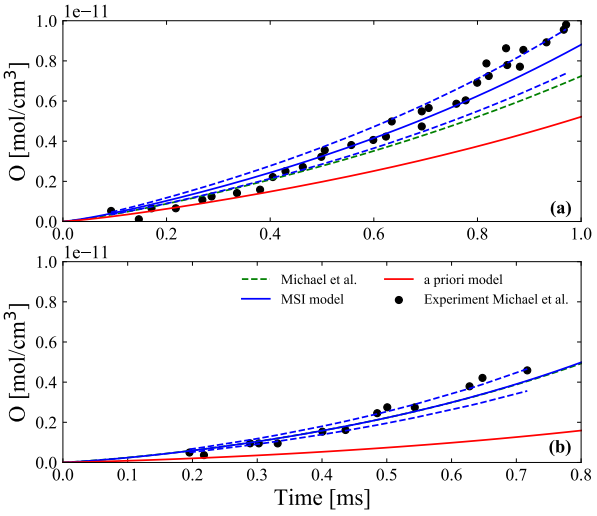


Figure 13: Time profiles of O in shock-heated $H_2/O_2/Ar$ near 1981 K 0.687 atm (a) and 1810 K 0.687 atm (b) [175], lines represent model predictions *a priori* model, MSI model, Michael et al. [175].

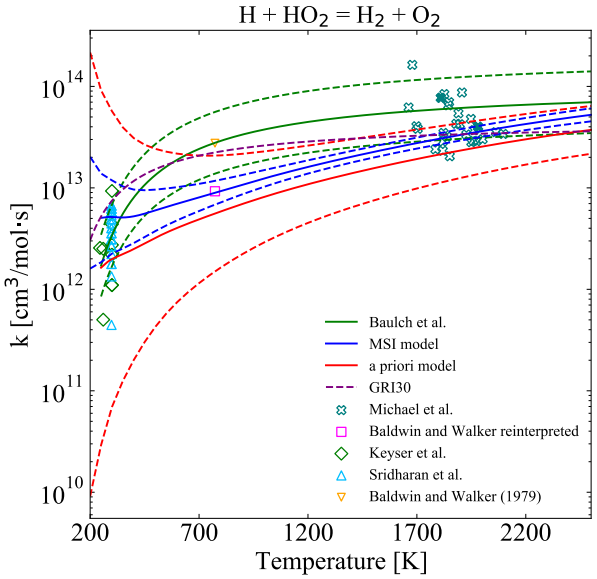


Figure 14: Rate constants for C-R1 at 1 atm. Open symbols represent experimental determinations [153, 154, 175, 219] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from Baulch et al. and Masten et al. [44, 167], the *a priori* model, and the MSI model.

be considered simultaneously. Davidson et al. [178] determined k_{D1} by simulating OH profiles and adjusting k_{D1} to match the peak of the experimental OH profiles– assuming a constant value for k_{A1} consistent with that of the the GRI 1.12 mechanism [214]. The k_{A1} value assumed by

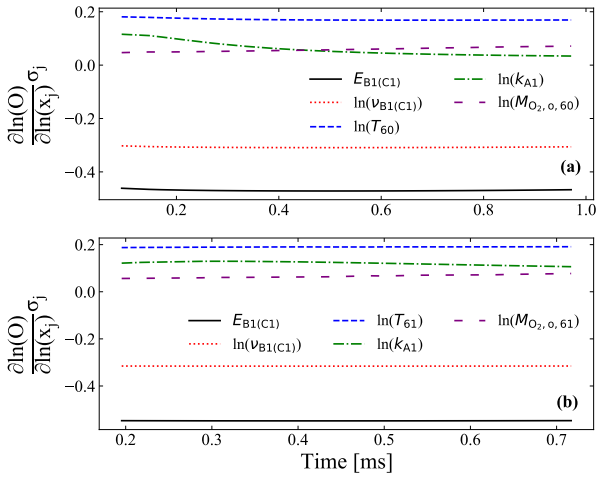


Figure 15: Uncertainty-weighted sensitivity analysis for O of Michael et al. [170] shown in Fig. 13

Davidson et al. is approximately 14% larger than that of the optimized MSI model. This deviation is fairly substantial given the nature of R-A1 and likely explains the discrepancy between the MSI model and k_{D1} values determined by Davidson et al.

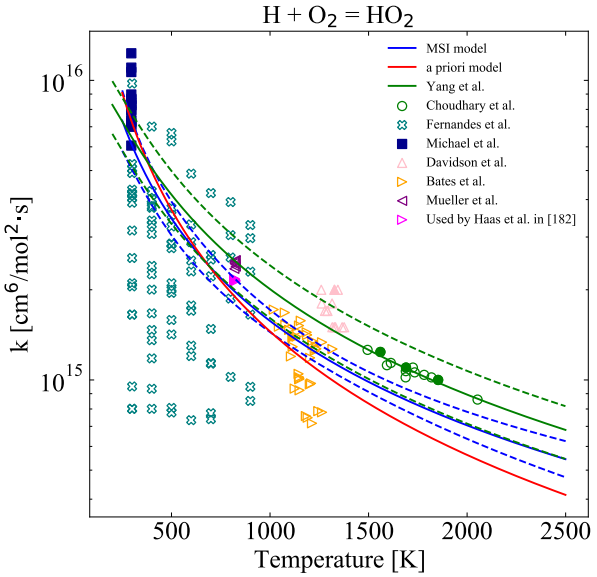


Figure 16: Rate constants for D-R1. Open symbols represent experimental determinations [221, 180, 177, 178, 176, 222] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from *a priori* model, and the MSI model.

Bates et al. [176] also performed a series of experiments to determine the rate constant for k_{D1}

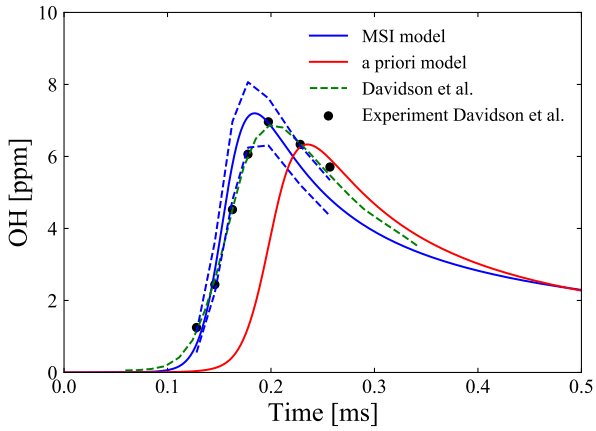


Figure 17: Time profiles of OH, in shock-heated $H_2/O_2/H/Ar$ near 1326 K 64.3 atm. Symbols represent experimental data from Davidson et al. [178], lines represent model predictions using the *a priori* model, MSI model, and Davidson et al. [178].

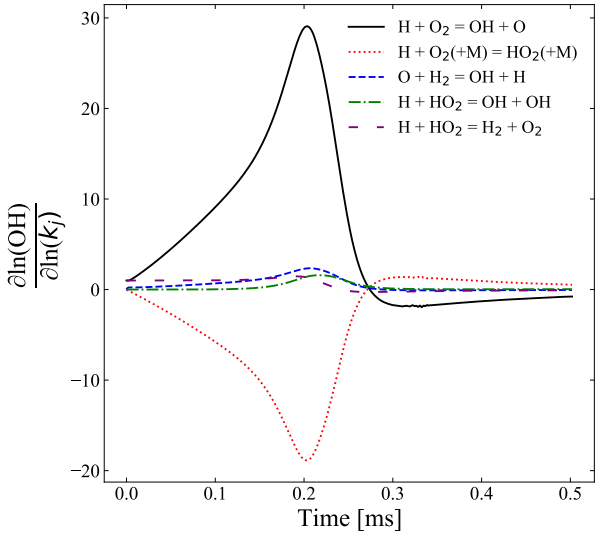


Figure 18: Uncertainty-weighted sensitivity analysis for OH of Davidson et al. [178] shown in Fig. 17.

where a small amount of NO was added to the H_2/O_2 mixture and the rate of k_{D1} was determined by monitoring the absorption of NO_2 profiles. After induction, the major competing reaction with k_{D1} is k_{E1a} and the rate constant for k_{D1} can be determined using the analytical expression $[NO_2]_{\text{plateau}} = k_{D1}[O_2][M]/k_{E1a}$ assuming a reliable value for k_{E1a} is known. Davidson et al. used a value for k_{E1a} from the GRI 1.12 mechanism [214], which is within 11% of the MSI model. The MSI model is able to reproduce the the NO_2 profile presented by Bates et al., and has good

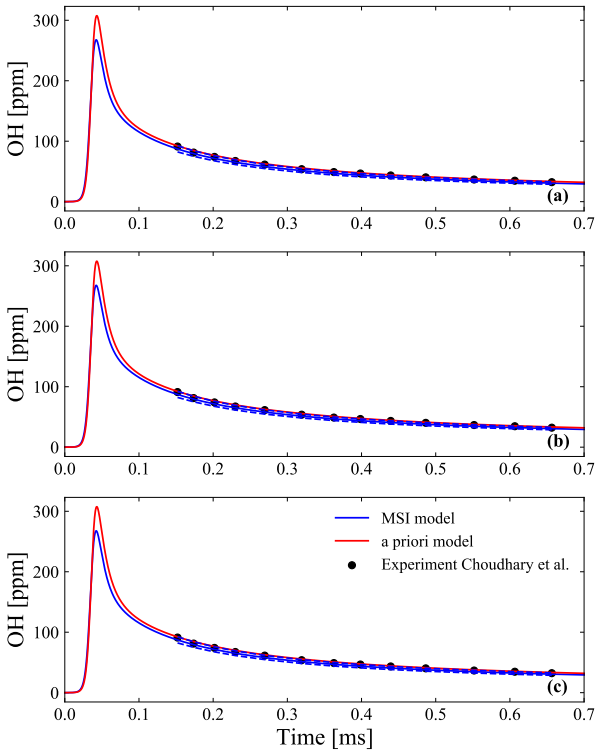


Figure 19: Time profiles of OH H_2/O_2/Ar near 1600 K 12 atm (a), and 1853 K 12.4 atm (b). Time profiles of OH $H_2/O_2/CO_2$/Ar near 1564 K 13.9 atm (c). Symbols represent experimental data from Choudhary et al. [180], lines represent model predictions using the *a priori* model and MSI model.

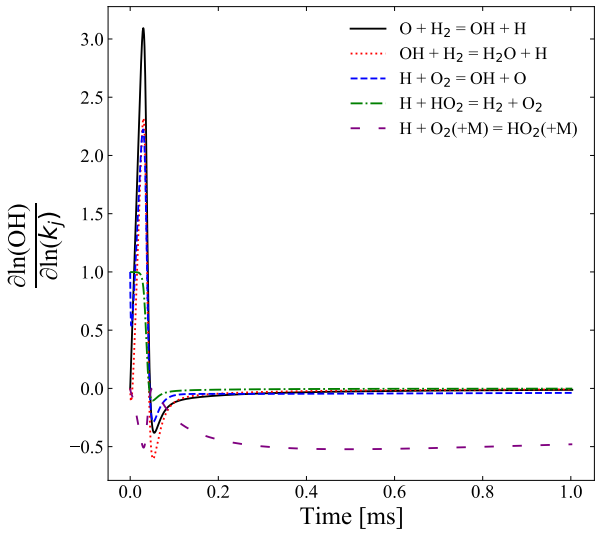


Figure 20: Traditional sensitivity analysis for OH of Choudhary et al. [180] shown in Fig. 19 (a).

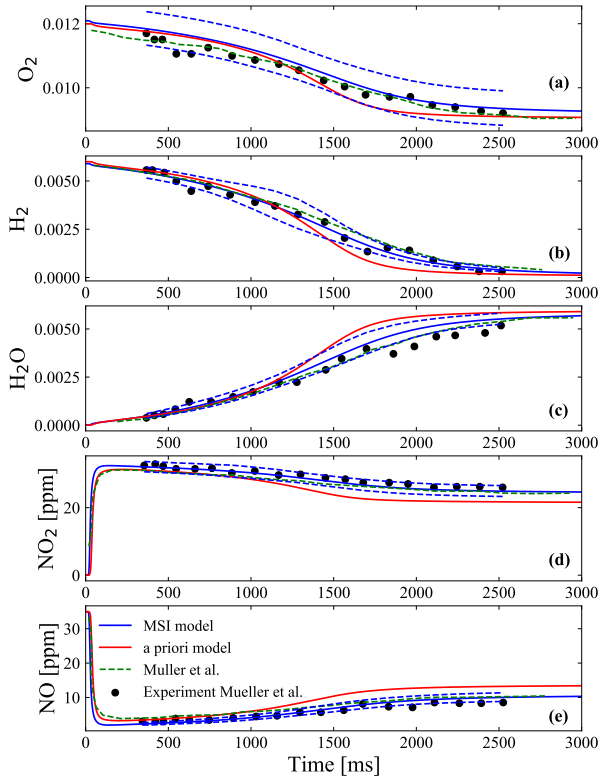


Figure 21: Time profiles of $H_2/O_2/NO/Ar$ near 800 K 10 atm Symbols represent experimental data from Mueller et al. [222], lines represent model predictions using the *a priori* model, MSI model, and Mueller et al. [222].

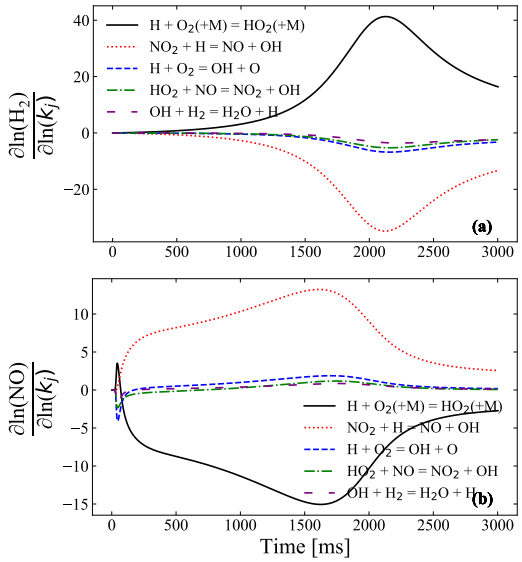


Figure 22: Traditional sensitivity analysis for H_2 and NO of Mueller et al. [222] shown in Fig. 21. NO_2 sensitivities are similar to those of NO but are of opposite sign.

agreement with the k_{D1} value determined by Bates et al. [176] from this profile– likely because the MSI model has good agreement with the k_{E1a}.

The variation in k_{D1} determinations due to secondary chemistry or competing reactions has not gone unnoticed by the combustion community. In fact, a publication by Yang et al. [223] explores the idea of reinterpreting the original results from the Bates et al. [176] and Davidson et al. [178] using updated (and more accurate) rate constants for the secondary chemistry to determine a more accurate value for k_{D1}.

The premise of the Yang et al. [223] study is similar in spirit to that of the the MSI study conducted here– acknowledging the importance of secondary chemistry to accurately untangle reaction systems. However, the MSI approach seeks to actively determine values for entangled rate constants simultaneously instead of relying on stagnant updated chemistry in the *a priori* model. The results of the Yang et al. study, which only consider target data from Davidson et al. and Bates et al., can be seen in Fig. 16 and agree with the results of the MSI model within error bars.

The MSI study also included additional data sets to aid in the determination from k_{D1}, one of which is from from Choudhary et al. [180]. The MSI model agrees with determinations from Choudhary et al. [180] within the stated 9% uncertainties (Fig. 16) and can reproduce the OH time-histories in Fig. 19 used to determine k_{D1}. The slight discrepancy between the MSI determinations and the Choudhary et al. [180] determinations are tied to k_{A3} and k_{A2}. Choudhary et al. [180] explain that k_{A3} and k_{A2} have an affect on OH concentration during induction (this was confirmed here with a sensitivity analysis Fig. 20) and depending on the base mechanism chosen for their analysis, determined values of k_{D1} could vary. In fact, Choudhary explains how using the Aramco 2.0 mechanism [220] for the base mechanism instead of FFCM [84] produced results for k_{D1}, which were 8% lower than those presented in the paper and displayed in Fig. 16. Given the MSI values for k_{A3} and k_{A2} do not agree with those of FFCM [84] (Fig. 3 and Fig. 4), it is unsurprising the final determinations for k_{D1} differ slightly. This is another example of how even small uncertainties in k_{A3} and k_{A2} can have a large effect on combustion model predictions, and in

this case even other rate constant determinations.

Mueller et al. [222] used a flow reactor setup and mixtures of $H_2/O_2/NO/Ar$ to determine values for k_{D1} by introducing NO into the initial experimental mixture the species profiles used to determine k_{D1} are also equally sensitive to k_{E1a} (Fig. 22). Mueller et al. [222] assumed a value for k_{E1a} citing a personal communication along with a publication from Ko et al. [156]. k_{D1} was determined by Mueller et al. by adjusting the rate constant value in simulations to reproduce H_2, NO and NO_2 species time histories determined in the study. Fig. 16 shows good agreement between the MSI model and the value for k_{D1} determined by Mueller et al. [222] and Fig. 21 shows good agreement with the experimental species profiles. Given the undeniable entanglement between k_{D1} and k_{E1a} for these results, as one would expect the MSI model and the k_{E1a} value used by Mueller et al. are also in good agreement.

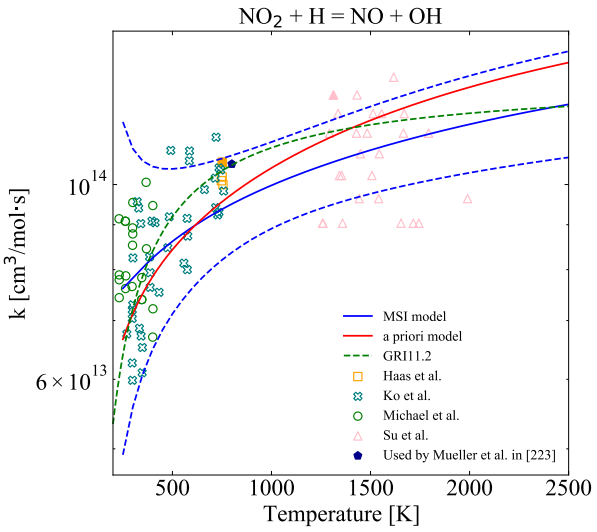


Figure 23: Rate constants for E-R1. Open symbols represent experimental determinations [222, 181, 155, 156, 182] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from *a priori* model, and the MSI model.

Similarly, Haas et al. [181] used flow reactor experiments to determine values for k_{E1a} using mixtures of $H_2/O_2/NO/Ar$, by using the ratio of $[O_2]/[NO_2]$ and assuming a value for k_{D1}. Sensitivity profiles for O_2 and NO_2 are similar to that seen in Fig. 22 where k_{E1a} and k_{D1} are the dominating sensitivities. Haas et al. assumes a value for k_{D1}, which is the average of the ten most

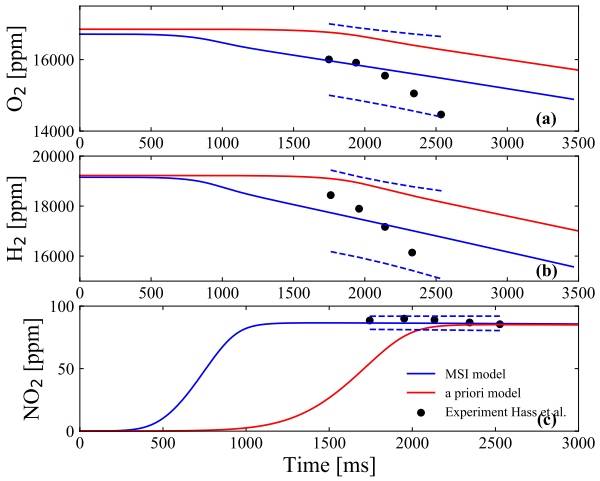


Figure 24: Time profiles of O_2, H_2 and NO_2 in a flow-reactor H_2/O_2/NO/Ar near 752 K 15 atm Symbols represent experimental data from Haas et al. [181], lines represent model predictions using the *a priori* model, MSI model.

common fits used for the reaction– this assumption is in good agreement with the MSI model (see Fig. 16). Given the agreement between the MSI model and k_{D1}, and the ability of the MSI model to reproduce the species profiles in Fig. 24, it is unsurprising there is good agreement between the Haas et al. determinations for k_{E1a} and the MSI model.

Analysis of reaction groups D and E demonstrate the advantages of implementing MSI for determining rate constants in coupled reaction systems. Exemplifying how often times experimental data can not be analyzed without considering a two (or possibly more) reaction system and how assuming a value for one or more reactions in an intertwined system could systematically skew final determinations. Analysis of group D and E reactions also reinforces the importance of using raw experimental data sets when possible in the model structure rather than relying on derived data (such as rate constant determinations), which could systematically be skewed by assumptions made during their determinations.

Similar to the group A rate constants, even small uncertainties in the rate constant for D-R1 can propagate to large uncertainties in combustion model predictions. Relatively minor adjustments in k_{D1} in the combustion temperature region can have a fairly large impact on combustion model predictions (Fig. 17 and Fig. 21).

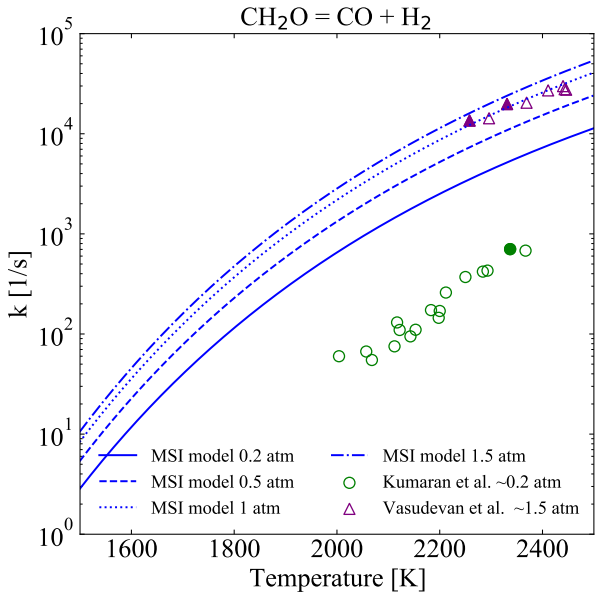


Figure 25: Rate constants for K-R1a. Open symbols represent experimental determinations [193, 194] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from *a priori* model, and the MSI model.

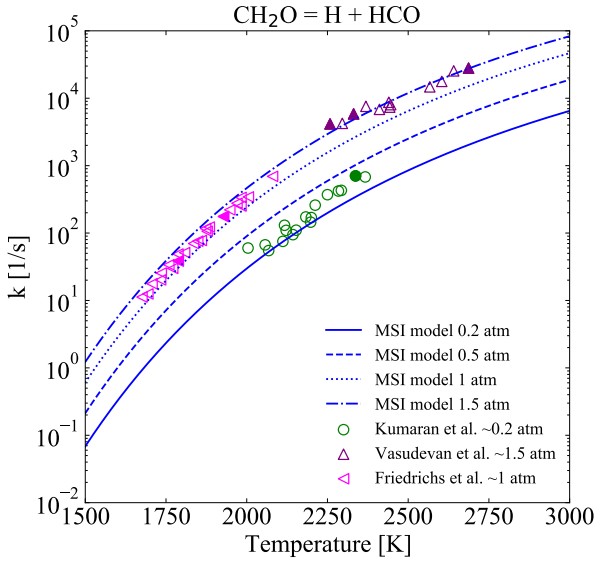


Figure 26: Rate constants for K-R1c. Open symbols represent experimental determinations [193, 194, 191, 192] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from *a priori* model, and the MSI model.

Group K analysis continues to affirm some of the previous trends. Examining the rate constant determinations for k_{K1a} and k_{K1c} in Fig. 25 and Fig. 26 it can be noted the MSI determinations

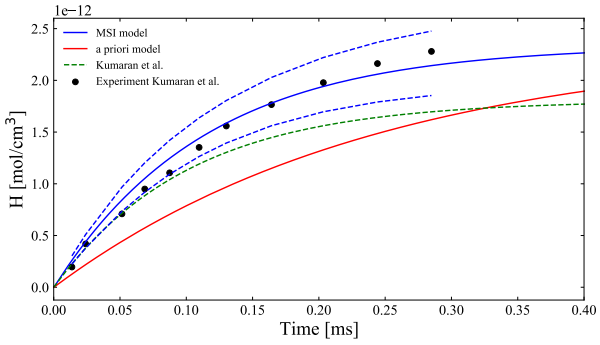


Figure 27: Time profiles of H in a shock-heated CH_2O/Ar near 2337 K 0.329 atm Symbols represent experimental data from Kumaran et al. [193], lines represent model predictions using [193] et al., *a priori* model, and MSI model.

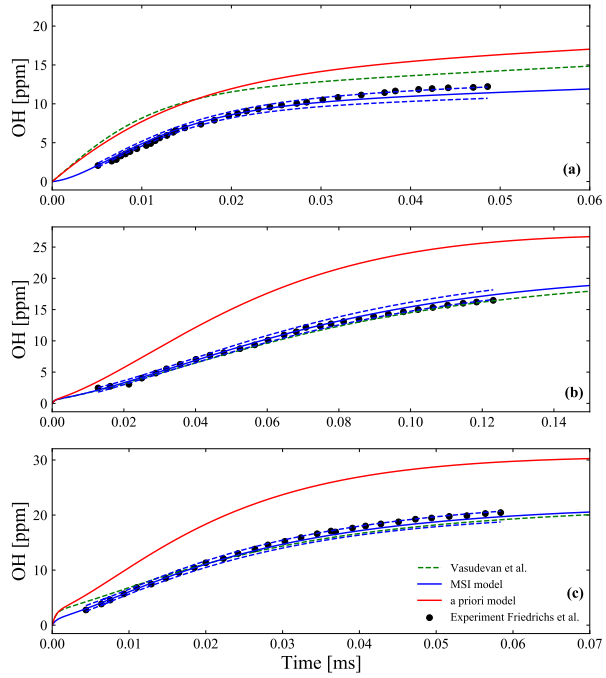


Figure 28: Time profiles of OH in a shock-heated CH_2O/O_2/Ar near 2687 K 1.52 atm and 2068 K 1.26 atm and 2331 K and 1.16 atm. Symbols represent experimental data from Vasudevan et al. [194], lines represent model predictions using Vasudevan et al. [194], *a priori* model, and MSI model.

are larger than the determinations from Kumaran et al. [193], despite having good agreement with H species profile (Fig. 27) Kumaran et al. [193] used to derive the rate constant determinations. A sensitivity analysis of the the H profile reveals it is predominately sensitive to k_{K1a} and k_{K1c} and unsurprisingly, an uncertainty weighted sensitivity analyses reveals the profile is only sensitive to

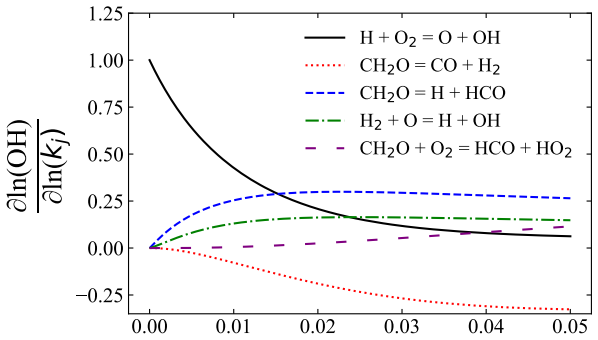


Figure 29: Traditional sensitivity analysis for OH of Vasudevan et al. [194] shown in Fig. 28 (c).

parameters that inform k_{K1}. Deriving accurate values for k_{K1a} and k_{K1c} using the H time history requires untangling the relationship between the two rate constants.

Kumaran et al. [193] determines k_{K1a} and k_{K1c} by varying the rate constants until a good fit of the experimental data is obtained, however no additional constraints were considered when determining the rate constants. The lack of additional constraints (such as a branching ratio) means the values of k_{K1a} and k_{K1c} are free to take any value relative to each-other. The final branching ratio between k_{K1a} and k_{K1c} determined by Kumaran et al. does not agree with the final branching ratio determined by the MSI model; additionally, the rate constants presented by Kumaran et al. [193] fail to reproduce the experimental data as accurately as MSI model (Fig. 27). In the MSI model k_{K1a} and k_{K1c} are treated by the same set of theoretical molecular parameters because they physically exist on the same PES. The advantage of the MSI implementation is that throughout the optimization a physically meaningful branching ratio between the two rate constants is maintained. The values of k_{K1a} and k_{K1c} (as well as the other rate constants on the PES) can only move relative to each-other in a physically meaningful way defined by the *ab initio* calculations used as the *a priori* determinations.

The most interesting result from our group L reactions pertains to k_{L2} and can be seen in Fig. 30. Despite having good agreement with the species profiles present by Fabherber et al. [196] (Fig. 31), there is significant disagreement with the final determined rate constant for k_{RL2}. In their work, Fabherber uses a shock-heated glyoxal decomposition and the resulting HCO profiles to determine the rate constant for k_{L2}. However, sensitivity analysis as well as an UWSA reveal

the HCO profiles measured by Fabherber are not all that sensitive to the reaction of interest (32), thus it is not surprising the MSI rate constant determined seems to be completely agnostic of the Fabherber profiles. This result really highlights the advantage of including raw data into the model structure so the true information content of a data set can be extracted. However, this result also indicates the inherent uncertainty in this rate constant determination. The scatter in the data seen in Fig. 30 along with the sensitivity analysis in Fig. 32 indicate that this reaction would be a good candidate for future studies as well as a good candidate to be implemented with theoretical molecular parameters.

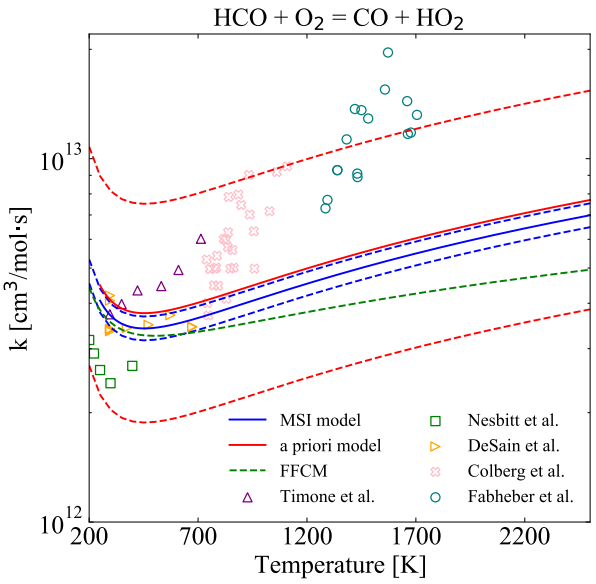


Figure 30: Rate constants for L-R2. Open symbols represent experimental determinations [160, 161, 162, 196, 224] as indicated in the legend. Solid symbols represent experimental conditions for which raw data was included in the model. Lines represent model data from FFCM [84], *a priori* model, and the MSI model.

6.4 Concluding Remarks: Coupled Reaction Systems

Previous theoretical and experimental data for coupled reaction systems were analyzed using the MultiScale Informatics approach, which optimizes values and quantifies uncertainties for a set of molecular parameters (within theoretical kinetics calculations), rate parameters, and physical model parameters (within simulations of macroscopic observables) based on data from various

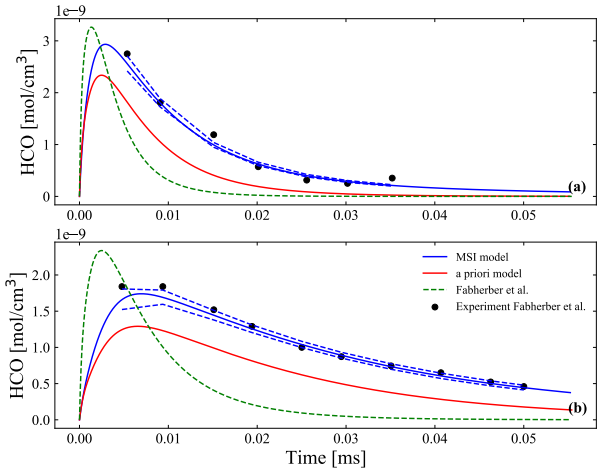


Figure 31: Time profiles of HCO in a shock-heated OCHCHO/Ar near 1618 K 1.39 atm and 1519 K 1.28 atm. Symbols represent experimental data from Fabherber et al. [196]. Lines represent model data from the *a priori* model, and MSI model.

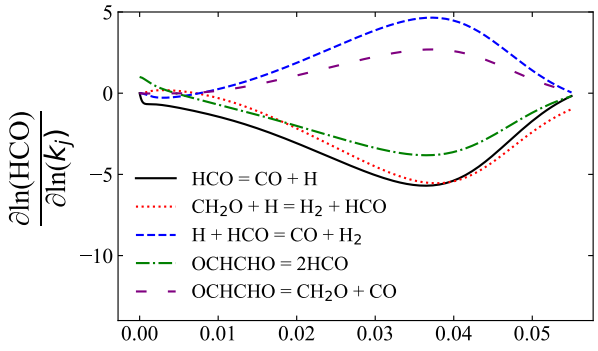


Figure 32: Traditional sensitivity analysis for OH of Vasudevan et al. [196] shown in Fig. 31 (b).

sources and scales. Results and discussion reinforce key elements of developing a high-accuracy kinetic database, including: the importance of raw data for quantifying the information content of experimental data, and the utility of theoretical kinetics calculations for constraining experimental interpretations and providing an independent data source and finally, the subtleties of target data selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties. Results and discussion also places emphasises on the importance of considering all rate constants in coupled reaction systems when deriving rate constant values and the systematic errors that can occur when secondary chemistry is inaccurate. The vignettes in the results presented here highlight the key elements of the approach and the clear advantages of the MSI method for

untangling coupled reaction systems and deriving accurate rate constant values.

A general theme of the present work (consistent with similar analyses [69, 70, 71, 72] for other reactions) is that apparent inconsistencies among theoretically calculated and experimentally derived rate constants can often be resolved when the constraints imposed by *ab initio* electronic structure calculations and measured macroscopic observables are considered simultaneously within multiscale, physics-based frameworks for data analysis. We therefore recommend that future experimental studies report raw data (e.g. measured time profiles) and future theoretical studies provide theoretical kinetics input files (e.g. for Variflex or MESS) in addition to derived rate constants—in order to enable rigorous comparisons of theory and experiment in a manner that better reflects the information content and associated uncertainties in each.

Chapter 7: Model Validation

7.1 Introduction

The model derived from results in Chapter 6 is tested against a wide variety of combustion targets to assess its performance. The validation set consists of targets used for validation by Burke et al. [111] and Keromnes et al. [225]. If the data had already been used in the target set for the model, it was not included in validation. The validation data consists of ignition delay times, species profiles from flow reactors, and laminar flame speeds. The MSI modules built on top of the Cantera [83] framework are used to simulate experimental conditions.

Accurate predictions for the validation targets require both accurate chemical kinetics parameters and accurate treatments of chemical physics in realistic environments. The work presented here addresses the determination of accurate chemical kinetics parameters by leveraging the information content of experimental data sets; these data sets were specifically selected because the conditions under which they were performed were experiments in which chemical physics is well defined and reliable.

More holistically accurate combustion modeling needs to be done in conjunction with better treatments of chemical physics of realistic environments that have large fractions of energy-transferring and reactive colliders [73, 77, 75, 45, 74, 79]. Both of which are active areas of research and are on the precipice of resolution. Thus, given the current state of combustion simulation predictions, for certain experiments, predictions will not be 100% accurate until these challenges are addressed. All told, the MSI model presented here still does better than the *a priori* model for the validation data chosen.

7.2 Validation Plots

Overall, the MSI model performs well on the validation data, with a few exceptions that will be discussed. The ignition delay validation tests are shown in Fig. 1 - Fig. 21. The ignition delay conditions spanned a temperature range from 1000 - 3000 K and a pressure range from 1 - 64 atm. The current mechanism is in good agreement with data from Herzler et al. [226] for 1 bar and 16 bar experiments (Fig. 1 and Fig. 3). However, it can be noted the model has trouble reproducing the Herzler et al. data determined at 4 bar between 950 K and 1100 K (Fig. 2). At these intermediate temperature conditions, and under certain experiment conditions, the data show an unusual behavior. This behavior is created by the competition between the chain-branching reaction H + O_2 = O + H and the chain propagating reaction H + O_2 (+M) = HO_2 (+M). At low temperatures, the reactivity is controlled by the chain propagating reaction, and at high temperatures, the reactivity is controlled by the chain branching reaction. With increasing pressure, the rate of collisions increase, which results in the predominance of H + O_2 (+M) = HO_2 (+M) over H + O_2 = O + H. Being relatively inactive, the HO_2 radicals produced contribute to an overall decrease in reactivity. Thus, in this intermediate temperature range, low pressure experiments show a stronger reactivity than high pressure experiment resulting in this unusual cross-over behavior. The MSI model seems to have trouble reproducing this specific behavior. Generally though, outside of this exception, the MSI model has improved performance over that of the *a priori* model and is able to reproduce the ignition delay validation data. Further examples of the model having trouble reproducing this specefic behavior can be noted in Fig. 19 and Fig. 21.

The flame speed validation tests are shown in Fig. 22 - Fig. 68. The MSI model also performs well on the experimental data used for flame speed validation. The laminar flame speed conditions spanned a temperature range from 298-325 K and a pressure range from 1 - 25 atm. Laminar flame speed experimental data spanned a range of conditions. These conditions ranged from lean (ϕ = 0.4) to rich mixtures (ϕ = 2.5) and contained various diluents Ar/He/N_2. Flame speed predictions for the MSI model generally performed well at room temperature on lean H_2/O_2 mixtures with

various diluents. The MSI model is able to capture the different effects of the diluent at these lower pressure conditions. The model also performs well at 1 atm conditions when CO and CO_2 are added to the initial mixture composition. The MSI model generally shows improvement over the *a priori model*.

However, it can be noted that the model has difficulty reproducing rich mixtures at higher pressures (Fig. 39, Fig. 42, Fig. 43, Fig. 44, and Fig. 45). Specifically, the model has the most difficulty with pure H_2/O_2/Ar mixtures where no CO is added. The source of this discrepancy is likely due to a combination of two factors. The first likely source of discrepancy is the rate constant of H + H + M = H_2 + M. The rate constant for this reaction becomes significantly more important for accurate predictions of flames at rich conditions. The MSI model does not specifically focus on improving the kinetics parameters for this reaction– revisiting this reaction and optimizing its kinetics parameters may be necessary to accurately simulate flames at rich conditions.

The second likely source of discrepancy is the effects of mixture rules and collider efficiencies. Mixture rules help account for how a rate constant (and thus an overall reaction rate) would be affected by the kinetic energy transfer occurring for molecular collisions in the mixture. The ability of a specific molecule to effectively transfer energy is accounted for by assigning it a collision efficiency. The collision efficiency for each molecule present in the mixture is multiplied by that specific molecule's concentration and they are then linearly summed. This linear summation is then multiplied into the expression for calculating the reaction rate. As shown by Lei et al. [77] this assumption of linearity often breaks down at different conditions (specifically as pressure increases). Additionally, the values of collision efficiencies (which are temperature dependent) assigned to specific molecules have an impact on the calculation of the overall reaction rate. These simulations were run using Cantera, which only has the ability to implement a linear mixture rule and collider efficiencies at a single temperature. In this model, collider efficiencies were assigned for $\sim$1000 K. The deviations of these high pressure rich mixture condition predictions from the experimental data are likely due to some combination of these factors. However, further work needs to be done to explore the root of these deviations.

Species profile validation tests are shown in Fig. 69 - Fig. 104. The model also performs very well for species profiles generated in flow reactors. The species profiles conditions spanned a temperature range from 800-1068 K and 0.5-14 atm. H_2, O_2, H_2O, CO and CO_2 profiles were generated from various dilute $H_2/O_2/N_2$ or $H_2/O_2/H_2O/CO/N_2$ mixtures. The MSI model performs well on all $H_2/O_2/N_2$ mixtures and shows improvement over the *a priori* model. The MSI model seems to have more difficulty reproducing species profiles for initial mixtures containing CO. It is likely that this discrepancy stems from the rate constant determined for CO + OH = CO_2 + H. Unfortunately, the rate constant for CO + OH = CO_2 + H was not able to be properly addressed in the current model due to the lack of theory and/or experimental data within the MSI analysis.

As a general note, for flow reactor studies, it is necessary to shift the time in order to take into account any reactivity occurring during the initial mixing of the reactants, where non-ideal conditions are encountered. The MSI model and the *a priori* model have both been time shifted appropriately to account for these effects.

Overall, the MSI model performs well on the the initial validation data. However, until more holistically accurate combustion modeling can be achieved, the model predictions may not perfectly reproduce experimental data. This being said, there are still some potential discrepancies with experimental data, which need to be explored. First, how the branching ratio for the reactions of H + O_2 (+M) = HO_2 (+M) and H + O_2 = O + H is affecting ignition delay time predictions. Second, the effect of the rate constant for H + H + M = H_2 + M, mixture rules, and collider efficiencies are having on rich flame speed predictions. Finally, the effect the rate constant of CO + OH = CO_2 + H is having on CO profile predictions.

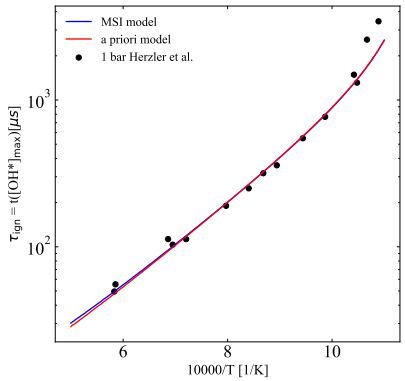


Figure 1: Comparison with ignition delay times measured by Herzler et al. [226] (5.87% H_2 + 2.95 O_2 + 91.18 Ar).

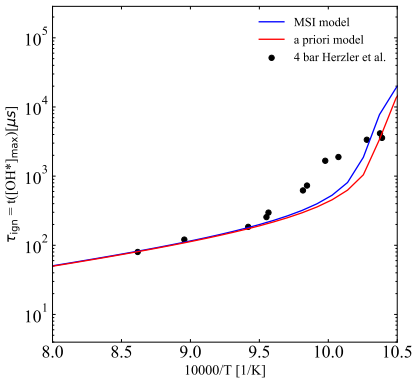


Figure 2: Comparison with ignition delay times measured by Herzler et al. [226] (5.87% H_2 + 2.95 O_2 + 91.18 Ar).

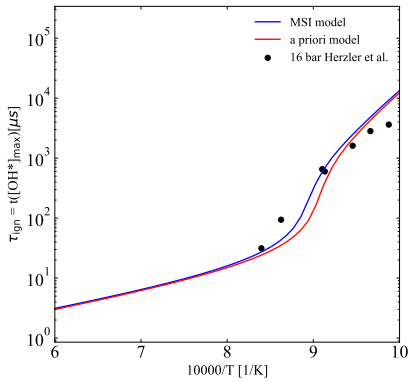


Figure 3: Comparison with ignition delay times measured by Herzler et al. [226] (5.87% H_2 + 2.95 O_2 + 91.18 Ar).

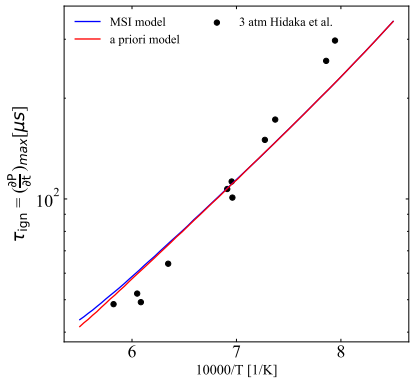


Figure 4: Comparison with ignition delay times measured by Hidaka et al. [227] (1.0% H_2 + 1.0% O_2 + 98.0% Ar).

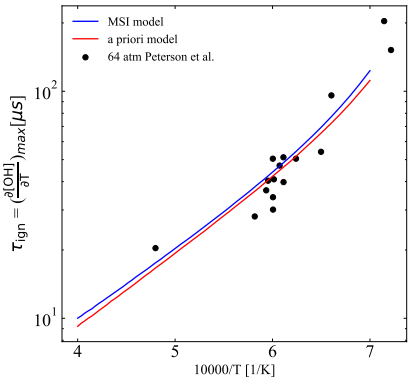


Figure 5: Comparison with ignition delay times measured by Petersen et al. [228] (2.0% H_2 + 1.0% O_2 + 97.0% Ar).

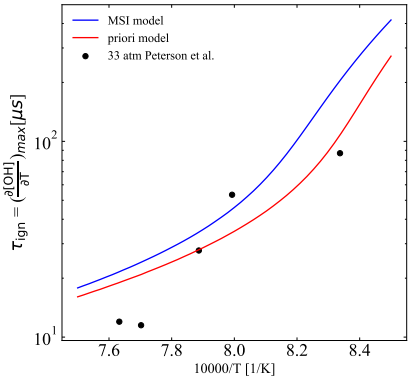


Figure 6: Comparison with ignition delay times measured by Petersen et al. [228] (2.0% H_2 + 1.0% O_2 + 97.0% Ar).

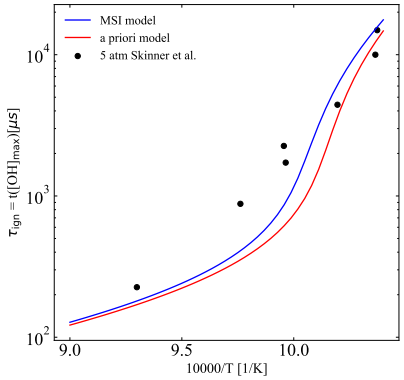


Figure 7: Comparison with ignition delay times measured by Skinner et al. [229] (8.0% H_2 + 3.0% O_2 + 90.0% Ar).

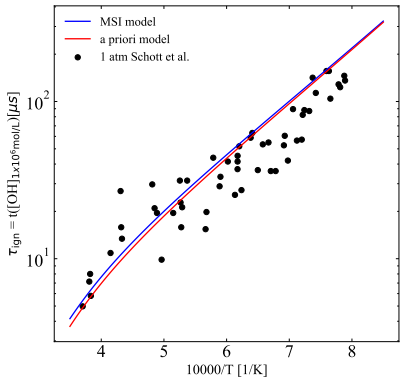


Figure 8: Comparison with ignition delay times measured by Schott et al. [230] (4.0% H_2 + 2.0% O_2 + 94.0% Ar).

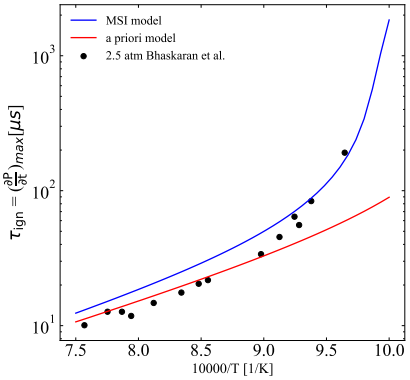


Figure 9: Comparison with ignition delay times measured by Bhaskaran et al. [231] (22.59% H_2 + 14.79% O_2 + 63.41% N_2).

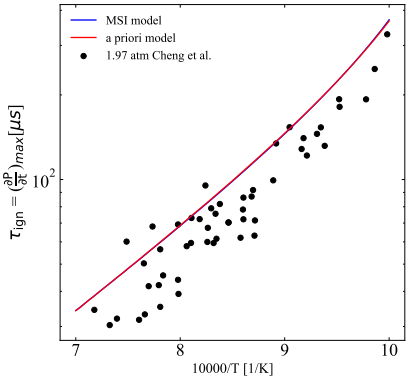


Figure 10: Comparison with ignition delay times measured by Cheng et al. [232] (5.0% H_2 + 5.0% O_2 + 90.00% N_2).

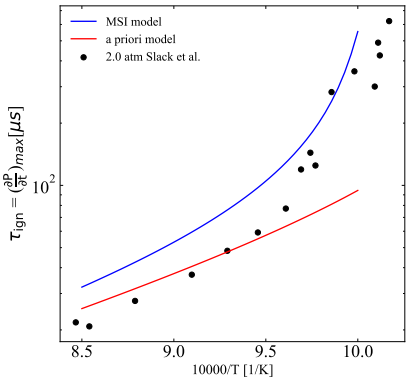


Figure 11: Comparison with ignition delay times measured Slack et al. [233] (22.59% H_2 + 14.79% O_2 + 63.41% N_2).

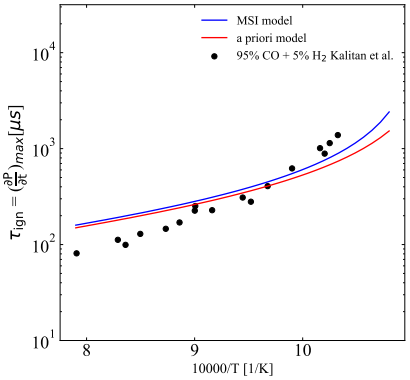


Figure 12: Comparison with ignition delay times measured by Kalitan et al. [234] at 1.1 atm (17.3%(αH_2 + (1-α)CO) + 17.5% O_2 + 65.2% Ar).

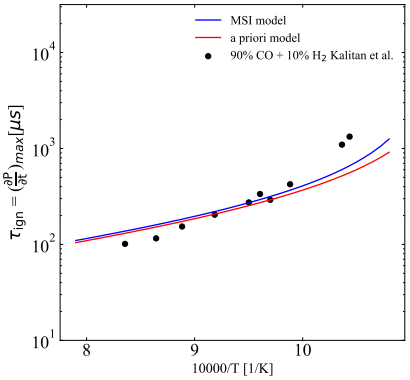


Figure 13: Comparison with ignition delay times measured by Kalitan et al. [234] at 1.1 atm (17.3%(αH_2 + (1-α)CO) + 17.5% O_2 + 65.2% Ar).

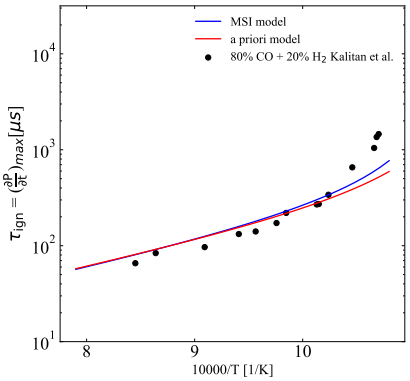


Figure 14: Comparison with ignition delay times measured by Kalitan et al. [234] at 1.1 atm (17.3%(αH_2 + (1-α)CO) + 17.5% O_2 + 65.2% Ar).

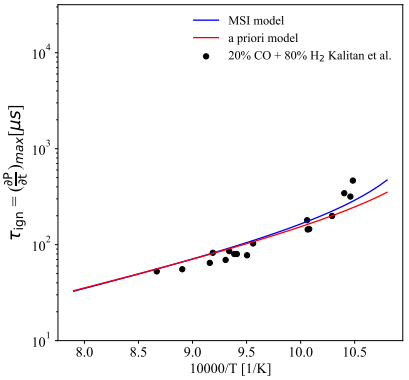


Figure 15: Comparison with ignition delay times measured by Kalitan et al. [234] at 1.1 atm (17.3%(αH_2 + (1-α)CO) + 17.5% O_2 + 65.2% Ar).

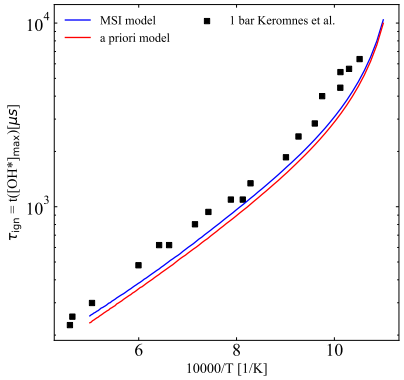


Figure 16: Comparison with ignition delay times measured by Keromnes et al. [225] at 1 bar (0.3% H_2 + 2.96% O_2 + 5.62% CO + 91.12% Ar).

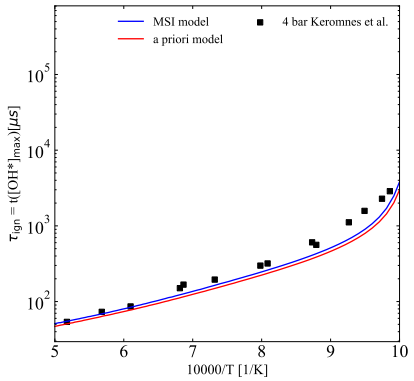


Figure 17: Comparison with ignition delay times measured by Keromnes et al. [225] at 4 bar (0.3% H_2 + 2.96% O_2 + 5.62% CO + 91.12% Ar).

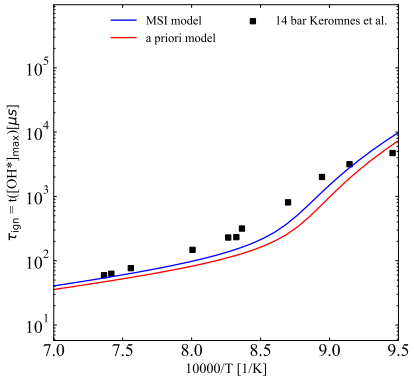


Figure 18: Comparison with ignition delay times measured by Keromnes et al. [225] at 16 bar (0.3% H_2 + 2.96% O_2 + 5.62% CO + 91.12% Ar).

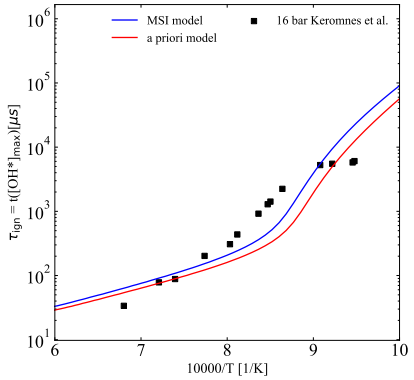


Figure 19: Comparison with ignition delay times measured by Keromnes et al. [225] at 16 bar (0.17% H_2 + 3.47% O_2 + 3.30% CO + 93.06% Ar).

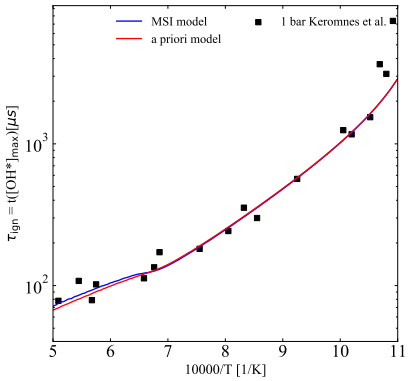


Figure 20: Comparison with ignition delay times measured by Keromnes et al. [225] at 16 bar (2.96% H_2 + 2.96% O_2 + 2.96% CO + 92.12% Ar).

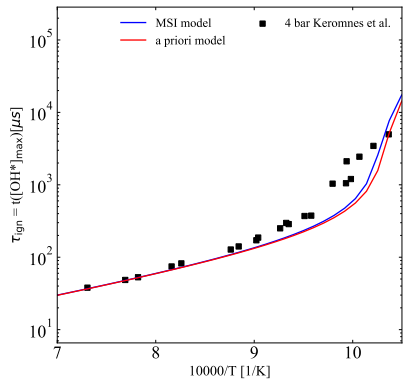


Figure 21: Comparison with ignition delay times measured by Keromnes et al. [225] at 4 bar (2.96% H_2 + 2.96% O_2 + 2.96% CO + 92.12% Ar).

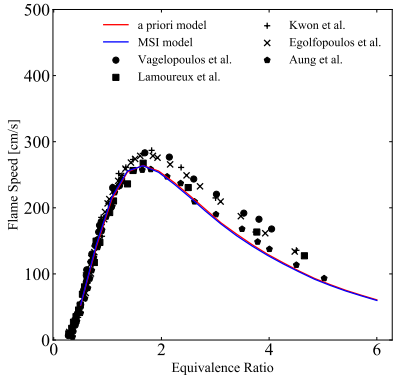


Figure 22: Laminar flame speed of $H_2/O_2/N_2$ mixture at 1 atm and room temperature. Where N_2/O_2 = 3.76 Experimental data are taken from [235, 236, 237, 238, 239].

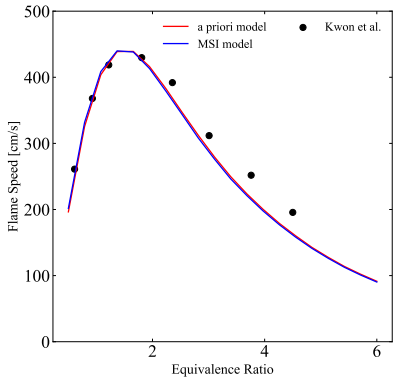


Figure 23: Laminar flame speed of $H_2/O_2/Ar$ mixture at 1 atm and room temperature. Where $Ar/O_2 = 3.76$. Experimental data are taken from [237].

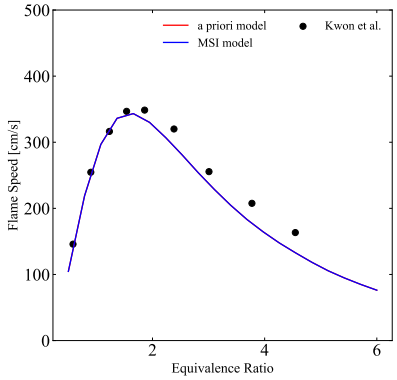


Figure 24: Laminar flame speed of $H_2/O_2/He$ mixture at 1 atm and room temperature. Where $He/O_2 = 3.76$. Experimental data are taken from [237].

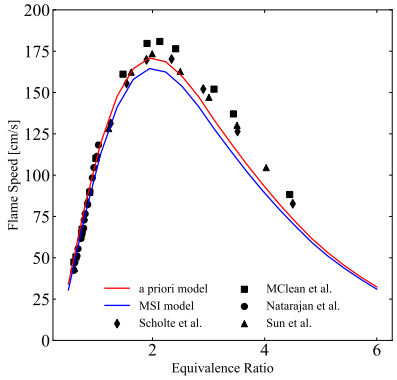


Figure 25: Laminar flame speed of $H_2/O_2/CO/N_2$ mixture at 1 atm and room temperature. Where $H_2 = 50\%$, $CO = 50\%$, and $N_2/O_2 = 3.76$. Experimental data are taken from [240, 241, 242, 243].

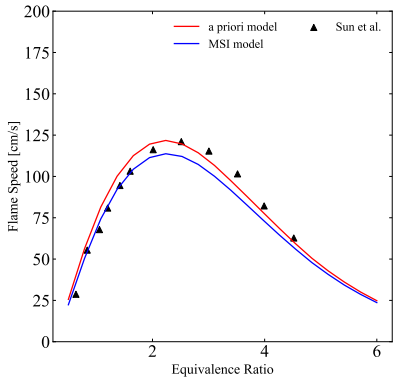


Figure 26: Laminar flame speed of $H_2/O_2/CO/N_2$ mixture at 1 atm and room temperature. Where H_2 = 25%, CO = 75% and N_2/O_2 = 3.76. Experimental data are taken from [243].

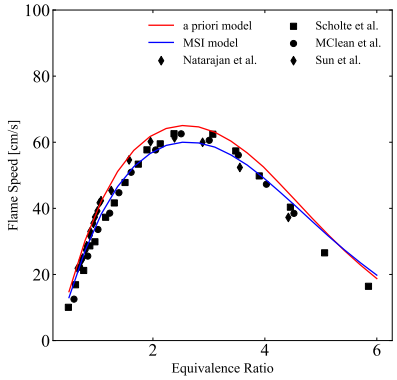


Figure 27: Laminar flame speed of $H_2/O_2/CO/N_2$ mixture at 1 atm and room temperature. Where H_2 = 5%, CO = 95%, and N_2/O_2 = 3.76. Experimental data are taken from [240, 241, 242, 243].

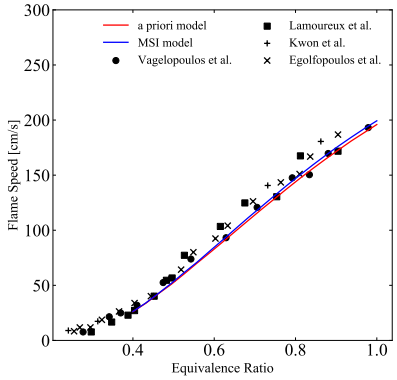


Figure 28: Laminar flame speed of $H_2/O_2/N_2$ mixture at 1 atm and room temperature. Where N_2/O_2 = 3.76. Experimental data are taken from [236, 237, 238, 235].

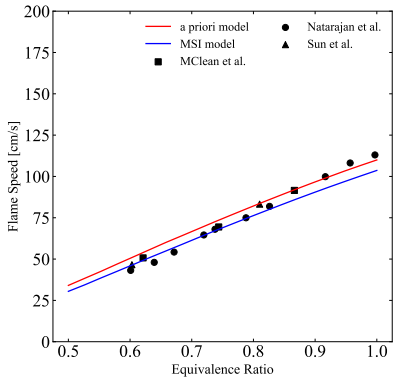


Figure 29: Laminar flame speed of $H_2/O_2/CO/N_2$ mixture at 1 atm and room temperature. Where H_2 = 50%, CO = 50% and N_2/O_2 = 3.76. Experimental data are taken from [240, 242, 243].

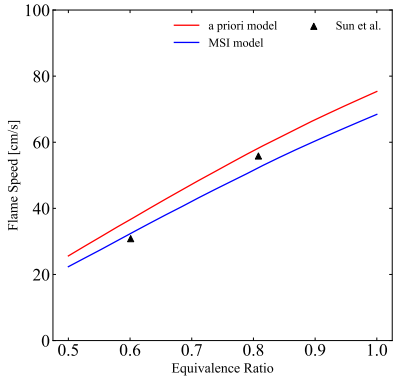


Figure 30: Laminar flame speed of $H_2/O_2/CO/N_2$ mixture at 1 atm and room temperature. Where H_2 = 25%, CO = 75% and N_2/O_2 = 3.76. Experimental data are taken from[243].

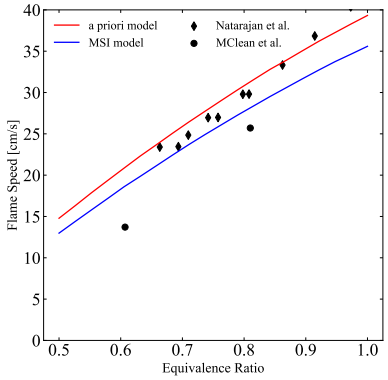


Figure 31: Laminar fLaminar flame speed of $H_2/O_2/CO/N_2$ mixture at 1 atm and room temperature. Where H_2 = 5%, CO = 95% and N_2/O_2 = 3.76. Experimental data are taken from [242, 240].

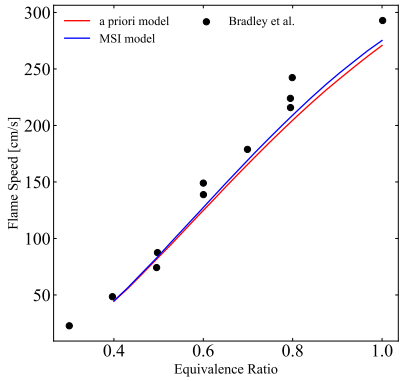


Figure 32: Laminar flame speed of $H_2/O_2/N_2$ mixture at 1 atm and room temperature. Where N_2/O_2 = 3.76. Experimental data are taken from [244].

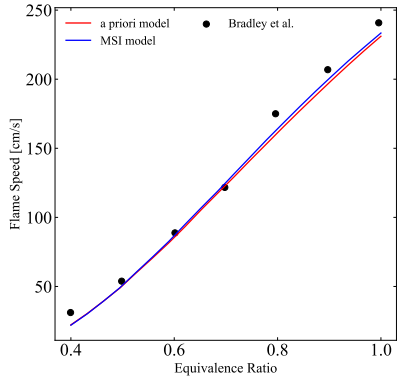


Figure 33: Laminar flame speed of $H_2/O_2/N_2$ mixture at 5 atm and room temperature. Where N_2/O_2 = 3.76. Experimental data are taken from [244].

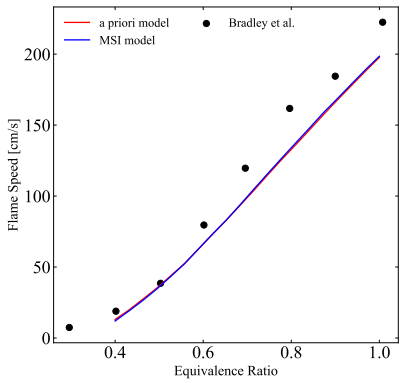


Figure 34: Laminar flame speed of $H_2/O_2/CO/N_2$ mixture at 10 atm and room temperature. Where N_2/O_2 = 3.76. Experimental data are taken from [244].

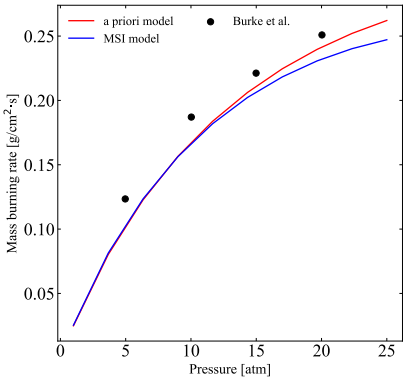


Figure 35: Mass burning rates of hydrogen for a stoichiometric mixture of H_2/O_2 diluted in He for a 1800 K flame temperature. Experimental data are taken from [245].

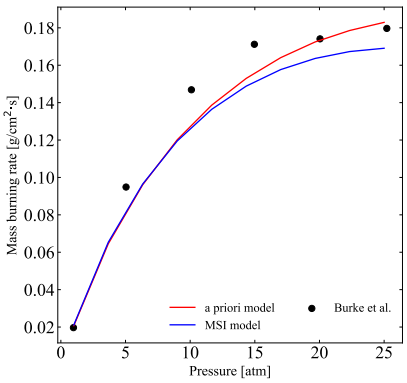


Figure 36: Mass burning rates of hydrogen for a stoichiometric mixture of H_2/O_2 diluted in He for a 1700 K flame temperature. Experimental data are taken from [245].

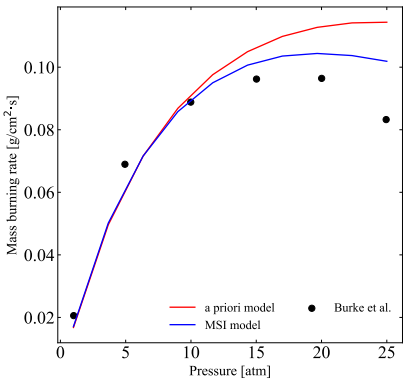


Figure 37: Mass burning rates of hydrogen for a stoichiometric mixture of H_2/O_2 diluted in He for a 1600 K flame temperature. Experimental data are taken from [245].

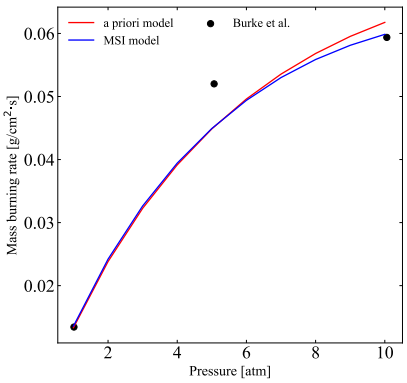


Figure 38: Mass burning rates of hydrogen for a stoichiometric mixture of H_2/O_2 diluted in He for a 1500 K flame temperature. Experimental data are taken from [245].

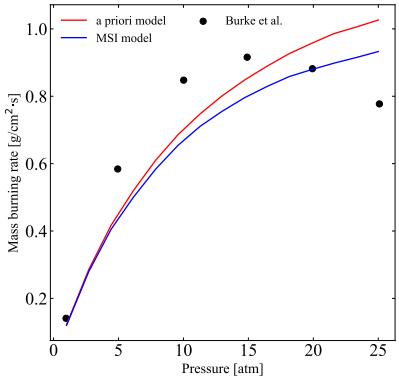


Figure 39: Mass burning rates of hydrogen for a rich mixture (ϕ = 2.5) of H_2/O_2 diluted in Ar for a 1600 K flame temperature. Experimental data are taken from [245].

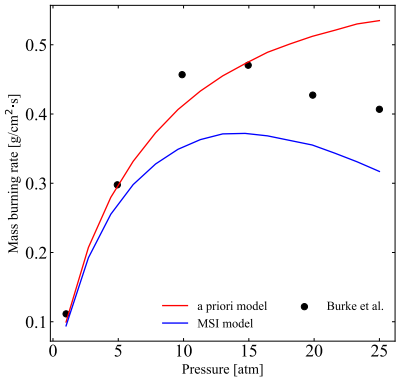


Figure 40: Mass burning rates of hydrogen for a rich mixture (ϕ = 2.5) of H_2/O_2/CO diluted in Ar for a 1600 K flame temperature. Where H_2 = 50% and CO = 50%. Experimental data are taken from [245].

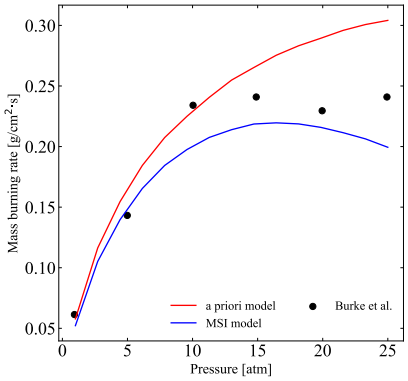


Figure 41: Mass burning rates of hydrogen for a rich mixture (ϕ = 2.5) of $H_2/O_2/CO$ diluted in Ar for a 1600 K flame temperature. Where H_2 = 10% and CO = 90%. Experimental data are taken from [245].

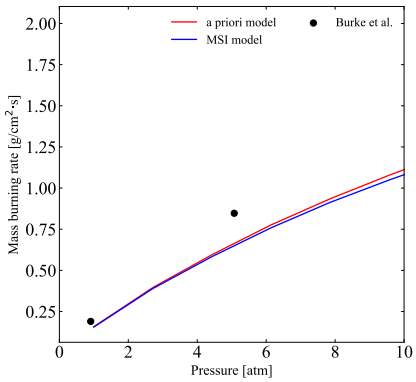


Figure 42: Mass burning rates of hydrogen for a rich mixture (ϕ = 2.5) of H_2/O_2 diluted in Ar for a 1500 K flame temperature. Experimental data are taken from [245].

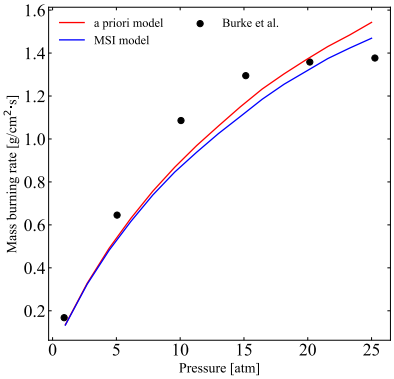


Figure 43: Mass burning rates of hydrogen for a rich mixture (ϕ = 2.5) of H_2/O_2 diluted in Ar for a 1600 K flame temperature. Experimental data are taken from [245].

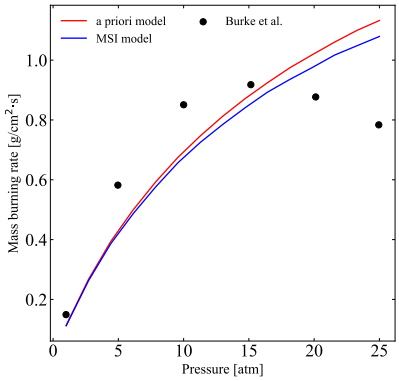


Figure 44: Mass burning rates of hydrogen for a rich mixture ($\phi = 2.5$) of H_2/O_2 diluted in Ar for a 1700 K flame temperature. Experimental data are taken from [245].

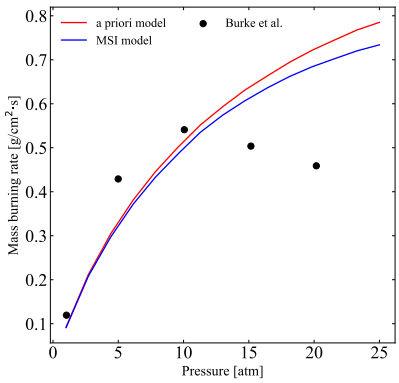


Figure 45: Mass burning rates of hydrogen for a rich mixture ($\phi = 2.5$) of H_2/O_2 diluted in Ar for a 1800 K flame temperature and 1 atm. Experimental data are taken from [245].

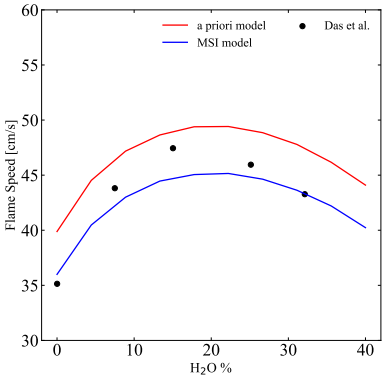


Figure 46: Laminar flame speed for a lean mixture ($\phi = 0.9$) of $H_2/O_2/CO/N_2$ at 1 atm and 323 K. Where CO = 95%, H_2 = 5% and N_2/O_2 = 3.76. Experimental data are taken from [246].

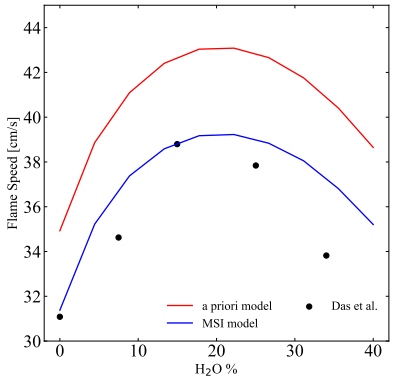


Figure 47: Laminar flame speed for a lean mixture ($\phi = 0.8$) of $H_2/O_2/CO/N_2$ at 1 atm and 323 K. Where CO = 95%, H_2 = 5% and N_2/O_2 = 3.76. Experimental data are taken from [246].

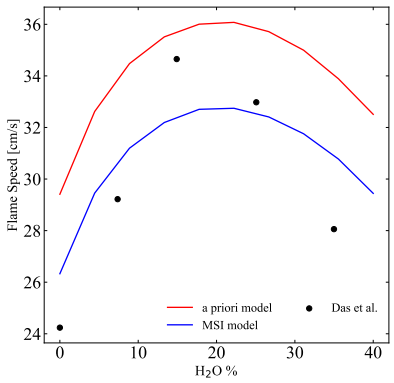


Figure 48: Laminar flame speed for a lean mixture ($\phi = 0.7$) of $H_2/O_2/CO/N_2$ at 1 atm and 323 K. Where CO = 95%, H_2 = 5% and N_2/O_2 = 3.76. Experimental data are taken from [246].

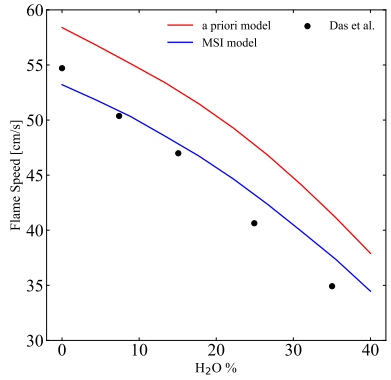


Figure 49: Laminar flame speed for a lean mixture ($\phi = 0.6$) of $H_2/O_2/CO/N_2$ at 1 atm and 323 K. Where CO = 50%, H_2 = 50% and N_2/O_2 = 3.76. Experimental data are taken from [246].

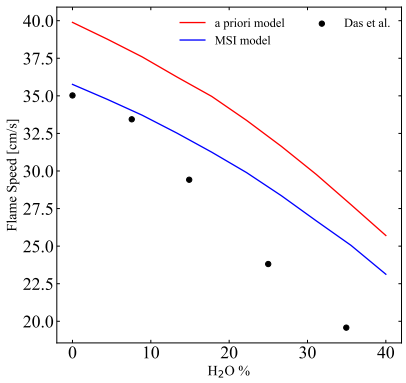


Figure 50: Laminar flame speed for a lean mixture (ϕ = 0.5) of $H_2/O_2/CO/N_2$ at 1 atm and 323 K. Where CO = 50%, H_2 = 50% and N_2/O_2 = 3.76. Experimental data are taken from [246].

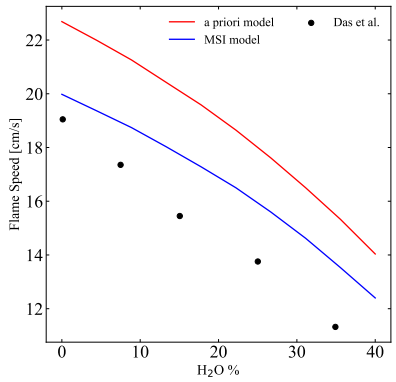


Figure 51: Laminar flame speed for a lean mixture (ϕ = 0.4) of $H_2/O_2/CO/N_2$ at 1 atm and 323 K. Where CO = 50%, H_2 = 50% and N_2/O_2 = 3.76. Experimental data are taken from [246].

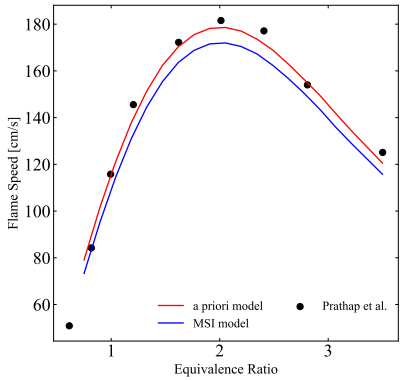


Figure 52: Laminar flame speed for a $H_2/O_2/N_2$ mixture at 1 bar and 302 K. Where N_2/O_2 = 3.76. Experimental data are taken from [247].

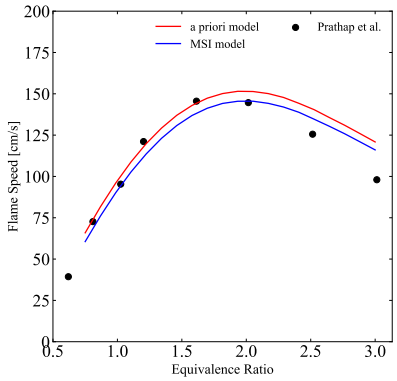


Figure 53: Laminar flame speed for a $H_2/O_2/N_2$ mixture at 1 bar and 302 K. Where N_2 = 20%. Experimental data are taken from [247].

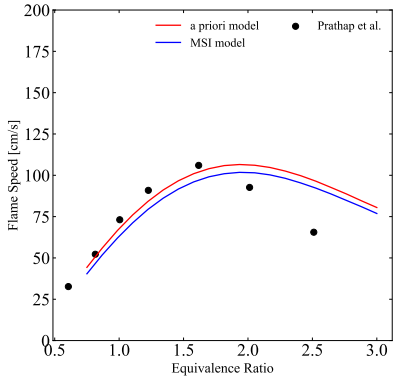


Figure 54: Laminar flame speed for a $H_2/O_2/N_2$ mixture at 1 bar and 302 K. Where N_2 = 40%. Experimental data are taken from [247].

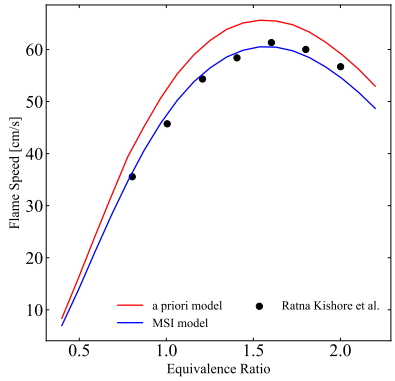


Figure 55: Laminar flame speed for a $H_2/O_2/CO/CO_2/N_2$ mixture at 1 bar and 302 K. Where CO = 30%, H_2 = 30%, CO = 40% and N_2/O_2 = 3.76. Experimental data are taken from [248].

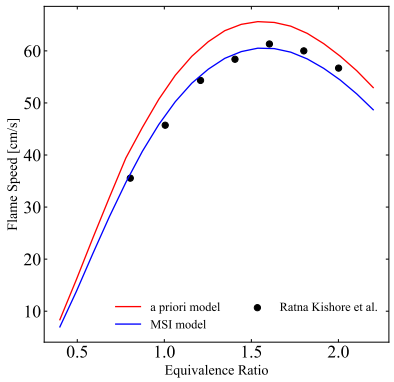


Figure 56: Laminar flame speed for a $H_2/O_2/CO/CO_2/N_2$ mixture at 1 bar and 302 K. Where CO = 30%, H_2 = 30%, CO_2 = 40% and N_2/O_2 = 3.76. Experimental data are taken from [248].

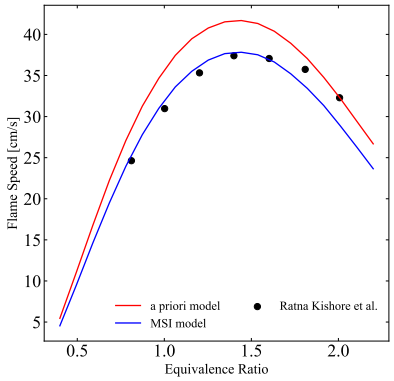


Figure 57: Laminar flame speed for a $H_2/O_2/CO/CO_2/N_2$ mixture at 1 bar and 302 K. Where CO = 25%, H_2 = 25%, CO_2 = 50% and N_2/O_2 = 3.76. Experimental data are taken from [248].

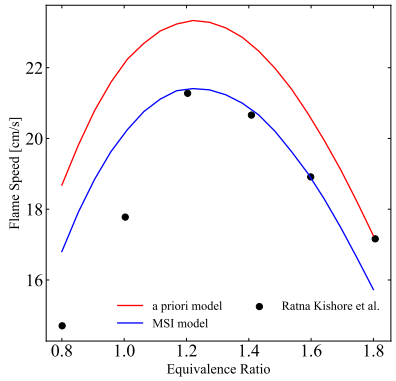


Figure 58: Laminar flame speed for a $H_2/O_2/CO/CO_2/N_2$ mixture at 1 bar and 302 K. Where CO = 20%, H_2 = 20%, CO_2 = 40% and N_2/O_2 = 3.76. Experimental data are taken from [248].

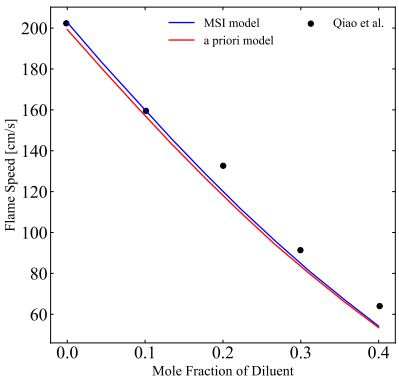


Figure 59: Laminar flame speed for a stoichiometric $H_2/O_2/N_2$ mixture at 0.5 bar and 298 K, as a function of N_2 suppressant. Experimental data are taken from [249].

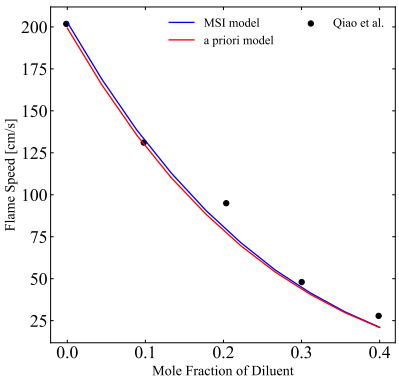


Figure 60: Laminar flame speed for a stoichiometric $H_2/O_2/CO_2/N_2$ at 0.5 bar and 298 K, as a function of CO_2 suppressant. Experimental data are taken from [249].

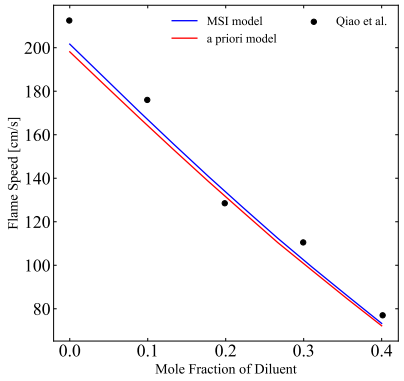


Figure 61: Laminar flame speed for a stoichiometric $H_2/O_2/N_2$/Ar mixture at 1 bar and 298 K, as a function of Ar suppressant. Experimental data are taken from [249].

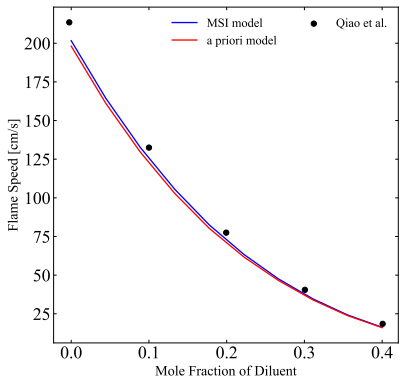


Figure 62: Laminar flame speed for a stoichiometric $H_2/O_2/N_2$ mixture at 1 bar and 298 K, as a function of N_2 suppressant. Experimental data are taken from [249].

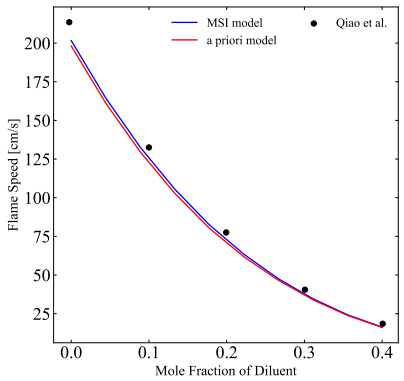


Figure 63: Laminar flame speed for a stoichiometric $H_2/O_2/N_2/CO_2$ mixture at 1 bar and 298 K, as a function of CO_2 suppressant. Experimental data are taken from [249].

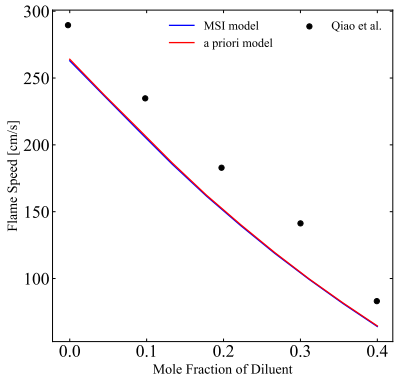


Figure 64: Laminar flame speed for a rich mixture (ϕ = 1.8) $H_2/O_2/N_2$ at 1 bar and 298 K, as a function of N_2 suppressant. Experimental data are taken from [249].

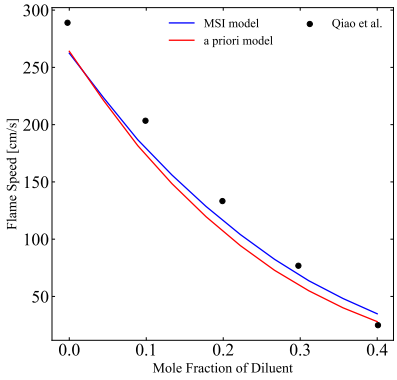


Figure 65: Laminar flame speed for a rich mixture (ϕ = 1.8) $H_2/O_2/N_2/CO_2$ at 1 bar and 298 K, as a function of O_2 suppressant. Experimental data are taken from [249].

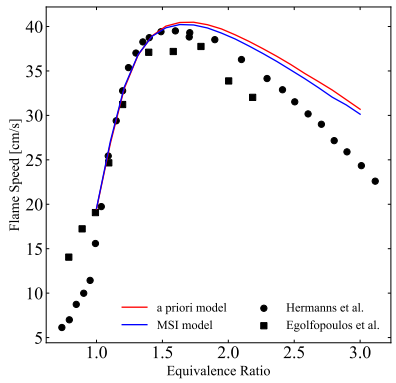


Figure 66: Laminar flame speed for a $H_2/O_2/N_2$ mixture at 1 bar and 298 K. Where $O_2/(O_2+N_2)$=0.077 Experimental data are taken from [238, 250].

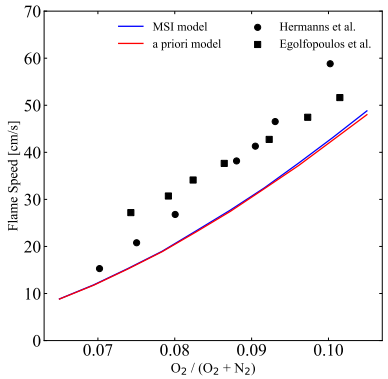


Figure 67: Laminar flame speed for a rich mixture (ϕ = 1.058) $H_2/O_2/N_2$ at 1 bar and 298 K. Experimental data are taken from [238, 250].

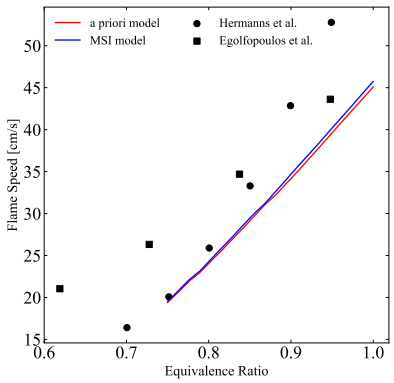


Figure 68: Laminar flame speed for a $H_2/O_2/N_2$ mixture diluted in at 1 bar and 298 K. Where $O_2/(O_2+N_2)$=0.1077 Experimental data are taken from [238, 250].

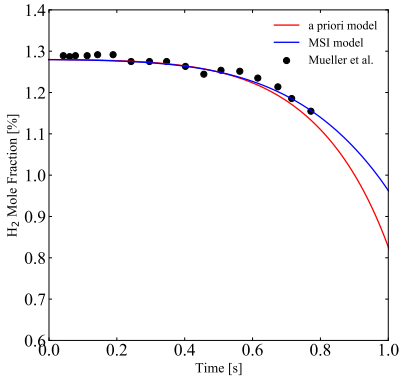


Figure 69: H_2 mole fraction profile measured in a flow reactor at 6.5 atm and 884 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

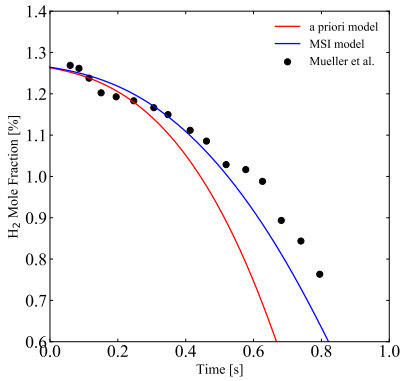


Figure 70: H_2 mole fraction profile measured in a flow reactor at 6.5 atm and 889 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

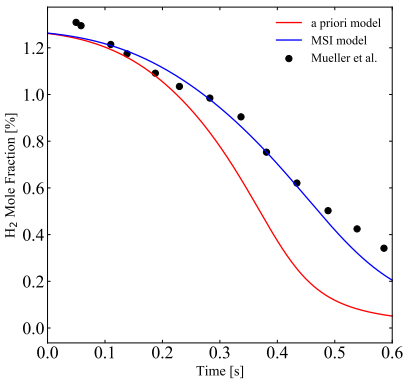


Figure 71: H_2 mole fraction profile measured in a flow reactor at 6.5 atm and 906 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

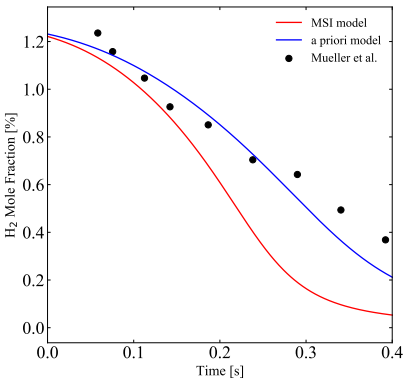


Figure 72: H_2 mole fraction profile measured in a flow reactor at 6.5 atm and 914 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

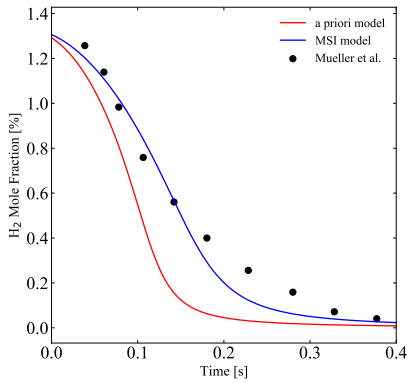


Figure 73: H_2 mole fraction profile measured in a flow reactor at 6.5 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

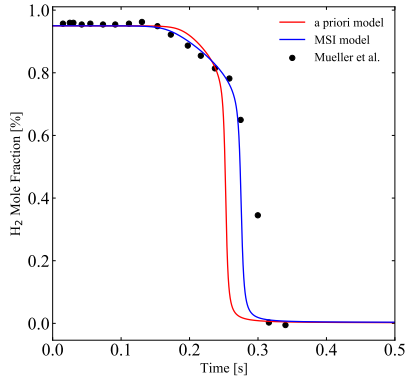


Figure 74: H_2 mole fraction profile measured in a flow reactor at 3.02 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

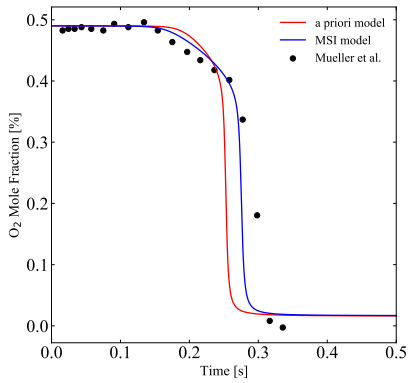


Figure 75: O_2 mole fraction profile measured in a flow reactor at 3.02 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

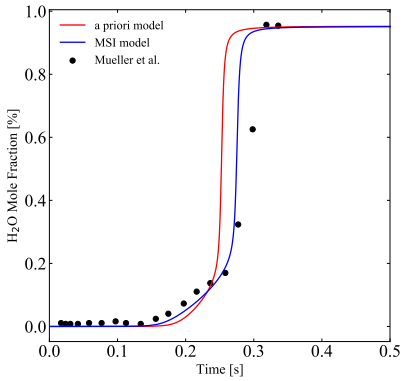


Figure 76: H_2O mole fraction profile measured in a flow reactor at 3.02 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

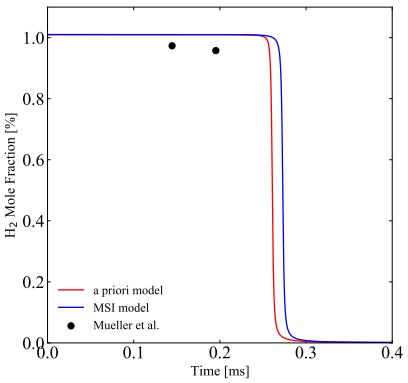


Figure 77: H_2 mole fraction profile measured in a flow reactor at 2.55 atm and 935 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

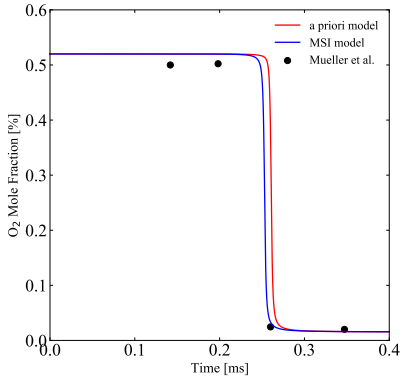


Figure 78: O_2 mole fraction profile measured in a flow reactor at 2.55 atm and 935 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

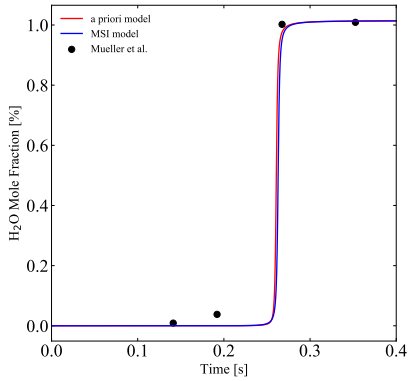


Figure 79: H_2O mole fraction profile measured in a flow reactor at 2.55 atm and 935 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

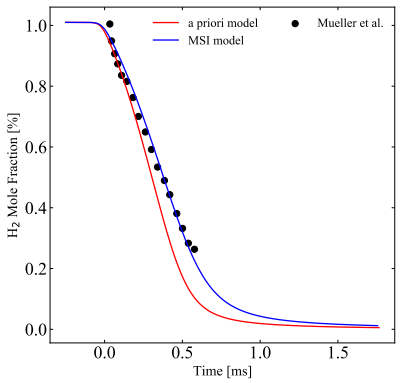


Figure 80: H_2 mole fraction profile measured in a flow reactor at 6.0 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

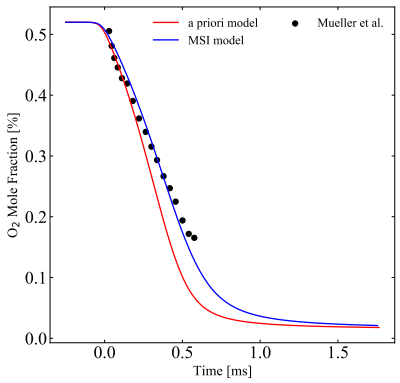


Figure 81: O_2 mole fraction profile measured in a flow reactor at 6.0 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

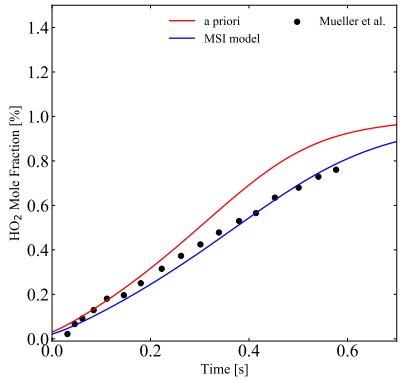


Figure 82: H_2O mole fraction profile measured in a flow reactor at 6.0 atm and 934 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

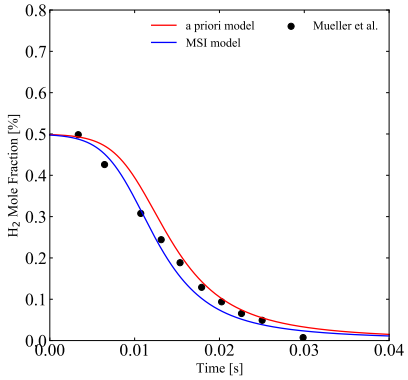


Figure 83: H_2 mole fraction profile measured in a flow reactor at 0.3 atm and 880 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

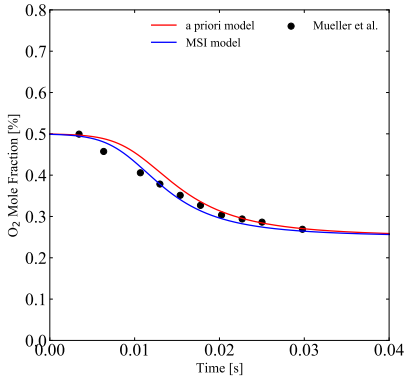


Figure 84: O_2 mole fraction profile measured in a flow reactor at 0.3 atm and 880 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

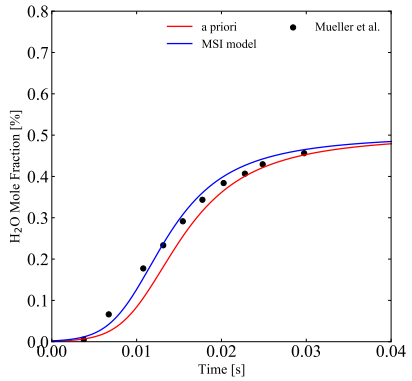


Figure 85: H_2O mole fraction profile measured in a flow reactor at 0.3 atm and 880 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

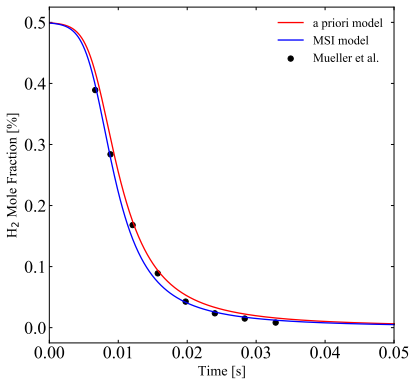


Figure 86: H_2 mole fraction profile measured in a flow reactor at 0.6 atm and 896 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

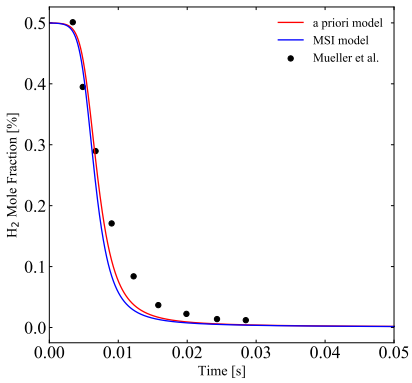


Figure 87: H_2 mole fraction profile measured in a flow reactor at 0.6 atm and 896 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

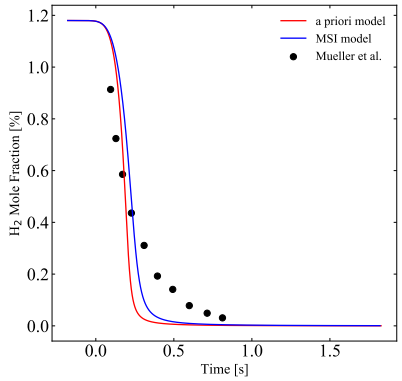


Figure 88: H_2 mole fraction profile measured in a flow reactor at 17.7 atm and 914 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

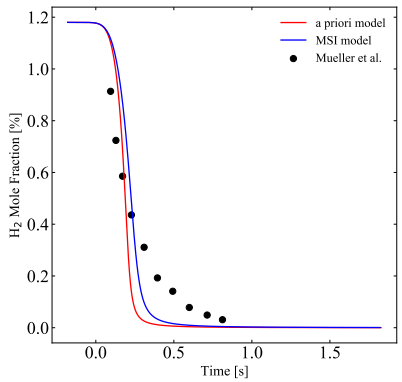


Figure 89: H_2 mole fraction profile measured in a flow reactor at 15.7 atm and 914 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Mueller et al. [222].

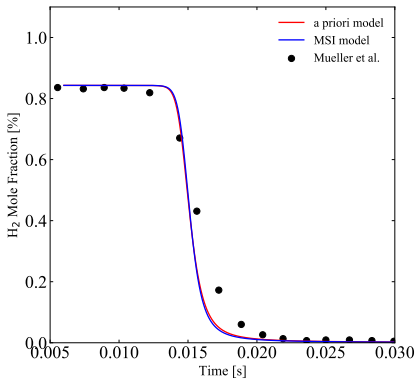


Figure 90: H_2 mole fraction profile measured in a flow reactor at 1.0 atm and 910 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Yetter et al. [251].

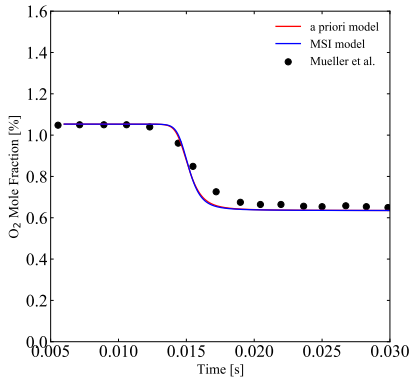


Figure 91: O_2 mole fraction profile measured in a flow reactor at 1.0 atm and 910 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Yetter et al. [251].

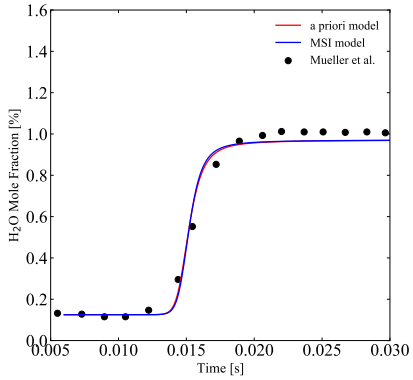


Figure 92: H_2O mole fraction profile measured in a flow reactor at 1.0 atm and 910 K from a $H_2/O_2/N_2$ mixture. Experimental data are from Yetter et al. [251].

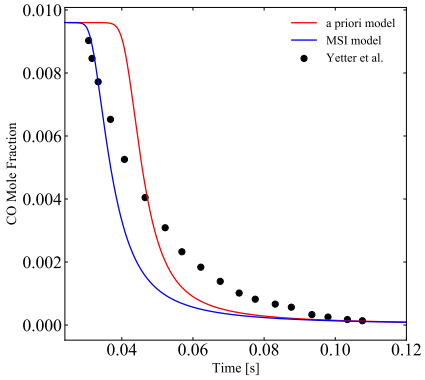


Figure 93: CO mole fraction profile measured in a flow reactor at 1.0 atm and 1033 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Yetter et al. [251].

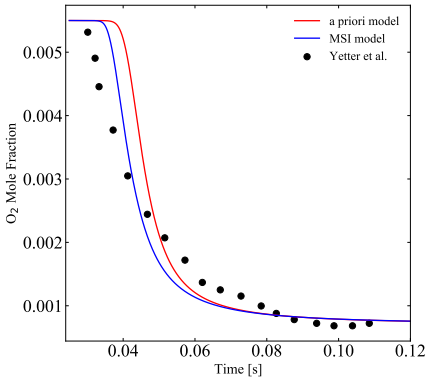


Figure 94: O_2 mole fraction profile measured in a flow reactor at 1.0 atm and 1033 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Yetter et al. [251].

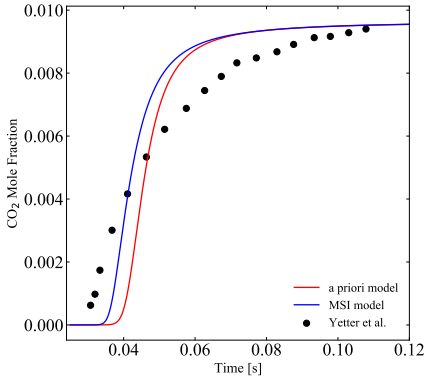


Figure 95: CO_2 mole fraction profile measured in a flow reactor at 1.0 atm and 1033 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Yetter et al. [251].

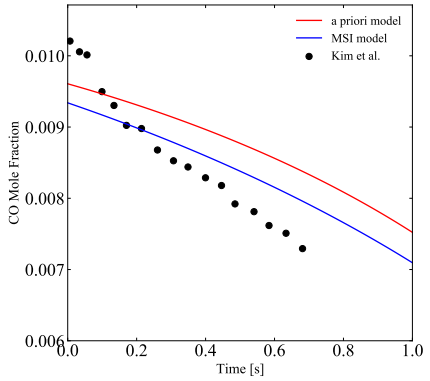


Figure 96: CO mole fraction profile measured in a flow reactor at 6.5 atm and 1030 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

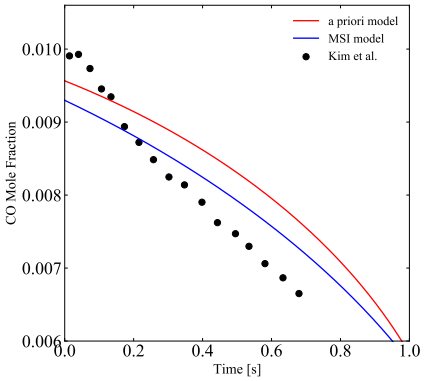


Figure 97: CO mole fraction profile measured in a flow reactor at 6.5 atm and 1040 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

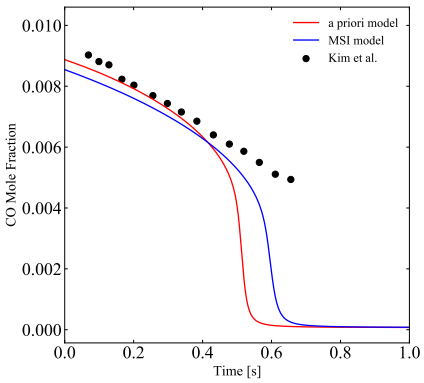


Figure 98: CO mole fraction profile measured in a flow reactor at 6.5 atm and 1052 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

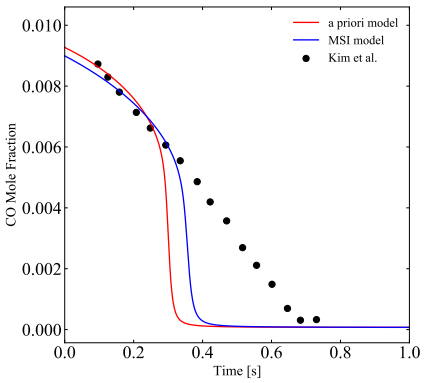


Figure 99: CO mole fraction profile measured in a flow reactor at 6.5 atm and 1068 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

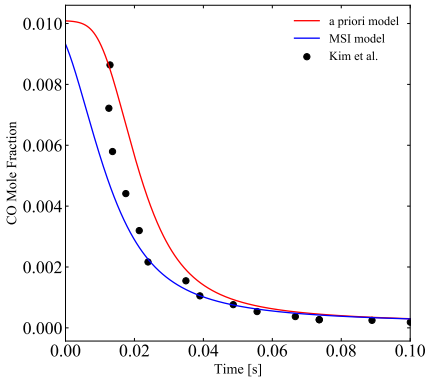


Figure 100: CO mole fraction profile measured in a flow reactor at 1.0 atm and 1040 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

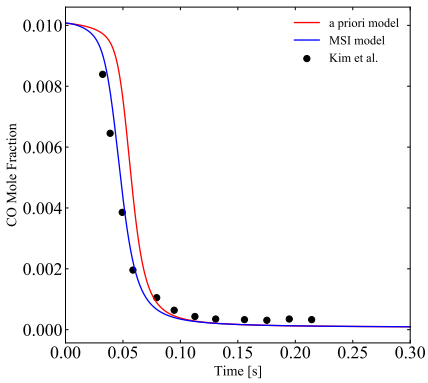


Figure 101: CO mole fraction profile measured in a flow reactor at 2.45 atm and 1040 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

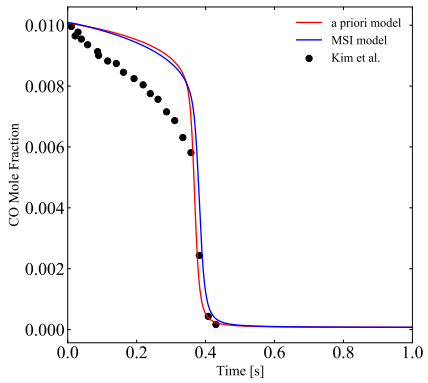


Figure 102: CO mole fraction profile measured in a flow reactor at 3.46 atm and 1040 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

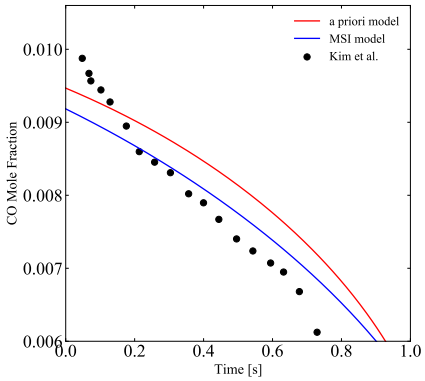


Figure 103: CO mole fraction profile measured in a flow reactor at 6.5 atm and 1040 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

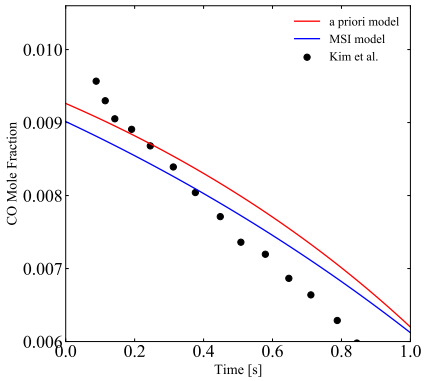


Figure 104: CO mole fraction profile measured in a flow reactor at 9.6 atm and 1040 K from a $CO/H_2/O_2/N_2$ mixture. Experimental data are from Kim et al. [252].

Chapter 8: Conclusion and Future Works

Combustibles are essential for powering numerous aspects of the transportation sector as well as providing long term storage to the energy grid. However, the harmful byproducts of traditional combustibles makes them an unappealing option for continued use. Therefore, the need to develop, study and understand renewable fuel options and the associated chemical energy conversion technology is imperative. These studies at their core are facilitated by the understanding of gas phase kinetics.

Computational simulations, specifically CFD, have become one of the most powerful approaches for developing and advancing energy technology in a safe and efficient manner; however, such simulations must have sound underlying chemical kinetics models to make accurate predictions. The specific challenges for developing accurate kinetics models at combustion conditions were outlined– specifically, the necessity of unprecedented accuracy to obtain truly predictive modeling.

Using the Active Thermochemical Tables (ATcT) [41] as inspiration, the MultiScale Informatics (MSI) method was proposed to aid in the development of accurate chemical kinetics models. Discussed within this work were the details regarding how the MSI method can be used to identify optimized values and quantified uncertainties for a set of molecular parameters (within theoretical kin tics calculations), rate parameters, and physical model parameters (within simulations of experimental observables) as informed by data from various sources and scales. These details included outlining the mathematical framework and the software implementation.

The MSI approach was implemented for two foundational combustion systems (CH_3 + HO_2 and HO_2 + HO_2) and a third much larger network of reactions. These sub-systems demonstrated some key elements of the MSI approach, including: the importance of raw data for quantifying the information content of experimental data, the utility of theoretical kinetics calculations for con-

straining experimental interpretations and providing an independent data source, and the subtleties of target data selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties. Exploring these three subsystems also demonstrated how MSI can unravel a web of interconnected rate constants in a complex system of reactions and demonstrated how MSI can resolve discrepancies among data sets as well as theoretical and experimental determinations.

All three of the sub-systems demonstrated the power of using the MSI method and its promise to aid in development of kinetics models in the future. The three reaction systems culminated in a hydrogen model, for which initial validation testing was performed. Preliminary results for validation against combustion targets of various types were presented to begin assessment of model performance.

While the research performed herein has successfully proved that the MSI approach is a powerful approach for further exploring chemical kinetics systems and developing more accurate chemical kinetics models, there are additional steps to consider for improving other important aspects of this research. Provided below is a preliminary list of recommended research topics requiring further investigation.

8.1 Software

As noted in Chapter 3, the MSI Theory software and the MSI Library are not currently, computationally connected. Ideally, the MSI Theory code would become part of the MSI Library allowing users to avoid the steps of manually transferring output from the MSI Theory code into the MSI Library input file. It would make it possible for users to circumvent the surrogate model currently implemented into the MSI Library, meaning the MSI Theory code could be executed each iteration of the optimization, thereby proving extremely beneficial for more complicated and less linear systems. Computationally linking the MSI Theory code and the MSI Library requires making some changes to the MSI Library source code, and writing additional software. First, the required inputs need to be updated to include the MSI Theory main file. Second, a call must be made to execute the MSI Theory code, which should be added directly after the MSI input file is parsed. Third,

additional software must be written to extract the sensitivity coefficients and Chebyshev rate constant fits from the text files the MSI Theory code returns. (These rate constant fits would need to be added to the master equation cti file and the sensitivity coefficients need to be passed into the MSI Library). Fourth, additional software must be written to update MESS or Variflex files with optimized values for molecular parameters.

An additional improvement worth implementing to the MSI software is parallelizing the Opt Runner module, which would allow for various physical modules to be executed simultaneously and decrease the computational time needed to run the MSI Library. This would require a wrapper to be written around the Opt Runner, thereby allowing for the parallel execution of the sub-modules. The wrapper would need to keep track of the simulation executions and return the data in the same format as the current simulation dictionary.

There are also additional physical modules that should be added to the MSI library, including flame speeds and batch reactors that use a prescribed pressure trace. These additional physical modules would make possible the incorporation of more raw data into the MSI framework as optimization targets. This would require the necessary sub-module(s) be constructed and added to the Opt Runner.

Another feature worth implementing into the MSI software is the ability of different physical models to be executed using different, or slightly modified, cti files. Doing so would allow for the use of slightly different Chebyshev fits for the same rate constant to be utilized based on the experiment being simulated. This is advantageous because Chebyshev rate constants do not currently have the ability to handle different bath gasses (which have different collisional efficiencies). The ability to simulate different experiments using different cti files would also allow physical modules that might not require the full set of chemistry contained in a larger cti file to be executed faster. This would reduce the codes overall run time. In order to implement this change the cti Core module would need the ability to load multiple processors. The code would keep track of each processor and a method would need to be developed to identify which processor should be used to execute a specific simulation.

8.2 Additional MESS / Variflex Calculations

To have the most comprehensive model, it would be extremely beneficial to have some additional reactions represented by theoretical molecular parameters rather than Arrhenius coefficients. These reactions would include all group A reactions, reaction L2 and reaction P1. As previously mentioned, group A reactions are fundamental building blocks of many combustion systems, so it would be exceptionally advantageous to represent them with theoretical molecular parameters. However, given the current state of theoretical kinetics calculations and the inherent complexities associated with these reactions, this is perhaps more of a long term goal.

Presently, implementing reactions L2 and P1 with theoretical molecular parameters would make the model even more robust. As mentioned in the results of Chapter 6, there are various rate constant determinations for reaction L2 that have a much steeper temperature dependence than the MSI model would suggest. A step towards providing a resolution for these discrepancies would be to represent this reaction with theoretical molecular parameters. Similar to the group A reactions, R-P1 is also a fundamental building block in many combustion systems. This reaction has less variation in the rate constant determinations at combustion relevant temperatures than that of R-L2, however constraining it across the full temperature range with theoretical molecular parameters would help add additional robustness to the model. Specifically, it may help address the models issues regarding production of species profiles generated from initial mixtures containing CO.

8.3 Comparison Against Other Published Models

After the remaining model inconsistencies discussed in Chapter 7 are addressed and the MSI model is performing well for all validation data, the next step is to compare the MSI model performance against other well known and published hydrogen models. Potential models for comparison are Burke [111], FFCM [84], USC [209] and Keromnes [225]. Comparing the MSI model against other published hydrogen models can be used to assess where the model is performing well and

if there are any shortcomings in predictions as compared to other models. After comparing the MSI model against other published models to benchmark performance, the next step would be to publish the model thereby making it available for use by the combustion community.

References

[1] Y. Matsuo, A. Yanagisawa, and Y. Yamashita, "A global energy outlook to 2035 with strategic considerations for Asia and Middle East energy supply and demand interdependencies," *Energy Strategy Reviews*, vol. 2, no. 1, pp. 79–91, 2013.

[2] S. Pacala and R. H. Socolow, "Stabilization wedges: Solving the climate problem for the next 50 years with current technologies," *Science*, vol. 305, no. 5686, pp. 968–972, 2004.

[3] R. H. Socolow and S. H. Lam, "Good enough tools for global warming policy making," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 365, no. 1853, pp. 897–934, 2007.

[4] B. E. Layton, "A comparison of energy densities of prevalent energy sources in units of joules per cubic meter," *International Journal of Green Energy*, vol. 5, no. 6, pp. 438–455, 2008.

[5] R. Heinberg and D. Fridley, *Our renewable future laying the path for 100% clean energy*. Island Press/Center for Resource Economics, 2016.

[6] S. Verhelst, J. W. Turner, L. Sileghem, and J. Vancoillie, "Methanol as a fuel for internal combustion engines," *Progress in Energy and Combustion Science*, vol. 70, pp. 43–88, 2019.

[7] C. Liu, Z. Li, Y. Pei, and Y. An, "Methanol as a fuel for internal combustion engines," *Engines and Fuels for Future Transport*, pp. 281–324, 2022.

[8] C. Hobson, C. Márquez, *et al.*, "Renewable methanol report," *Methanol Institute, Singapore*, 2018.

[9] P. G. Cifre and O. Badr, "Renewable hydrogen utilisation for the production of methanol," *Energy conversion and management*, vol. 48, no. 2, pp. 519–527, 2007.

[10] X. Zhen and Y. Wang, "An overview of methanol as an internal combustion engine fuel," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 477–493, 2015.

[11] W. Yanju, L. Shenghua, L. Hongsong, Y. Rui, L. Jie, and W. Ying, "Effects of methanol/-gasoline blends on a spark ignition engine performance and emissions," *Energy & Fuels*, vol. 22, no. 2, pp. 1254–1259, 2008.

[12] V. Masson-Delmotte *et al.*, "2021: Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change," *Cambridge University Press*, 2021.

[13] O. Edenhofer, *Climate change 2014: mitigation of climate change*. Cambridge University Press, 2015, vol. 3.

[14] P. S. Shah, "Cold starting of methanol-fueled engines using direct fuel injection system," Ph.D. dissertation, 1992.

[15] V. Kumar and A. K. Agarwal, "Material compatibility aspects and development of methanol-fueled engines," *Advanced Combustion Techniques and Engine Technologies for the Automotive Sector*, pp. 37–51, 2020.

[16] S Verhelst and T Wallner, "Hydrogen-fueled internal combustion engines," *Progress in energy and combustion science*, vol. 35, no. 6, pp. 490–527, 2009.

[17] R. Babayev, H. G. Im, A. Andersson, and B. Johansson, "Hydrogen double compression-expansion engine: A sustainable internal combustion engine with 60%+ brake thermal efficiency potential at 45 bar," *Energy Conversion and Management*, vol. 264, p. 115 698, 2022.

[18] R. Daiyan, I. MacGill, and R. Amal, "Opportunities and challenges for renewable power-to-x," *ACS Energy Letters*, vol. 5, no. 12, pp. 3843–3847, 2020.

[19] J. Koponen *et al.*, "Review of water electrolysis technologies and design of renewable hydrogen production systems," 2015.

[20] Y. H. Teoh *et al.*, "A review on production and implementation of hydrogen as a green fuel in internal combustion engines," *Fuel*, vol. 333, p. 126 525, 2023.

[21] L. M. Das, "Hydrogen-fueled internal combustion engines," in *Compendium of hydrogen energy*, Elsevier, 2016, pp. 177–217.

[22] S. Verhelst, "Recent progress in the use of hydrogen as a fuel for internal combustion engines," *international journal of hydrogen energy*, vol. 39, no. 2, pp. 1071–1085, 2014.

[23] M Berckmüller *et al.*, "Potentials of a charged SI-hydrogen engine," SAE Technical Paper, Tech. Rep., 2003.

[24] R. Babayev, A. Andersson, A. S. Dalmau, H. G. Im, and B. Johansson, "Computational characterization of hydrogen direct injection and nonpremixed combustion in a compression-ignition engine," *International Journal of Hydrogen Energy*, vol. 46, no. 35, pp. 18 678–18 696, 2021.

[25] K. Schmietendorf, J. Peinke, and O. Kamps, "The impact of turbulent renewable energy production on power grid stability and quality," *The European Physical Journal B*, vol. 90, no. 11, pp. 1–6, 2017.

[26] O. Smith, O. Cattell, E. Farcot, R. D. O'Dea, and K. I. Hopcraft, "The effect of renewable energy incorporation on power grid stability and resilience," *Science advances*, vol. 8, no. 9, 2022.

[27] S. C. Johnson, J. D. Rhodes, and M. E. Webber, "Understanding the impact of non-synchronous wind and solar generation on grid stability and identifying mitigation pathways," *Applied Energy*, vol. 262, p. 114 492, 2020.

[28] P. Breeze, "Chapter 8 - hydrogen energy storage," in *Power System Energy Storage Technologies*, P. Breeze, Ed., Academic Press, 2018, pp. 69–77, ISBN: 978-0-12-812902-9.

[29] E. Wolf, "Chapter 9 - large-scale hydrogen energy storage," in *Electrochemical Energy Storage for Renewable Sources and Grid Balancing*, P. T. Moseley and J. Garche, Eds., Amsterdam: Elsevier, 2015, pp. 129–142, ISBN: 978-0-444-62616-5.

[30] N. A. Sepulveda, J. D. Jenkins, A. Edington, D. S. Mallapragada, and R. K. Lester, "The design space for long-duration energy storage in decarbonized power systems," *Nature Energy*, vol. 6, no. 5, pp. 506–516, 2021.

[31] "Pathways to sustainable energy: Accelerating energy transition in the UNECE region," *United Nations Economic Commision for Europe*, vol. Ece Energy Series No. 67, 2020.

[32] S. Tracey and B. Anne, "OECD contribution to the united nations commission on sustainable development," *Organization for Economic Cooperation and Development*, vol. 15, 2007.

[33] W. Hwang, J. Dec, and M. Sjöberg, "Spectroscopic and chemical-kinetic analysis of the phases of HCCI autoignition and combustion for single- and two-stage ignition fuels," *Combustion and Flame*, vol. 154, no. 3, pp. 387–409, 2008.

[34] Y. Shi, H. Ge, and R. D. Reitz, *Computational Optimization of Internal Combustion Engines*. 2011.

[35] H. Wang, M. Yao, and R. D. Reitz, "Development of a reduced primary reference fuel mechanism for internal combustion engine combustion simulations," *Energy and Fuels*, vol. 27, no. 12, pp. 7843–7853, 2013.

[36] R. D. Reitz, "Directions in internal combustion engine research," *Combustion and Flame*, vol. 160, no. 1, pp. 1–8, 2013.

[37] S. Som *et al.*, "Quantum tunneling affects engine performance," *The Journal of Physical Chemistry Letters*, vol. 4, no. 12, pp. 2021–2025, 2013.

[38] R. D Reitz and G. Duraisamy, "Review of high efficiency and clean reactivity controlled compression ignition (RCCI) combustion in internal combustion engines," *Progress in Energy and Combustion Science*, vol. 46, pp. 12–71, 2015.

[39] G. Coskun *et al.*, "Analysis of in-cylinder H_2O_2 and HO_2 distributions in an HCCI engine–Comparison of laser-diagnostic results with CFD and SRM simulations," *Combustion and Flame*, vol. 162, no. 9, pp. 3131–3139, 2015.

[40] C. K. Law, *Combustion Physics*. Cambridge university press, 2010.

[41] B. Ruscic *et al.*, "Introduction to active thermochemical tables: Several Key enthalpies of formation revisited," *Journal of Physical Chemistry A*, vol. 108, no. 45, pp. 9979–9997, 2004.

[42] M. P. Burke, F. L. Dryer, and Y. Ju, "Assessment of kinetic modeling for lean $H_2/CH_4/O_2$/diluent flames at high pressures," *Proceedings of the Combustion Institute*, vol. 33, no. 1, pp. 905–912, 2011.

[43] H. Wang and D. A. Sheen, "Combustion kinetic model uncertainty quantification, propagation and minimization," *Progress in Energy and Combustion Science*, vol. 47, pp. 1–31, 2015.

[44] D. L. Baulch *et al.*, "Evaluated Kinetic Data for Combustion Modeling: Supplement II," *Journal of Physical and Chemical Reference Data*, vol. 34, no. 3, pp. 757–1397, 2005.

[45] M. P. Burke and S. J. Klippenstein, "Ephemeral collision complexes mediate chemically termolecular transformations that affect system chemistry," *Nature Chemistry*, vol. 9, no. 11, pp. 1078–1082, 2017.

[46] L. Lei and M. P. Burke, "An extended methodology for automated calculations of non-Boltzmann kinetic sequences: H + C_2H_2 + X and combustion impact," *Proceedings of the Combustion Institute*, vol. 38, no. 1, pp. 661–669, 2021.

[47] N. J. Labbe, R. S., C. F. Goldsmith, Y. Georgievskii, J. A. Miller, and S. J. Klippenstein, "Weakly Bound Free Radicals in Combustion: Prompt Dissociation of Formyl Radicals and Its Effect on Laminar Flame Speeds," *Journal of Physical Chemistry Letters*, vol. 7, no. 1, pp. 85–89, 2016.

[48] M. A. Blitz *et al.*, "Reanalysis of rate data for the reaction CH_3 + CH_3 = C_2H_6 using revised cross sections and a linearized second-order master equation," *Journal of Physical Chemistry A*, vol. 119, no. 28, pp. 7668–7682, 2015.

[49] R. J. Shannon, A. S. Tomlin, S. H. Robertson, M. A. Blitz, M. J. Pilling, and P. W. Seakins, "Global uncertainty propagation and sensitivity analysis in the $CH_3OCH_2 + O_2$ system: Combining experiment and theory to constrain key rate coefficients in DME combustion," *Journal of Physical Chemistry A*, vol. 119, no. 28, pp. 7430–7438, 2015.

[50] J. Prager, H. N. Najm, and J. Zádor, "Uncertainty quantification in the ab initio rate-coefficient calculation for the $CH_3CH(OH)CH_3 + OH = CH_3C(OH)CH_3 + H_2O$ reaction," *Proceedings of the Combustion Institute*, vol. 34, no. 1, pp. 583–590, 2013.

[51] M. Frenklach, A. Packard, P. Seiler, and R. Feeley, "Collaborative data processing in developing predictive models of complex reaction systems," *International Journal of Chemical Kinetics*, vol. 36, no. 1, pp. 57–66, 2004.

[52] R. Feeley, P. Seiler, A. Packard, and M. Frenklach, "Consistency of a Reaction Dataset," *The Journal of Physical Chemistry A*, vol. 108, no. 44, pp. 9573–9583, 2004.

[53] K. Braman, T. A. Oliver, and V. Raman, "Bayesian analysis of syngas chemistry models," *Combustion Theory and Modelling*, vol. 17, no. 5, pp. 858–887, 2013.

[54] K. Miki, E. E. Prudencio, S. H. Cheung, and G. Terejanu, "Using Bayesian analysis to quantify uncertainties in the $H+O_2$=OH+O reaction," *Combustion and Flame*, vol. 160, no. 5, pp. 861–869, 2013.

[55] L. Cai and H. Pitsch, "Mechanism optimization based on reaction rate rules," *Combustion and Flame*, vol. 161, no. 2, pp. 405–415, 2014.

[56] S. Mosbach, J. H. Hong, G. P. E. Brownbridge, M. Kraft, S. Gudiyella, and K. Brezinsky, "Bayesian error propagation for a kinetic model of n-propylbenzene oxidation in a shock tube," *International Journal of Chemical Kinetics*, vol. 46, no. 7, pp. 389–404, 2014.

[57] D. R. Yeates *et al.*, "Integrated data-model analysis facilitated by an instrumental model," *Proceedings of the Combustion Institute*, vol. 35, no. 1, pp. 597–605, 2015.

[58] M. Frenklach, "Transforming data into knowledge-Process Informatics for combustion chemistry," *Proceedings of the Combustion Institute*, vol. 31 I, no. 1, pp. 125–140, 2007.

[59] D. A. Sheen and H. Wang, "Modeling High-Pressure $H_2/CO/O_2$/diluent Mass Burning Rates with the Method of Uncertainty Minimization using Polynomial Chaos Expansions.," in *7th US National Technical Meeting of the Combustion Institute*, Atlanta GA, 2011, pp. 20–22.

[60] T. Varga *et al.*, "Optimization of a hydrogen combustion mechanism using both direct and indirect measurements," *Proceedings of the Combustion Institute*, vol. 35, no. 1, pp. 589–596, 2015.

[61] C. Olm, I. G. Zsély, T. Varga, H. J. Curran, and T. Turányi, "Comparison of the performance of several recent syngas combustion mechanisms," *Combustion and Flame*, vol. 162, no. 5, pp. 1793–1812, 2015.

[62] D. A. Sheen, X. You, H. Wang, and T. Løvås, "Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion," *Proceedings of the Combustion Institute*, vol. 32, no. 1, pp. 535–542, 2009.

[63] D. A. Sheen and H. Wang, "Combustion kinetic modeling using multispecies time histories in shock-tube oxidation of heptane," *Combustion and Flame*, vol. 158, no. 4, pp. 645–656, 2011.

[64] T. Turányi *et al.*, "Determination of rate parameters based on both direct and indirect measurements," *International Journal of Chemical Kinetics*, vol. 44, no. 5, pp. 284–302, 2012.

[65] C. K. Westbrook and F. L. Dryer, "Chemical kinetic modeling of hydrocarbon combustion," *Progress in Energy and Combustion Science*, vol. 10, no. 1, pp. 1–57, 1984.

[66] H Curran, *Saudi Aramco Mechanism 2.0.* 2016.

[67] J. Li, Z. Zhao, A. Kazakov, M. Chaos, F. L. Dryer, and J. I. Scire, "A comprehensive kinetic mechanism for CO, CH_2O, and CH_3OH combustion," *International Journal of Chemical Kinetics*, vol. 39, no. 3, pp. 109–136, 2007.

[68] M. Christensen, E. J. K. Nilsson, and A. A. Konnov, "A Systematically Updated Detailed Kinetic Model for CH_2O and CH_3OH Combustion," *Energy and Fuels*, vol. 30, no. 8, pp. 6709–6726, 2016.

[69] M. P. Burke, S. J. Klippenstein, and L. B. Harding, "A quantitative explanation for the apparent anomalous temperature dependence of OH + HO_2=H_2O+O_2 through multi-scale modeling," *Proceedings of the Combustion Institute*, vol. 34, no. 1, pp. 547–555, 2013.

[70] M. P. Burke *et al.*, "Multiscale Informatics for Low-Temperature Propane Oxidation: Further Complexities in Studies of Complex Reactions," *Journal of Physical Chemistry A*, vol. 119, no. 28, pp. 7095–7115, 2015.

[71] M. P. Burke, "Harnessing the Combined Power of Theoretical and Experimental Data through Multiscale Informatics," *International Journal of Chemical Kinetics*, vol. 48, no. 4, pp. 212–235, 2016.

[72] C. E. LaGrotta, M. C. Barbet, L. Lei, and M. P. Burke, "Towards a high-accuracy kinetic database informed by theoretical and experimental data: CH_3 + HO_2 as a case study," *Proceedings of the Combustion Institute*, vol. 38, pp. 1043–1051, 2020.

[73] M. P. Burke and R. Song, "Evaluating Mixture Rules for Multi-Component Pressure Dependence: H + O_2 (+M) = HO_2 (+M)," *Proceedings of the Combustion Institute*, vol. 36, no. 1, pp. 245–253, 2017.

[74] M. C. Barbet, K. McCullough, and M. P. Burke, "A framework for automatic discovery of chemically termolecular reactions," *Proceedings of the Combustion Institute*, vol. 37, no. 1, pp. 347–354, 2019.

[75] L. Lei and M. P. Burke, "Evaluating mixture rules and combustion implications for multicomponent pressure dependence of allyl + HO_2 reactions," *Proceedings of the Combustion Institute*, vol. 37, no. 1, pp. 355–362, 2019.

[76] L. Lei and M. P. Burke, "Bath gas mixture effects on multichannel reactions: Insights and representations for systems beyond single-channel reactions," *The Journal of Physical Chemistry A*, vol. 123, no. 3, pp. 631–649, 2018.

[77] L. Lei and M. P. Burke, "Mixture rules and falloff are now major uncertainties in experimentally derived rate parameters for H + O_2 (+M) = HO_2 (+M)," *Combustion and Flame*, vol. 213, pp. 467–474, 2020.

[78] L. Lei and P Burke M., "Dynamically Evaluating Mixture Effects on Multi-Channel Reactions in Flames: A Case Study for the CH_3 + OH Reaction," *Proceedings of The Combustion Institute (in review)*, pp. 433–440, 2020.

[79] R. E. Cornell, M. C. Barbet, and M. P. Burke, "Automated discovery of influential chemically termolecular reactions in energetic material combustion: A case study for RDX," *Proceedings of the Combustion Institute*, vol. 38, no. 1, pp. 787–794, 2021.

[80] S. J. Klippenstein and C. Cavallotti, "Ab initio kinetics for pyrolysis and combustion systems," *Computer Aided Chemical Engineering*, vol. 45, pp. 115–167, 2019.

[81] Y. Georgievskii, J. A. Miller, M. P. Burke, and S. J. Klippenstein, "Reformulation and solution of the master equation for multiple-well chemical reactions," *Journal Physical Chemistry A*, vol. 117, no. 46, pp. 12 146–12 154, 2013.

[82] S. J. Klippenstein, A. F. Wagner, R. C. Dunbar, D. M. Wardlaw, S. H. Robertson, and J. A. Miller, "Variflex," 2005, version 1.14m.

[83] D. G. Goodwin, R. L. Speth, H. K. Moffat, and B. W. Weber, *Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes*, Version 2.5.1, 2021.

[84] G. Smith, Y. Tao, and H. Wang, *Foundational Fuel Chemistry Model Version 1.0 (FFCM-1)*, 2016.

[85] C. Olm, T. Varga, É. Valkó, H. J. Curran, and T. Turányi, "Uncertainty quantification of a newly optimized methanol and formaldehyde combustion mechanism," *Combustion and Flame*, vol. 186, pp. 45–64, 2017.

[86] L. Lei, C. E. LaGrotta, and M. C. Barbet, "MSI theory," *GitHub repository*, 2020, version 1.0.

[87] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.

[88] K. Simonov and et al., "Yaml is a human-friendly data serialization language for all programming languages," vol. 400, pp. 20–25, 2005.

[89] B. W. Weber and K. E. Niemeyer, "Chemked: A human-and machine-readable data standard for chemical kinetics experiments," *International Journal of Chemical Kinetics*, vol. 50, no. 3, pp. 135–148, 2018.

[90] W. McKinney and et al., "Data structures for statistical computing in python," vol. 445, pp. 51–56, 2010.

[91] C. R. Harris *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020.

[92] M. Sangwan and L. N. Krasnoperov, "Kinetics of the gas phase reaction CH_3 + HO_2," *Journal of Physical Chemistry A*, vol. 117, no. 14, pp. 2916–2923, 2013.

[93] A. S. Orhan, Z. Dulger, N. Kahraman, and N. T. Veziroglu, "Internal combustion engines fueled by natural gas—hydrogen mixtures," *Journal of Hydrogen Energy*, vol. 29, no. 14, pp. 1527–1539, 2004.

[94] A. C. Lipardi, J. M. Bergthorson, and G. Bourque, "No_x emissions modeling and uncertainty from exhaust-gas-diluted flames," *Journal of Engineering for Gas Turbines and Power*, vol. 138, no. 5, 2016.

[95] S. J. Klippenstein, Y. Georgievskii, and L. B. Harding, "Statistical theory for the kinetics and dynamics of roaming reactions," *Journal of Physical Chemistry A*, vol. 115, no. 50, pp. 14 370–14 381, 2011.

[96] N. K. Srinivasan, J. V. Michael, L. B. Harding, and S. J. Klippenstein, "Experimental and theoretical rate constants for CH_4 + O_2 = CH_3 + HO_2," *Combustion and Flame*, vol. 149, pp. 104–111, 2007.

[97] A. W. Jasper, S. J. Klippenstein, and L. B. Harding, "Theoretical rate coefficients for the reaction of methyl radical with hydroperoxyl radical and for methylhydroperoxide decomposition," *Proceedings of the combustion institute*, vol. 32, no. 1, pp. 279–286, 2009.

[98] D. M. Medvedev, L. B. Harding, and S. K. Gray, "Methyl radical: Ab initio global potential surface, vibrational levels and partition function," *Journal of Molecular Physics*, vol. 104, no. 1, pp. 73–81, 2006.

[99] N. J. Labbe, A. W. Jasper, J. A. Miller, and S. J. Klippenstein, "High Accuracy Thermochemical Kinetics for H+CH_3(+M)↔CH_4(+M)," *10^{th} U.S. National Combustion Meeting*, 1A01, 2017.

[100] T. Gierczak, R. K. Talukdar, S. C. Herndon, C. L. Vaghjiani, and A. R. Ravishankara, "Rate coefficients for the reactions of hydroxyl radicals with methane and deuterated methanes," *Journal of Physical Chemistry A*, vol. 101, no. 17, pp. 3125–3134, 1997.

[101] J. R. Dunlop and F. P. Tully, "A kinetic study of hydroxyl radical reactions with methane and perdeuterated methane," *Journal of Physical Chemistry*, vol. 97, no. 43, pp. 11 148–11 150, 1993.

[102] Z. Hong, R. D. Cook, D. F. Davidson, and R. K. Hanson, "A shock tube study of OH + H_2O_2 = H_2O + HO_2 and H_2O_2 + M = 2OH + M using laser absorption of H_2O and OH," *Journal of Physical Chemistry A*, vol. 114, no. 18, pp. 5718–5727, 2010.

[103] Z. Hong, A. Farooq, E. A. Barbour, D. F. Davidson, and R. K. Hanson, "Hydrogen peroxide decomposition rate: a shock tube study using tunable laser absorption of H_2O near 2.5 micro meters," *Journal of Physical Chemistry A*, vol. 113, no. 46, pp. 12 919–12 925, 2009.

[104] Z. Hong, S. S. Vasu, D. F. Davidson, and R. K. Hanson, "Experimental study of the rate of OH + HO_2 = H_2O + O_2 at high temperatures using the reverse reaction," *Journal Physical Chemistry A*, vol. 114, no. 17, pp. 5520–5525, 2010.

[105] C. Kappel, K. Luther, and J. Troe, "Shock wave study of the unimolecular dissociation of H_2O_2 in its falloff range and of its secondary reactions," *Physical Chemistry Chemical Physics*, vol. 4, no. 18, pp. 4392–4398, 2002.

[106] Z. Hong, K-Y Lam, R. Sur, S. Wang, D. F. Davidson, and R. K. Hanson, "On the rate constants of OH + HO_2 and HO_2 + HO_2: A comprehensive study of H_2O_2 thermal decomposition using multi-species laser absorption," *Proceedings of the Combustion Institute*, vol. 34, no. 1, pp. 565–571, 2013.

[107] C. C. Kircher and S. P. Sander, "Kinetics and mechanism of hydroperoxo and hydroperoxod disproportionations," *The Journal of Physical Chemistry*, vol. 88, no. 10, pp. 2082–2091, 1984.

[108] Z. Hong, D. F. Davidson, K-Y Lam, and R. K. Hanson, "A shock tube study of the rate constants of HO_2 and CH_3 reactions," *Combustion and Flame*, vol. 159, no. 10, pp. 3007–3013, 2012.

[109] J. J. Scire Jr., R. A. Yetter, and F. L. Dryer, "Flow reactor studies of methyl radical oxidation reactions in methane-perturbed moist carbon monoxide oxidation at high pressure with model sensitivity analysis," *International Journal of Chemical Kinetics*, vol. 33, no. 2, pp. 75–100, 2001.

[110] M. C. Barbet and M. P. Burke, "Impact of "Missing" Third-Body Efficiencies on Kinetic Model Predictions of Combustion Properties," *Proceedings of the Combustion Institute*, vol. 38, pp. 425–432, 2020.

[111] M. P. Burke, M. Chaos, Y. Ju, F. L. Dryer, and S. J. Klippenstein, "Comprehensive H_2/O_2 kinetic model for high-pressure combustion," *International Journal of Chemical Kinetics*, vol. 44, no. 7, pp. 444–474, 2012.

[112] M. P. Burke, M. Chaos, F. L. Dryer, and Y. Ju, "Negative pressure dependence of mass burning rates of $H_2/CO/O_2$/diluent flames at low flame temperatures," *Combustion and Flame*, vol. 157, no. 4, pp. 618–631, 2010.

[113] S. J. Klippenstein *et al.*, "$H\dot{O}_2$ + $H\dot{O}_2$: High level theory and the role of singlet channels," *Combustion and Flame*, p. 111 975, 2022.

[114] W. R. Stockwell, "On the HO_2+ HO_2 reaction: Its misapplication in atmospheric chemistry models," *Journal of Geophysical Research: Atmospheres*, vol. 100, no. D6, pp. 11 695–11 698, 1995.

[115] S. R. Albrecht *et al.*, "Measurements of hydroperoxy radicals (HO_2) at atmospheric concentrations using bromide chemical ionisation mass spectrometry," *Atmospheric Measurement Techniques*, vol. 12, no. 2, pp. 891–902, 2019.

[116] R. Patrick and M. J. Pilling, "The temperature dependence of the HO_2+ HO_2 reaction," *Chemical Physical Letters*, vol. 91, no. 5, pp. 343–347, 1982.

[117] P. D. Lightfoot, B. Veyret, and R. Lesclaux, "The rate constant for the HO_2+ HO_2 reaction at elevated temperatures," *Chemical Physical Letters*, vol. 150, no. 1-2, pp. 120–126, 1988.

[118] H. Hippler, J. Troe, and J. Willner, "Shock wave study of the reaction HO_2 + HO_2 = H_2O_2 + O_2: Confirmation of a rate constant minimum near 700 k," *The Journal of Chemical Physics*, vol. 93, no. 3, pp. 1755–1760, 1990.

[119] R. A. Cox, J. P. Burrows, and T. J. Wallington, "Rate coefficient for the reaction OH+ HO_2= H_2O+ O_2 at 1 atmosphere pressure and 308 K," *Chemical Physical Letters*, vol. 84, no. 2, pp. 217–221, 1981.

[120] W. B. DeMore, "Reaction of hydroperoxo radicals with ozone and the effect of water vapor on hydroperoxo kinetics," *Journal of Physical Chemistry*, vol. 83, no. 9, pp. 1113–1118, 1979.

[121] W. DeMore, "Rate constant and possible pressure dependence of the reaction hydroxyl+ hydroperoxo," *The Journal of Physical Chemistry*, vol. 86, no. 1, pp. 121–126, 1982.

[122] R-R Lii, R. A. Gorse Jr., M. C. Sauer Jr., and S. Gordon, "Temperature dependence of the gas-phase self-reaction of hydroperoxo radicals in the presence of ammonia," *Journal of Physical Chemistry*, vol. 84, no. 8, pp. 813–817, 1980.

[123] M. J. Kurylo, O. Klais, and A. H. Laufer, "Mechanistic investigation of the hydroxyl+ hydroperoxo reaction," *Journal of Physical Chemistry*, vol. 85, no. 24, pp. 3674–3678, 1981.

[124] M. Braun, A. Hofzumahaus, and F. Stuhl, "VUV flash photolysis study of the reaction of ho with HO_2 at 1 atm and 298 K," *Berichte der Bunsengesellschaft für physikalische Chemie*, vol. 86, no. 7, pp. 597–602, 1982.

[125] L. F. Keyser, "Kinetics of the reaction hydroxyl + hydroperoxo = water + oxygen from 254 to 382 K," *Journal of Physical Chemistry*, vol. 92, no. 5, pp. 1193–1200, 1988.

[126] N. K. Srinivasan, M. Su, J. W. Sutherland, J. V. Michael, and B. Ruscic, "Reflected shock tube studies of high-temperature rate constants for OH+ NO_2= HO_2+ NO and OH+ HO_2= H_2O+ O_2," *Journal of Physical Chemistry A*, vol. 110, no. 21, pp. 6602–6607, 2006.

[127] D. D. Zhou, K. Han, P. Zhang, L. B. Harding, M. J. Davis, and R. T. Skodje, "Theoretical determination of the rate coefficient for the HO_2 + HO_2 = H_2O_2 + O_2 reaction: Adiabatic treatment of anharmonic torsional effects," *Journal of Physical Chemistry A*, vol. 116, no. 9, pp. 2089–2100, 2012.

[128] H. Hippler, H. Neunaber, and J. Troe, "Shock wave studies of the reactions HO+ H_2O_2= H_2O+ HO_2 and HO+ HO_2= H_2O+ O_2 between 930 and 1680 K," *Journal Chemical Physics.*, vol. 103, no. 9, pp. 3510–3516, 1995.

[129] S. J. Klippenstein, L. B. Harding, and B. Ruscic, "Ab initio computations and active thermochemical tables hand in hand: Heats of formation of core combustion species," *The Journal of Physical Chemistry A*, vol. 121, no. 35, pp. 6580–6602, 2017.

[130] A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, "W4 theory for computational thermochemistry: In pursuit of confident sub-kj/mol predictions," *The Journal of Chemical Physics*, vol. 125, no. 14, p. 144 108, 2006.

[131] M. E. Harding, J. Vázquez, B. Ruscic, A. K. Wilson, J. Gauss, and J. F. Stanton, "High-accuracy extrapolated ab initio thermochemistry. III. additional improvements and overview," *The Journal of Chemical Physics*, vol. 128, no. 11, p. 114 111, 2008.

[132] H. M. Jaeger, H. F. Schaefer III, J. Demaison, A. G. Császár, and W. D. Allen, "Lowest-lying conformers of alanine: Pushing theory to ascertain precise energetics and semiex-

perimental r_e structures," *Journal of Chemical Theory and Computation*, vol. 6, no. 10, pp. 3066–3078, 2010.

[133] K. A. Peterson, D. Feller, and D. A. Dixon, "Chemical accuracy in ab initio thermochemistry and spectroscopy: Current strategies and future challenges," *Theoretical Chemistry Accounts*, vol. 131, no. 1, pp. 1–20, 2012.

[134] W. B. DeMore *et al.*, "Chemical kinetics and photochemical data for use in stratospheric modeling, jpl publication 97-4," *Jet Propulsion Laboratory, Pasadena, CA*, pp. 1755–1760, 1997.

[135] M. Mozurkewich and S. W. Benson, "Self-reaction of ho_2 and do_2: Negative temperature dependence and pressure effects," *International Journal of Chemical Kinetics*, vol. 17, no. 8, pp. 787–807, 1985.

[136] R. Atkinson *et al.*, "Evaluated kinetic and photochemical data for atmospheric chemistry: Volume i-gas phase reactions of O_x, HO_x, $NO_x$and SO_x species," *Atmospheric Chemistry and Physics*, vol. 4, no. 6, pp. 1461–1738, 2004.

[137] S. R. Sellevåg, Y. Georgievskii, and J. A. Miller, "Kinetics of the gas-phase recombination reaction of hydroxyl radicals to form hydrogen peroxide," *The Journal of Physical Chemistry A*, vol. 113, no. 16, pp. 4457–4467, 2009.

[138] S. P. Sander *et al.*, "Chemical kinetics and photochemical data for use in atmospheric studies, evaluation number 14," *Journal of Physical Chemistry A*, vol. 116, no. 9, pp. 2089–2100, 2011.

[139] A. A. Konnov, "On the role of excited species in hydrogen combustion," *Combustion and Flame*, vol. 162, no. 10, pp. 3755–3772, 2015.

[140] M. P. Burke, Q. Meng, and C. Sabaitis, "Dissociation-induced depletion of high-energy reactant molecules as a mechanism for pressure-dependent rate constants for bimolecular reactions," *Faraday Discussions*, pp. 1–29, 2022.

[141] J. A. Miller and S. J. Klippenstein, "Master equation methods in gas phase chemical kinetics," *The Journal of Physical Chemistry A*, vol. 110, no. 36, pp. 10 528–10 544, 2006.

[142] L. Lei and M. P. Burke, "Understanding and representing the distinct kinetics induced by reactive collisions of rovibrationally excited ephemeral complexes across reactive collider mole fractions and pressures," *The Journal of Physical Chemistry A*, vol. 124, no. 52, pp. 10 937–10 953, 2020.

[143] M. Khalil and H. N. Najm, "Probabilistic inference of reaction rate parameters from summary statistics," *Combustion Theory and Modelling*, vol. 22, no. 4, pp. 635–665, 2018.

[144] A. B. S. Alquaity, K. C. Utsav, A. Popov, and A. Farooq, "Detection of shock-heated hydrogen peroxide (H_2O_2) by off-axis cavity-enhanced absorption spectroscopy OA-CEAS," *Applied Physics B*, vol. 123, p. 280, 2017.

[145] X. Yang, X. Shen, P. Zhao, and C. K. Law, "Statistical analysis on rate parameters of the H_2–O_2 reaction system," *The Journal of Physical Chemistry A*, vol. 125, no. 47, pp. 10 223–10 234, 2021.

[146] C. E. LaGrotta, Q. Meng, L. Lei, M. C. Barbet, Z. Hong, and M. P. Burke, "Resolving discrepancies between state-of-the-art theory and experiment for HO_2 + HO_2 via multiscale informatics," *The Journal of Physical Chemistry A*, 2023.

[147] X. Chen, M. E. Fuller, and C. F. Goldsmith, "Decomposition kinetics for HONO and HNO_2," *Reaction Chemistry & Engineering*, vol. 4, no. 2, pp. 323–333, 2019.

[148] S. J. Klippenstein, Y. Georgievskii, and L. B. Harding, "Predictive theory for the combination kinetics of two alkyl radicals," *Physical Chemistry Chemical Physics*, vol. 8, no. 10, pp. 1133–1147, 2006.

[149] A. W. Jasper, S. J. Klippenstein, L. B. Harding, and B. Ruscic, "Kinetics of the reaction of methyl radical with hydroxyl radical and methanol decomposition," *The Journal of Physical Chemistry A*, vol. 111, no. 19, pp. 3932–3950, 2007.

[150] J. V. Michael, "Measurement of thermal rate constants by flash or laser photolysis in shock tubes: Oxidations of H_2 and D_2," *Progress in Energy and Combustion Science*, vol. 18, no. 4, pp. 327–347, 1992.

[151] R. C. Oldenborg, G. W. Loge, D. M. Harradine, and K. R. Winn, "Kinetic study of the hydrogel + hydrogen reaction from 800 to 1550 K," *The Journal of Physical Chemistry*, vol. 96, no. 21, pp. 8426–8430, 1992.

[152] J. R. Rumble, "Nist 40. ndrl/nist solution kinetics database: Version 3.0," 2008.

[153] L. F. Keyser, "Absolute rate constant and branching fractions for the atomic hydrogen+ hydroperoxyl radical reaction from 245 to 300 K," *The Journal of Physical Chemistry*, vol. 90, no. 13, pp. 2994–3003, 1986.

[154] U. Sridharan, L. Qiu, and F Kaufman, "Kinetics and product channels of the reactions of perhydroxyl with oxygen and hydrogen atoms at 296 k," *The Journal of Physical Chemistry*, vol. 86, no. 23, pp. 4569–4574, 1982.

[155] J. Michael, D. Nava, W. Payne, J. Lee, and L. Stief, "Rate constant for the reaction atomic hydrogen+ nitrogen dioxide from 195 to 400 k with fp-rf and df-rf techniques," *Journal of Physical Chemistry*, vol. 83, no. 22, pp. 2818–2823, 1979.

[156] T Ko and A Fontijn, "High-temperature photochemistry kinetics study of the reaction h+ no2 yields oh+ no from 296 to 760 k.(reannouncement with new availability information)," Rensselaer Polytechnic Inst., Troy, NY (United States), Tech. Rep. 10, 1991, pp. 3984–3987.

[157] I. R. Slagle, D. Gutman, J. W. Davies, and M. J. Pilling, "Study of the recombination reaction methyl+ methyl. fwdarw. ethane. 1. experiment," *The Journal of Physical Chemistry*, vol. 92, no. 9, pp. 2455–2462, 1988.

[158] J. Kiefer, S Santhanam, N. Srinivasan, R. Tranter, S. Klippenstein, and M. Oehlschlaeger, "Dissociation, relaxation, and incubation in the high-temperature pyrolysis of ethane, and a successful rrkm modeling," *Proceedings of the Combustion Institute*, vol. 30, no. 1, pp. 1129–1135, 2005.

[159] L. N. Krasnoperov, E. N. Chesnokov, H. Stark, and A. R. Ravishankara, "Unimolecular dissociation of formyl radical, HCO = H + CO, studied over 1- 100 bar pressure range," *The Journal of Physical Chemistry A*, vol. 108, no. 52, pp. 11 526–11 536, 2004.

[160] R. S. Timonen, E. Ratajczak, D. Gutman, and A. F. Wagner, "The addition and dissociation reaction atomic hydrogen+ carbon monoxide. dblharw. oxomethyl. 2. experimental studies and comparison with theory," *Journal of Physical Chemistry*, vol. 91, no. 20, pp. 5325–5332, 1987.

[161] F. L. Nesbitt, J. F. Gleason, and L. J. Stief, "Temperature dependence of the rate constant for the reaction HCO + O_2 = HO_2 + CO at T = 200-398 K," *The Journal of Physical Chemistry A*, vol. 103, no. 16, pp. 3038–3043, 1999.

[162] J. D. DeSain, L. E. Jusinski, A. D. Ho, and C. A. Taatjes, "Temperature dependence and deuterium kinetic isotope effects in the HCO (DCO) + O_2 reaction between 296 and 673 K," *Chemical physics letters*, vol. 347, no. 1-3, pp. 79–86, 2001.

[163] A. A. Westenberg and N DeHaas, "Measurement of the rate constant for H + H_2CO = h_2 + HCO at 297-652. deg. K," *The Journal of Physical Chemistry*, vol. 76, no. 16, pp. 2213–2214, 1972.

[164] R. B. Klemm, "Absolute rate parameters for the reactions of formaldehyde with o atoms and h atoms over the temperature range 250–500 k," *The Journal of Chemical Physics*, vol. 71, no. 5, pp. 1987–1993, 1979.

[165] V Sivakumaran, D. Hölscher, T. J. Dillon, and J. N. Crowley, "Reaction between oh and hcho: Temperature dependent rate coefficients (202–399 k) and product pathways (298 k)," *Physical Chemistry Chemical Physics*, vol. 5, no. 21, pp. 4821–4827, 2003.

[166] Z. Hong, D. F. Davidson, E. A. Barbour, and R. K. Hanson, "A new shock tube study of the H+ O_2 = OH+ O reaction rate using tunable diode laser absorption of H_2O near 2.5 μm," *Proceedings of the Combustion Institute*, vol. 33, no. 1, pp. 309–316, 2011.

[167] D. A. Masten, R. K. Hanson, and C. T. Bowman, "Shock tube study of the reaction hydrogen atom+ oxygen. fwdarw. hydroxyl+ oxygen atom using hydroxyl laser absorption," *Journal of Physical Chemistry*, vol. 94, no. 18, pp. 7119–7128, 1990.

[168] J. T. Herbon, R. K. Hanson, D. M. Golden, and C. T. Bowman, "A shock tube study of the enthalpy of formation of OH," *Proceedings of the Combustion Institute*, vol. 29, no. 1, pp. 1201–1208, 2002.

[169] A. N. Pirraglia, J. V. Michael, J. W. Sutherland, and R. B. Klemm, "A flash photolysis-shock tube kinetic study of the H atom reaction with O_2: H+ O_2 = OH + O 962 K <T <1705 K and H + O_2+ Ar= HO_2 + Ar (746 K <T <987 K)," *Journal of Physical Chemistry*, vol. 93, no. 1, pp. 282–291, 1989.

[170] S. Wang, D. F. Davidson, and R. K. Hanson, "Shock tube and laser absorption study of CH_2O oxidation via simultaneous measurements of OH and CO," *The Journal of Physical Chemistry A*, vol. 121, no. 45, pp. 8561–8568, 2017.

[171] J. V. Michael and J. W. Sutherland, "Rate constant for the reaction of H with H_2O and OH with H_2 by the flash photolysis-shock tube technique over the temperature range 1246-2297 k," *Journal of Physical Chemistry*, vol. 92, no. 13, pp. 3853–3857, 1988.

[172] L. N. Krasnoperov and J. V. Michael, "Shock tube studies using a novel multipass absorption cell: Rate constant results for OH + H_2 and OH + C_2H_6," *The Journal of Physical Chemistry A*, vol. 108, no. 26, pp. 5643–5648, 2004.

[173] J. W. Sutherland, J. V. Michael, A. N. Pirraglia, F. L. Nesbitt, and R. B. Klemm, "Rate constant for the reaction of O (3P) with H_2 by the flash photolysis—shock tube and flash photolysis—resonance fluorescence techniques; 504 K <T <2495 K," vol. 21, no. 1, pp. 929–941, 1988.

[174] K-Y Lam, D. F. Davidson, and R. K. Hanson, "A shock tube study of H_2+ OH = H_2O + H using OH laser absorption," *International Journal of Chemical Kinetics*, vol. 45, no. 6, pp. 363–373, 2013.

[175] J. V. Michael, J. W. Sutherland, L. B. Harding, and A. F. Wagner, "Initiation in H_2/O_2: Rate constants for H_2 + O_2 yields H + HO_2 at high temperature," Tech. Rep. 2, 2000, pp. 1471–1478.

[176] R. W. Bates, D. M. Golden, R. K. Hanson, and C. T. Bowman, "Experimental study and modeling of the reaction H + O_2 + M = HO_2 + M (M= Ar, N_2, H_2O) at elevated pressures

and temperatures between 1050 and 1250 K," *Physical Chemistry Chemical Physics*, vol. 3, no. 12, pp. 2337–2342, 2001.

[177] J. Michael, M.-C. Su, J. Sutherland, J. Carroll, and A. Wagner, "Rate constants for H + O_2 + M = HO_2 + M in seven bath gases," *The Journal of Physical Chemistry A*, vol. 106, no. 21, pp. 5297–5313, 2002.

[178] D. Davidson, E. Petersen, M Röhrig, R. Hanson, and C. Bowman, "Measurement of the rate coefficient of H + O_2 + M = HO_2 + M for M = Ar and N_2 at high pressures," in *Symposium (International) on Combustion*, Elsevier, vol. 26, 1996, pp. 481–488.

[179] M. A. Mueller, R. A. Yetter, and F. L. Dryer, "Measurement of the rate constant for H + O_2 + M = HO_2 + M (M = N_2, Ar) using kinetic modeling of the high-pressure $H_2/O_2/NO_x$ reaction," in *Symposium (International) on Combustion*, Elsevier, vol. 27, 1998, pp. 177–184.

[180] R. Choudhary, J. J. Girard, Y. Peng, J. Shao, D. F. Davidson, and R. K. Hanson, "Measurement of the reaction rate of H + O_2 + M = HO_2 + M, for M = Ar, N_2, CO_2, at high temperature with a sensitive OH absorption diagnostic," *Combustion and Flame*, vol. 203, pp. 265–278, 2019.

[181] F. M. Haas and F. L. Dryer, "Rate coefficient determinations for H + NO_2 = OH + NO from high pressure flow reactor measurements," *The Journal of Physical Chemistry A*, vol. 119, no. 28, pp. 7792–7801, 2015.

[182] M.-C. Su *et al.*, "Rate constants, 110 < T < 2000 K, for H + NO_2 = OH + NO using two shock tube techniques: Comparison of theory to experiment," *The Journal of Physical Chemistry A*, vol. 106, no. 36, pp. 8261–8270, 2002.

[183] M. A. Oehlschlaeger, D. F. Davidson, and R. K. Hanson, "High-temperature ethane and propane decomposition," *Proceedings of the Combustion Institute*, vol. 30, no. 1, pp. 1119–1127, 2005.

[184] M. Sangwan, C. Yan, E. N. Chesnokov, and L. N. Krasnoperov, "Reaction CH_3 + CH_3 = C_2H_6 studied over the 292–714 k temperature and 1–100 bar pressure ranges," *The Journal of Physical Chemistry A*, vol. 119, no. 28, pp. 7847–7857, 2015.

[185] D. F. Davidson, M. D. D. Rosa, E. J. Chang, R. K. Hanson, and C. T. Bowman, "A shock tube study of methyl-methyl reactions between 1200 and 2400 K," *International journal of chemical kinetics*, vol. 27, no. 12, pp. 1179–1196, 1995.

[186] K.-Y. Lam, D. F. Davidson, and R. K. Hanson, "High-temperature measurements of the reactions of oh with a series of ketones: Acetone, 2-butanone, 3-pentanone, and 2-pentanone," *The Journal of Physical Chemistry A*, vol. 116, no. 23, pp. 5549–5559, 2012.

[187] N. K. Srinivasan, M.-C. Su, and J. V. Michael, "High-temperature rate constants for CH_3OH + Kr = Products, OH + CH_3OH = products, OH +$(CH_3)_2CO$ = CH_2COCH_3 + h_2o, and OH + CH_3 = CH_2 + H_2O," *The Journal of Physical Chemistry A*, vol. 111, no. 19, pp. 3951–3958, 2007.

[188] P.-F. Lee, H. Matsui, D.-W. Xu, and N.-S. Wang, "Thermal decomposition and oxidation of ch3oh," *The Journal of Physical Chemistry A*, vol. 117, no. 3, pp. 525–534, 2013.

[189] V. Vasudevan, R. D. Cook, R. K. Hanson, C. T. Bowman, and D. M. Golden, "High-temperature shock tube study of the reactions CH_3 + OH = products and CH_3OH + Ar = products," *International Journal of Chemical Kinetics*, vol. 40, no. 8, pp. 488–495, 2008.

[190] E. A. Irdam, J. H. Kiefer, L. B. Harding, and A. F. Wagner, "The formaldehyde decomposition chain mechanism," *International journal of chemical kinetics*, vol. 25, no. 4, pp. 285–303, 1993.

[191] Y. Hidaka, T. Taniguchi, T. Kamesawa, H. Masaoka, K. Inami, and H. Kawano, "High temperature pyrolysis of formaldehyde in shock waves," *International journal of chemical kinetics*, vol. 25, no. 4, pp. 305–322, 1993.

[192] G. Friedrichs, D. F. Davidson, and R. K. Hanson, "Validation of a thermal decomposition mechanism of formaldehyde by detection of CH_2O and HCO behind shock waves," *International journal of chemical kinetics*, vol. 36, no. 3, pp. 157–169, 2004.

[193] S. Kumaran, J. Carroll, and J. Michael, "The branching ratio in the thermal decomposition of H_2CO," in *Symposium (International) on Combustion*, Elsevier, vol. 27, 1998, pp. 125–133.

[194] V. Vasudevan, D. F. Davidson, R. K. Hanson, C. T. Bowman, and D. M. Golden, "High-temperature measurements of the rates of the reactions CH_2O + Ar = products and CH_2O + O_2 = products," *Proceedings of the Combustion Institute*, vol. 31, no. 1, pp. 175–183, 2007.

[195] G. Friedrichs, J. T. Herbon, D. F. Davidson, and R. K. Hanson, "Quantitative detection of hco behind shock waves: The thermal decomposition of HCO," *Physical Chemistry Chemical Physics*, vol. 4, no. 23, pp. 5778–5788, 2002.

[196] N. Faßheber, G. Friedrichs, P. Marshall, and P. Glarborg, "Glyoxal oxidation mechanism: Implications for the reactions HCO + O_2 and OCHCHO + HO_2," *The Journal of Physical Chemistry A*, vol. 119, no. 28, pp. 7305–7315, 2015.

[197] S. Wang, E. E. Dames, D. F. Davidson, and R. K. Hanson, "Reaction rate constant of CH_2O + H = HCO + H_2 revisited: A combined study of direct shock tube measurement and transition state theory calculation," *The Journal of Physical Chemistry A*, vol. 118, no. 44, pp. 10 201–10 209, 2014.

[198] G. Friedrichs, D. F. Davidson, and R. K. Hanson, "Direct measurements of the reaction H + CH_2O = H_2 + HCO behind shock waves by means of vis–UV detection of formaldehyde," *International journal of chemical kinetics*, vol. 34, no. 6, pp. 374–386, 2002.

[199] S. Wang, D. F. Davidson, and R. K. Hanson, "High-temperature laser absorption diagnostics for CH_2O and CH_3CHO and their application to shock tube kinetic studies," *Combustion and Flame*, vol. 160, no. 10, pp. 1930–1938, 2013.

[200] S. Kumaran, M.-C. Su, K. Lim, and J. Michael, "The thermal decomposition of C_2H_5I," in *Symposium (International) on Combustion*, Elsevier, vol. 26, 1996, pp. 605–611.

[201] T. Bentz, M. Szőri, B. Viskolcz, and M. Olzmann, "Pyrolysis of ethyl iodide as hydrogen atom source: Kinetics and mechanism in the temperature range 950–1200 K," *Zeitschrift für Physikalische Chemie*, vol. 225, no. 9-10, pp. 1117–1128, 2011.

[202] T. Varga, I. G. Zsély, T. Turányi, T. Bentz, and M. Olzmann, "Kinetic analysis of ethyl iodide pyrolysis based on shock tube measurements," *International Journal of Chemical Kinetics*, vol. 46, no. 6, pp. 295–304, 2014.

[203] V. Vasudevan, D. F. Davidson, and R. K. Hanson, "Direct measurements of the reaction OH + CH_2O = HCO + H_2O at high temperatures," *International journal of chemical kinetics*, vol. 37, no. 2, pp. 98–109, 2005.

[204] M. S. Wooldridge, R. K. Hanson, and C. T. Bowman, "A shock tube study of the CO + OH = CO_2 + H reaction," in *Symposium (International) on Combustion*, Elsevier, vol. 25, 1994, pp. 741–748.

[205] J. Michael, S. Kumaran, and M.-C. Su, "Rate constants for CH_3 + O_2 = CH_3O + O at high temperature and evidence for H_2CO + O_2 = HCO+ HO_2," *The Journal of Physical Chemistry A*, vol. 103, no. 30, pp. 5942–5948, 1999.

[206] J. T. Herbon, R. K. Hanson, C. T. Bowman, and D. M. Golden, "The reaction of ch3+ o2: Experimental determination of the rate coefficients for the product channels at high temperatures," *Proceedings of the Combustion Institute*, vol. 30, no. 1, pp. 955–963, 2005.

[207] N. K. Srinivasan, M.-C. Su, J. W. Sutherland, and J. V. Michael, "Reflected shock tube studies of high-temperature rate constants for OH + CH_4 = CH_3 + H_2O and CH_3 + NO_2 = CH_3O + NO," *The Journal of Physical Chemistry A*, vol. 109, no. 9, pp. 1857–1863, 2005.

[208] N. Srinivasan, M.-C. Su, and J. Michael, "CH_3 + O_2 = H_2CO + OH revisited," *The Journal of Physical Chemistry A*, vol. 111, no. 45, pp. 11 589–11 591, 2007.

[209] H. Wang *et al.*, "USC mechanism version II. high-temperature combustion reaction model of H_2/CO/C1-C4 compounds.," 2007.

[210] L. B. Harding, S. J. Klippenstein, H. Lischka, and R. Shepard, "Comparison of multireference configuration interaction potential energy surfaces for H + O_2 = HO_2: The effect of internal contraction," *Thom H. Dunning, Jr. A Festschrift from Theoretical Chemistry Accounts*, pp. 115–121, 2015.

[211] S. J. Klippenstein, "From theoretical reaction dynamics to chemical modeling of combustion," *Proceedings of the Combustion Institute*, vol. 36, no. 1, pp. 77–111, 2017.

[212] J. A. Miller, M. J. Pilling, and J. Troe, "Unravelling combustion mechanisms through a quantitative understanding of elementary reactions," *Proceedings of the Combustion Institute*, vol. 30, no. 1, pp. 43–88, 2005.

[213] G. P. Smith *et al.*, *GRI-Mech3.0.*

[214] G. P. Smith *et al.*, *GRI-Mech1.1.*

[215] J Troe and V. Ushakov, "Theoretical studies of the HO + O = HO_2 = H + O_2 reaction. II. classical trajectory calculations on an ab initio potential for temperatures between 300 and 5000 K," *The Journal of Chemical Physics*, vol. 115, no. 8, pp. 3621–3628, 2001.

[216] T. L. Nguyen and J. F. Stanton, "Accurate ab initio thermal rate constants for reaction of O (3p) with H_2 and isotopic analogues," *The Journal of Physical Chemistry A*, vol. 118, no. 27, pp. 4918–4928, 2014.

[217] T. L. Nguyen, J. F. Stanton, and J. R. Barker, "Ab initio reaction rate constants computed using semiclassical transition-state theory: HO + H_2 = H_2O + H and isotopologues," *The Journal of Physical Chemistry A*, vol. 115, no. 20, pp. 5118–5126, 2011.

[218] N Balakrishnan, "Quantum calculations of the O (3 P)+ H_2 = OH + H reaction," *The Journal of chemical physics*, vol. 121, no. 13, pp. 6346–6352, 2004.

[219] R. R. Baldwin and R. W. Walker, "Rate constants for hydrogen+ oxygen system, and for h atoms and oh radicals+ alkanes," *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases*, vol. 75, pp. 140–154, 1979.

[220] C.-W. Zhou *et al.*, "A comprehensive experimental and modeling study of isobutene oxidation," *Combustion and Flame*, vol. 167, pp. 353–379, 2016.

[221] "Experimental and modelling study of the recombination reaction H + O_2 (+ M) = HO_2 (+ M) between 300 and 900 K, 1.5 and 950 bar, and in the bath gases M = He, Ar, and N_2, author = Fernandes, RX and Luther, K and Troe, J and Ushakov, VG, year = 2008, journal = Physical Chemistry Chemical Physics, publisher = Royal Society of Chemistry, volume = 10, number = 29, pages = 4313–4321,"

[222] M. A. Mueller, T. J. Kim, R. A. Yetter, and F. L. Dryer, "Flow reactor studies and kinetic modeling of the H_2/O_2 reaction," *International Journal of Chemical Kinetics*, vol. 31, no. 2, pp. 113–125, 1999.

[223] X. Yang, W. Liang, T. Tan, and C. K. Law, "Reevaluation of the reaction rate of H + O_2 (+ M) = HO_2 (+ M) at elevated pressures," *Combustion and Flame*, vol. 217, pp. 103–112, 2020.

[224] M Colberg and G Friedrichs, "Room temperature and shock tube study of the reaction HCO + O_2 using the photolysis of glyoxal as an efficient HCO source," *The Journal of Physical Chemistry A*, vol. 110, no. 1, pp. 160–170, 2006.

[225] A. Kéromnès *et al.*, "An experimental and detailed chemical kinetic modeling study of hydrogen and syngas mixture oxidation at elevated pressures," *Combustion and Flame*, vol. 160, no. 6, pp. 995–1011, 2013.

[226] J. Herzler and C. Naumann, "Shock-tube study of the ignition of methane/ethane/hydrogen mixtures with hydrogen contents from 0% to 100% at different pressures," *Proceedings of the combustion institute*, vol. 32, no. 1, pp. 213–220, 2009.

[227] Y. Hidaka *et al.*, "Shock-tube and modeling study of ethane pyrolysis and oxidation," *Combustion and flame*, vol. 120, no. 3, pp. 245–264, 2000.

[228] E. L. Petersen, D. F. Davidson, M Röhrig, and R Hanson, "High-pressure shock-tube measurements of ignition times in stoichiometric H_2/O_2/Ar mixtures," pp. 941–946, 1996.

[229] G. B. Skinner and G. H. Ringrose, "Ignition delays of a hydrogen—oxygen—argon mixture at relatively low temperatures," *The Journal of Chemical Physics*, vol. 42, no. 6, pp. 2190–2192, 1965.

[230] G. L. Schott and J. L. Kinsey, "Kinetic studies of hydroxyl radicals in shock waves. ii. induction times in the hydrogen-oxygen reaction," *The Journal of Chemical Physics*, vol. 29, no. 5, pp. 1177–1182, 1958.

[231] K. A. Bhaskaran, M. C. Gupta, and T. Just, "Shock tube study of the effect of unsymmetric dimethyl hydrazine on the ignition characteristics of hydrogen-air mixtures," *Combustion and Flame*, vol. 21, no. 1, pp. 45–48, 1973.

[232] R. K. Cheng and A. K. Oppenheim, "Autoignition in methane hydrogen mixtures," *Combustion and flame*, vol. 58, no. 2, pp. 125–139, 1984.

[233] M. W. Slack, "Rate coefficient for H + O_2 + M = HO_2 + M evaluated from shock tube measurements of induction times," *Combustion and Flame*, vol. 28, pp. 241–249, 1977.

[234] D. M. Kalitan, J. D. Mertens, M. W. Crofton, and E. L. Petersen, "Ignition and oxidation of lean CO/H_2 fuel blends in air," *Journal of propulsion and power*, vol. 23, no. 6, pp. 1291–1301, 2007.

[235] C. M. Vagelopoulos, F. N. Egolfopoulos, and C. K. Law, "Further considerations on the determination of laminar flame speeds with the counterflow twin-flame technique," in *Symposium (international) on combustion*, Elsevier, vol. 25, 1994, pp. 1341–1347.

[236] N. Lamoureux, N Djebaıli-Chaumeix, and C.-E. Paillard, "Laminar flame velocity determination for H_2–air–He–CO_2 mixtures using the spherical bomb method," *Experimental Thermal and Fluid Science*, vol. 27, no. 4, pp. 385–393, 2003.

[237] O. Kwon and G. Faeth, "Flame/stretch interactions of premixed hydrogen-fueled flames: Measurements and predictions," *Combustion and Flame*, vol. 124, no. 4, pp. 590–610, 2001.

[238] F. Egolfopoulos and C. K. Law, "An experimental and computational study of the burning rates of ultra-lean to moderately-rich $H_2/O_2/N_2$ laminar flames with pressure variations," in *Symposium (international) on combustion*, Elsevier, vol. 23, 1991, pp. 333–340.

[239] K. Aung, M. Hassan, and G. Faeth, "Flame stretch interactions of laminar premixed hydrogen/air flames at normal temperature and pressure," *Combustion and flame*, vol. 109, no. 1-2, pp. 1–24, 1997.

[240] I. C. McLean, D. B. Smith, and S. C. Taylor, "The use of carbon monoxide/hydrogen burning velocities to examine the rate of the co+ oh reaction," in *Symposium (international) on combustion*, Elsevier, vol. 25, 1994, pp. 749–757.

[241] T. Scholte and P. Vaags, "Burning velocities of mixtures of hydrogen, carbon monoxide and methane with air," *Combustion and Flame*, vol. 3, pp. 511–524, 1959.

[242] J Natarajan, T Lieuwen, and J Seitzman, "Laminar flame speeds of H_2/CO mixtures: Effect of CO_2 dilution, preheat temperature, and pressure," *Combustion and flame*, vol. 151, no. 1-2, pp. 104–119, 2007.

[243] H. Sun, S. Yang, G. Jomaas, and C. Law, "High-pressure laminar flame speeds and kinetic modeling of carbon monoxide/hydrogen combustion," *Proceedings of the Combustion Institute*, vol. 31, no. 1, pp. 439–446, 2007.

[244] D Bradley, M Lawes, K. Liu, S. Verhelst, and R Woolley, "Laminar burning velocities of lean hydrogen–air mixtures at pressures up to 1.0 mpa," *Combustion and Flame*, vol. 149, no. 1-2, pp. 162–172, 2007.

[245] M. P. Burke, M. Chaos, F. L. Dryer, and Y. Ju, "Negative pressure dependence of mass burning rates of H_2/CO/O_2/diluent flames at low flame temperatures," *Combustion and Flame*, vol. 157, no. 4, pp. 618–631, 2010.

[246] A. K. Das, K. Kumar, and C.-J. Sung, "Laminar flame speeds of moist syngas mixtures," *Combustion and Flame*, vol. 158, no. 2, pp. 345–353, 2011.

[247] C Prathap, A. Ray, and M. R. Ravi, "Investigation of nitrogen dilution effects on the laminar burning velocity and flame stability of syngas fuel at atmospheric condition," *Combustion and Flame*, vol. 155, no. 1-2, pp. 145–160, 2008.

[248] V. R. Kishore, M. Ravi, and A. Ray, "Adiabatic burning velocity and cellular flame characteristics of H_2–CO–CO_2–air mixtures," *Combustion and flame*, vol. 158, no. 11, pp. 2149–2164, 2011.

[249] L Qiao, Y Gu, W. Dahm, E. Oran, and G. Faeth, "A study of the effects of diluents on near-limit h2–air flames in microgravity at normal and reduced pressures," *Combustion and Flame*, vol. 151, no. 1-2, pp. 196–208, 2007.

[250] R. Hermanns, A. Konnov, R. Bastiaans, and L. De Goey, "Laminar burning velocities of diluted hydrogen- oxygen- nitrogen mixtures," *Energy & fuels*, vol. 21, no. 4, pp. 1977–1981, 2007.

[251] R. A. Yetter, F. Dryer, and H Rabitz, "A comprehensive reaction mechanism for carbon monoxide/hydrogen/oxygen kinetics," *Combustion Science and Technology*, vol. 79, no. 1-3, pp. 97–128, 1991.

[252] T. J. Kim, R. A. Yetter, and F. L. Dryer, "New results on moist co oxidation: High pressure, high temperature experiments and comprehensive kinetic modeling," in *Symposium (international) on combustion*, Elsevier, vol. 25, 1994, pp. 759–766.

Appendix A: Supporting Material

This appendix contains the supplemental / supporting materials referenced in Chapter 4 and Chapter 5.

Supplemental Material

Carly E. LaGrotta, Mark C. Barbet, Lei Lei, Michael P. Burke, "Towards a High-Accuracy Kinetic Database Informed by Theoretical and Experimental Data: CH_3 + HO_2 as a Case Study," *Proceedings of the Combustion Institute (2021).*

(a) Parametric Representations of Rate Constants

In practice, rather than calculating $k_{p,n}(T,P,M)$ for every T/P/M condition, it was convenient to represent the T/P/M dependence using some suitable m_n-term parametric representation

$$k_{p,n}(T,P,M) = \sum_{m=1}^{m_n} \xi_{n,m}(T,P,M)\hat{k}_{n,m}$$

(Eq. S1)

such as the three-term Arrhenius expressions used in [S1, S2], $k(T) = AT^b e^{-\frac{E}{RT}}$, where

$$\underline{\xi}_n(T,P,M) = \{1, lnT, T^{-1}\}$$
$$\underline{\hat{k}}_n = \{\ln A_n, b_n, -E_n/R\}$$

(Eq. S2)

are used for R1, R6, R7, and R8. For R2-R6, which are strongly non-Arrhenius, three additional parameters were used to ensure accurate representation over the full temperature range:

$$\underline{\xi}_n(T,P,M) = \{1, lnT, T^{-1}, -T^{-3}, -T, -T^3\}$$
$$\underline{\hat{k}}_n = \{\ln A_n, b_n, -\frac{E_n}{R}, c_n, d_n, e_n\}$$

(Eq. S3)

In this way, Eq. (5) in [S3] becomes

$$S_{i,j} = \frac{\partial \ln k_{p,n}(T_i,P_i,M_i)}{\partial x_j} = \sum_{m=1}^{m_n} \xi_{n,m}(T_i,P_i,M_i)\frac{\partial \hat{k}_{n,m}}{\partial x_j}$$

(Eq. S4)

and Eq. (6) in [S3] becomes

$$S_{i,j} = \sum_n \frac{\partial F_i}{\partial \ln k_{p,n}(T_i,P_i,M_i)} \frac{\partial \ln k_{p,n}(T_i,P_i,M_i)}{\partial x_j} = \sum_n \frac{\partial F_i}{\partial \ln k_{p,n}} \sum_{m=1}^{m_n} \xi_{n,m}(T_i,P_i,M_i)\frac{\partial \hat{k}_{n,m}}{\partial x_j}$$

(Eq. S5)

(b) Optimized Values and Uncertainties

Uncertainties in the optimization variables including only targets class I and IV, $\sigma_i(I, IV)$; differences between constrained model values and *a priori* values, $x_{j,opt} - x_{j,nom}$; and uncertainties in the optimization variables including all targets I, II, III and IV (constrained model), $\sigma_i(I, II, III, IV)$, are provided in Table S1.

Table S1: List of uncertainties in the optimized variables and differences between constrained model values and *a priori* values [a]

Parameter	$\sigma_i(I, IV)$	$x_{j,opt} - x_{j,nom}$	$\sigma_i(I, II, III, IV)$
$A'_{(1)}$	N/A	-0.006	N/A
$A'_{(8)}$	N/A	0.005	N/A
$n_{(1)}$	N/A	0.113	N/A
$n_{(8)}$	N/A	0.004	N/A
$E_{(1)}/R$	N/A	1119.6	N/A
$E_{(8)}/R$	N/A	-6.918	N/A
$E^{\dagger}_{(2)}$	3.000	-0.533	0.712
$\nu'_{all(2)}$	0.030	-0.008	0.029
$\nu'_{tr(2)}$	0.100	-0.080	0.075
$\nu'_{ss(2)}$	0.180	-0.039	0.176
$\nu'_{img(2)}$	0.180	-0.126	0.155
$E_{w(2)}$	1.00	-0.188	0.992
$\eta'_{H_2O_2}$	0.100	0.006	0.100
$\eta'_{TS(2)}$	0.262	-0.010	0.262
$E^{\dagger}_{(3)}$	2.300	-0.736	1.469
$\nu'_{all(3)}$	0.030	-0.001	0.030
$\nu'_{tr(3)}$	0.100	-0.009	0.097
$\nu'_{ss(3)}$	0.180	-0.003	0.180
$\nu'_{img(3)}$	0.180	0.005	0.180
$E_{w(3)}$	1.000	0.001	1.000
$\eta'_{TS(3)}$	0.405	-0.030	0.398
$E^{\dagger}_{(4g)}$	2.000	-0.981	0.796
$\nu'_{all(4)}$	0.030	0.002	0.030
$\nu'_{tr(4g)}$	0.100	0.014	0.087
$\nu'_{ss(4g)}$	0.180	0.010	0.176
$\nu'_{img(4g)}$	0.180	0.019	0.179
$E_{w(4g)}$	1.000	0.105	0.995
$\eta'_{TS(4g)}$	0.405	-0.027	0.397
$E^{\dagger}_{(4e)}$	2.000	-0.006	1.985
$\nu'_{TS(4e)}$	0.030	0.000	0.030

$\nu'_{tr(4e)}$	0.100	0.003	0.100
$\nu'_{ss(4e)}$	0.180	0.000	0.180
$\eta'_{TS(4e)}$	0.405	0.005	0.405
$f'_{VRCTST,c(4)}$	0.693	0.328	0.667
$E^{\dagger}_{(5g)}$	1.400	-1.411	0.487
$\nu'_{all(5)}$	0.030	-0.001	0.029
$\nu'_{tr(5g)}$	0.100	-0.015	0.086
$\nu'_{ss(5g)}$	0.180	-0.003	0.177
$\nu'_{img(5g)}$	0.180	0.028	0.172
$E_{w(5g)}$	1.000	0.044	0.980
$E^{\dagger}_{(5e)}$	1.400	-0.078	1.384
$\nu'_{TS(5e)}$	0.030	0.000	0.030
$\nu'_{tr(5e)}$	0.100	0.000	0.099
$\nu'_{ss(5e)}$	0.180	0.001	0.180
$E^{\dagger}_{(6)}$	2.300	0.405	0.949
$\nu'_{low(6)}$	0.100	0.005	0.097
$\left(Q_{anh,c,CH_3}\right)$	1.000	-0.160	0.955
$f'_{VRC,TST(7)}$	0.262	-0.067	0.229
$\left(Q_{anh,c,CH_3}\right)$	1.000	-0.160	0.955

	nominal	optimized	unit
T'_1	1398.000	1400.078	K
P'_1	1.909	1.906	atm
$M'_{H_2O,o,1}$	0.123	0.124	%
$M'_{H_2O_2,o,1}$	0.254	0.255	%
$M'_{O_2,o,1}$	0.062	0.062	%
T'_2	1192.000	1198.139	K
P'_2	1.950	1.957	atm
$M'_{H_2O_2,o,2}$	0.222	0.221	%
$M'_{H_2O,o,2}$	0.136	0.137	%
$M'_{O_2,o,2}$	0.068	0.068	%
T'_3	1057.000	1058.546	K
P'_3	1.830	1.830	atm
$M'_{H_2O_2,o,3}$	0.086	0.0863	%
$M'_{H_2O,o,3}$	0.066	0.067	%
$M'_{O_2,o,3}$	0.033	0.033	%
T'_4	1182.000	1172.289	K
P'_4	1.672	1.669	atm
$M'_{H_2O_2,o,4}$	0.205	0.205	%
$M'_{H_2O,o,4}$	0.111	0.111	%
$M'_{O_2,o,4}$	0.056	0.056	%

T_5'	0.010	0.007	0.008	T_5'	1880.000	1892.849	K
P_5'	0.020	0.001	0.020	P_5'	1.740	1.743	atm
$M'_{H_2O,o,5}$	0.100	0.033	0.095	$M'_{H_2O,o,5}$	1.300	1.342	%
$M'_{O_2,o,5}$	0.010	0.000	0.010	$M'_{O_2,o,5}$	0.990	0.990	%
$M'_{H,o,5}$	2.303	0.516	0.301	$M'_{H,o,5}$	0.700	1.173	ppm
T_6'	0.020	0.003	0.018	T_6'	968.000	970.800	K
P_6'	0.040	-0.001	0.039	P_6'	3.590	3.585	atm
$M'_{H_2O_2,o,6}$	0.100	-0.011	0.088	$M'_{H_2O_2,o,6}$	0.314	0.310	%
T_7'	0.020	0.014	0.011	T_7'	1015.000	1029.088	K
P_7'	0.040	0.001	0.039	P_7'	3.298	3.301	atm
$M'_{H_2O_2,o,7}$	0.100	-0.008	0.089	$M'_{H_2O_2,o,7}$	0.326	0.324	%
T_8'	0.020	0.003	0.009	T_8'	1079.000	1081.702	K
P_8'	0.040	0.003	0.039	P_8'	3.442	3.452	atm
$M'_{H_2O_2,o,8}$	0.100	0.015	0.088	$M'_{H_2O_2,o,8}$	0.195	0.198	%
T_9'	0.020	-0.011	0.008	T_9'	1132.000	1119.379	K
P_9'	0.040	-0.001	0.039	P_9'	3.550	3.548	atm
$M'_{H_2O_2,o,9}$	0.100	0.008	0.089	$M'_{H_2O_2,o,9}$	0.136	0.137	%
T_{10}'	0.020	-0.009	0.009	T_{10}'	1218.000	1207.016	K
P_{10}'	0.040	-0.002	0.039	P_{10}'	3.038	3.030	atm
$M'_{H_2O_2,o,10}$	0.100	-0.003	0.089	$M'_{H_2O_2,o,10}$	0.171	0.170	%
T_{11}'	0.010	-0.002	0.006	T_{11}'	1103.000	1101.202	K
P_{11}'	0.020	-0.002	0.020	P_{11}'	3.560	3.554	atm
$M'_{H_2O_2,o,11}$	0.050	0.019	0.042	$M'_{H_2O_2,o,11}$	0.124	0.126	%
$M'_{H_2O,o,11}$	0.050	0.015	0.047	$M'_{H_2O,o,11}$	0.082	0.083	%
$M'_{O_2,o,11}$	0.050	0.000	0.050	$M'_{O_2,o,11}$	0.041	0.041	%
$M'_{CH_4,o,11}$	0.050	-0.001	0.050	$M'_{CH_4,o,11}$	0.998	0.996	%
$\widehat{\sigma'}_{1,H_2O_2}$	0.700	-0.098	0.075	$\widehat{\sigma'}_{1,H_2O_2}$	140000.000	126935.115	cm^2mol^{-1}
$\widehat{\sigma'}_{1,HO_2}$	0.700	0.003	0.202	$\widehat{\sigma'}_{1,HO_2}$	1270000.000	1265925.696	cm^2mol^{-1}
σ'_{1,H_2O_2}	0.700	0.094	0.133	σ'_{1,H_2O_2}	96000.000	105473.500	cm^2mol^{-1}
σ'_{2,H_2O_2}	0.300	-0.007	0.299	σ'_{2,H_2O_2}	-19.000	-18.875	$cm^2mol^{-1}K^{-1}$
σ'_{1,HO_2}	0.700	-0.201	0.305	σ'_{1,HO_2}	1155000.000	944466.901	cm^2mol^{-1}
σ'_{2,HO_2}	0.300	-0.019	0.293	σ'_{2,HO_2}	1299.000	1275.112	K

[a] Note that any parameters in () are represented by only a single value but influence multiple reactions

(c) Additional Figures

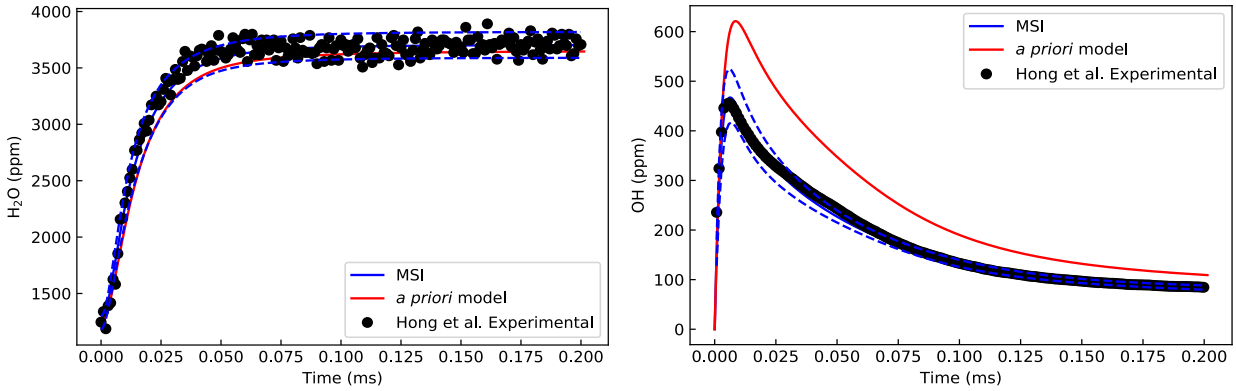


Fig. S1. H_2O (left) and OH (right) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1398 K and 1.909 atm. Symbols represent experimental data from Hong et al. [S4]; lines model predictions.

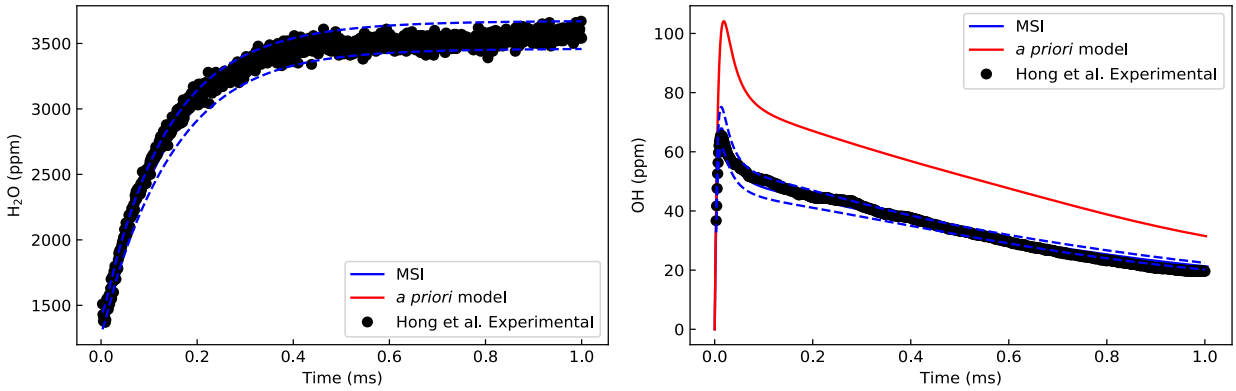


Fig. S2. H_2O (left) and OH (right) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1192 K and 1.950 atm. Symbols represent experimental data from Hong et al. [S4]; lines model predictions.

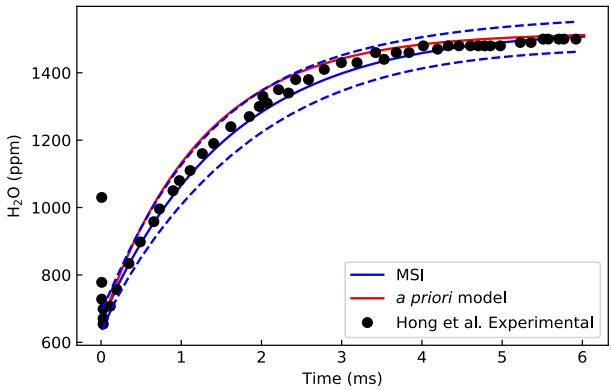


Fig. S3. H_2O time-history in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1057 K and 1.830 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

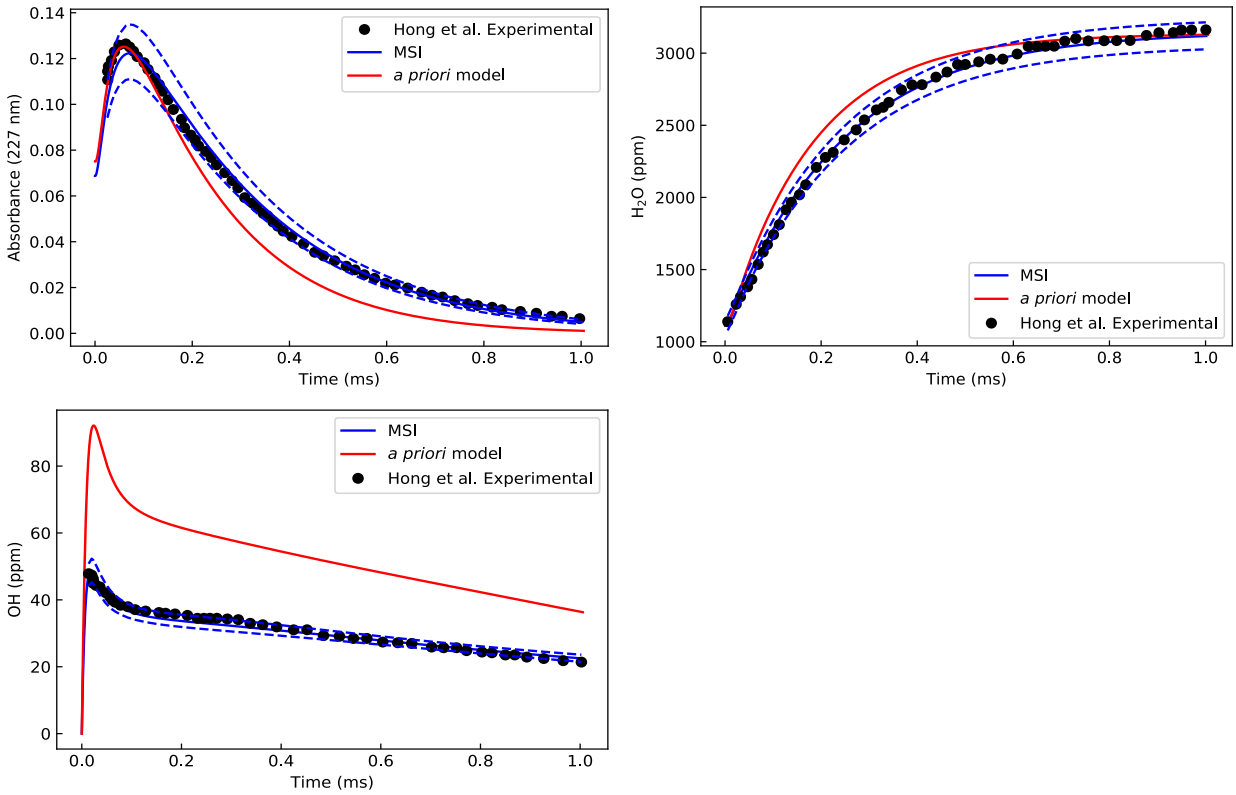


Fig. S4. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1182 K and 1.672 atm. Symbols represent experimental data from Hong et al. [S6]; lines model predictions.

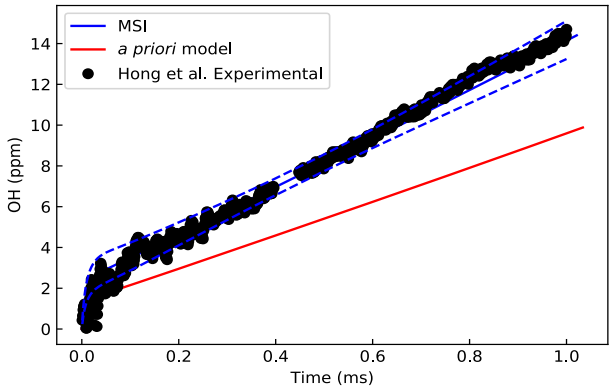


Fig. S5. OH time-history in shock heated $H_2O/O_2/H/Ar$ mixture near 1880 K and 1.740 atm. Symbols represent experimental data from Hong et al. [S7]; lines model predictions.

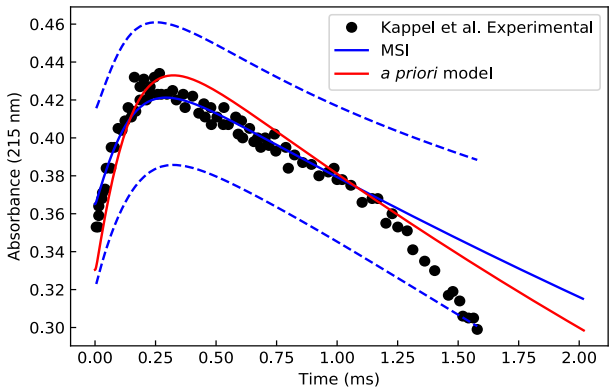


Fig. S6. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 968 K and 3.590 atm. Symbols represent experimental data from Kappel et al. [S8]; lines model predictions.

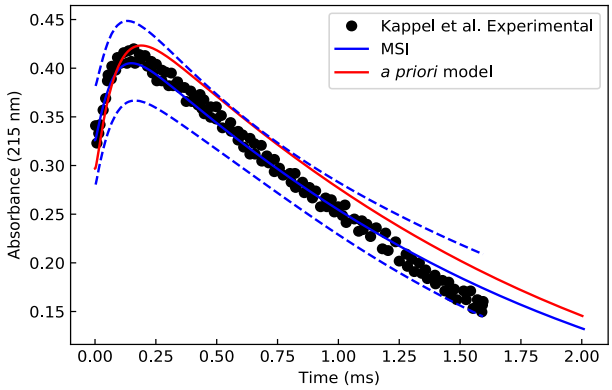


Fig. S7. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1015 K and 3.298 atm. Symbols represent experimental data from Kappel et al. [S8]; lines model predictions.

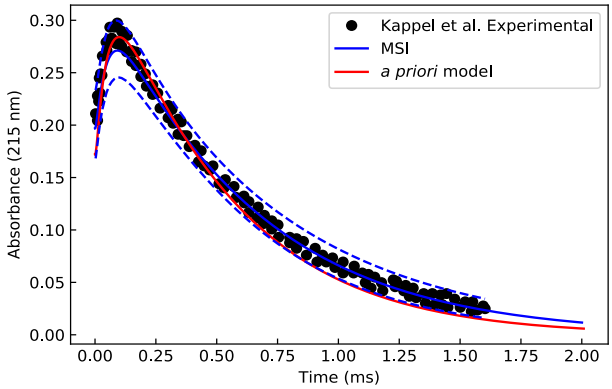


Fig. S8. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1079 K and 3.441 atm. Symbols represent experimental data from Kappel et al. [S8]; lines model predictions.

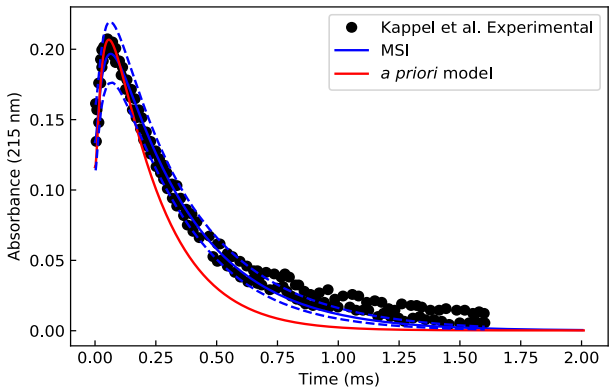


Fig. S9. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1132 K and 3.550 atm. Symbols represent experimental data from Kappel et al. [S8]; lines model predictions.

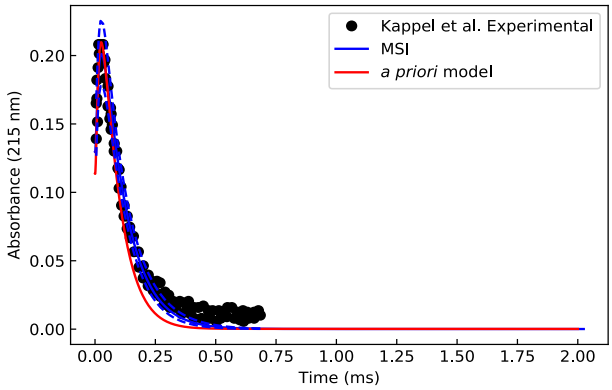


Fig. S10. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1218 K and 3.038 atm. Symbols represent experimental data from Kappel et al. [S8]; lines model predictions.

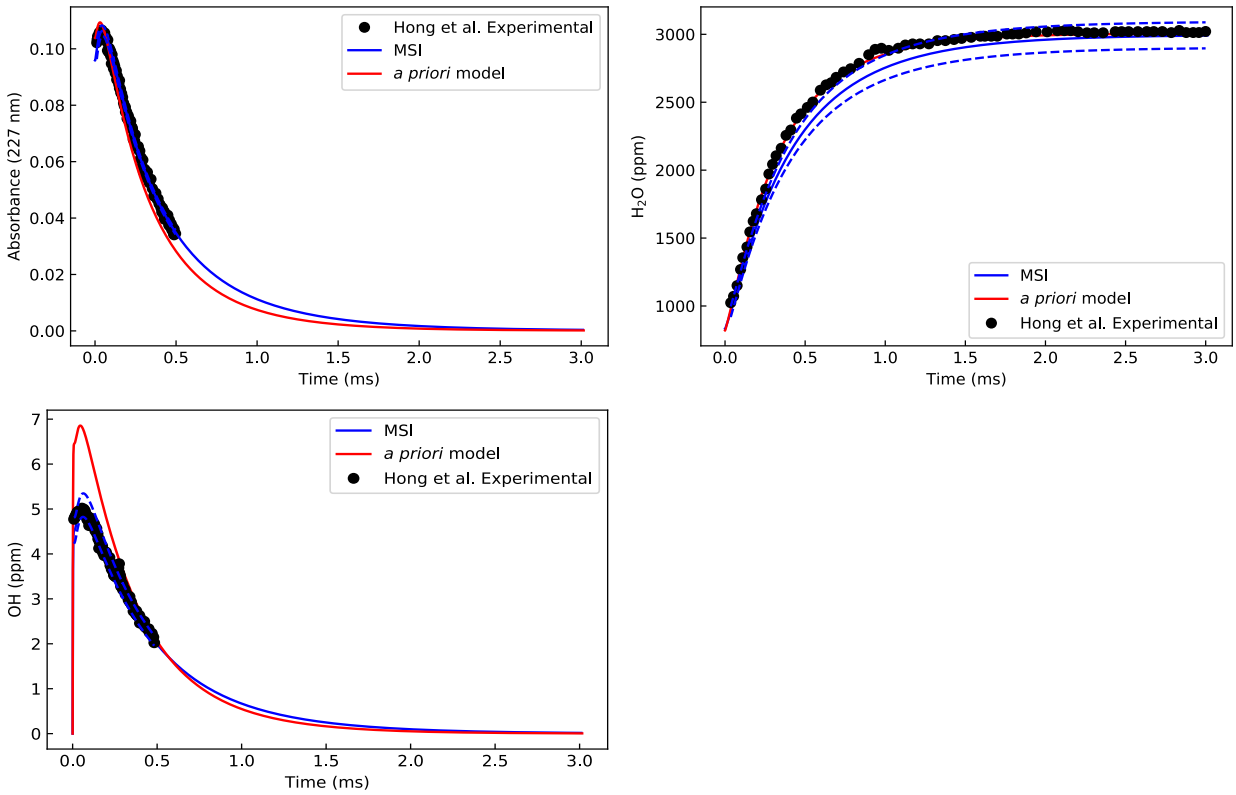


Fig. S11. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/CH_4/Ar$ mixture near 1103 K and 3.560 atm. Symbols represent experimental data from Hong et al. [S9]; lines model predictions.

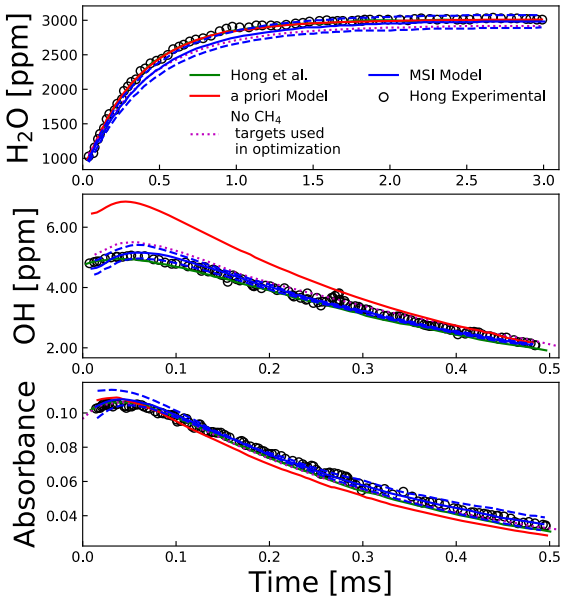


Fig. S12. H_2O (top) and OH (middle) and Absorbance 227 nm (bottom), time-histories in shock heated $H_2O_2/H_2O/O_2/CH_4/Ar$ mixture near 1103 K and 3.560 atm. Symbols represent experimental data from Hong et al. [S9]; lines model predictions. This figure is similar to that of Fig. 2 included in the main text, with the addition of a line to show the model predictions if CH_4/H_2O_2 experiments are not included as targets. As discussed in the main text the ability of the model to predict these time profiles so accurately, without including the CH_4/H_2O_2 experiments as targets, highlights the advantage of leveraging data across multiple sources and constraining the model with theory.

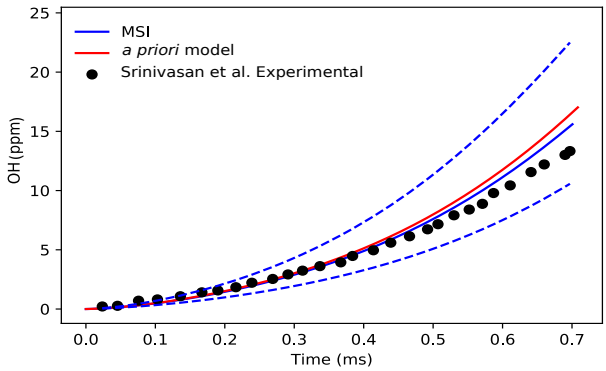


Fig. S13. OH time-history in shock heated $O_2/CH_4/Kr$ mixture near 1795 K and 0.948 atm. Symbols represent experimental data from Srinivasan et al. [S10]; lines model predictions. This data was not used as target data, but rather was just used to test the model.

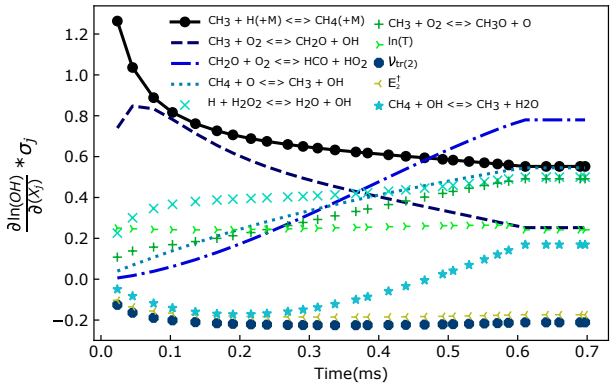


Fig. S14. Uncertainty weighted sensitivity analysis for OH, in shock heated $O_2/CH_4/Kr$ mixture near 1795 K and 0.948 atm [S10]. Here it can be noted that for each chemical reaction displayed in the legend the parameter being described (x_j) is the pre-exponential factor for that reaction. This experiment was used to determine the rate constant of, $CH_3 + HO_2 \Leftrightarrow CH_4 + O_2$, and exhibits some sensitivity to parameters describing this rate constant, but the largest contributors to uncertainty for this observable are parameters describing other reactions.

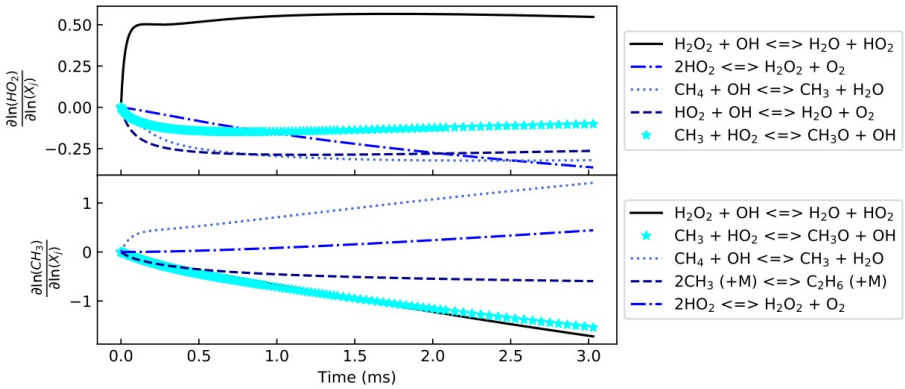


Fig. S15. Sensitivity analysis for HO_2 and CH_3 performed using the GRI30 Mechanism [S11] and $CH_3/HO_2/OH/H_2O_2/CH_4/H_2O/N_2O/Ar$ mixture near 295 K and 0.986 atm. [S12]. Here it can be noted that for each chemical reaction displayed in the legend the parameter being described (x_j) is the pre-exponential factor for that reaction. This experiment was used to determine rate constants for reactions of $CH_3 + HO_2$, and exhibits some sensitivity to parameters describing these rate constants, but the most sensitive parameters for these observables mainly seem to be parameters describing other reactions.

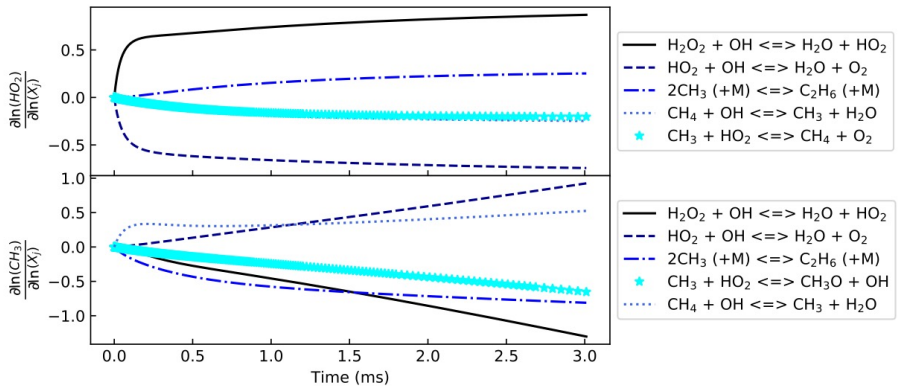


Fig. S16. Sensitivity analysis for HO_2 and CH_3 performed using the FFCM Mechanism [S13] and $CH_3/HO_2/OH/H_2O_2/CH_4/H_2O$/He mixture near 295 K and 0.986 atm. [S12]. Here it can be noted that for each chemical reaction displayed in the legend the parameter being described (x_j) is the pre-exponential factor for that reaction. This experiment was used to determine rate constants for reactions of CH_3 + HO_2, and exhibits some sensitivity to parameters describing these rate constants, but the most sensitive parameters for these observables mainly seem to be parameters describing other reactions.

Fig. S15 and Fig. S16 helped the authors make the decision to exclude the low temperature datasets presented in [S12]. To simulate the low temperature system with nitrogen containing species the GRI30 mechanism had to be used in order to avoid convergence issues that were occurring with other mechanisms. The simulation was also repeated using the FFCM mechanism, disregarding the nitrogen containing species, in order to help verify the results received from the simulation which used the GRI30 mechanism.

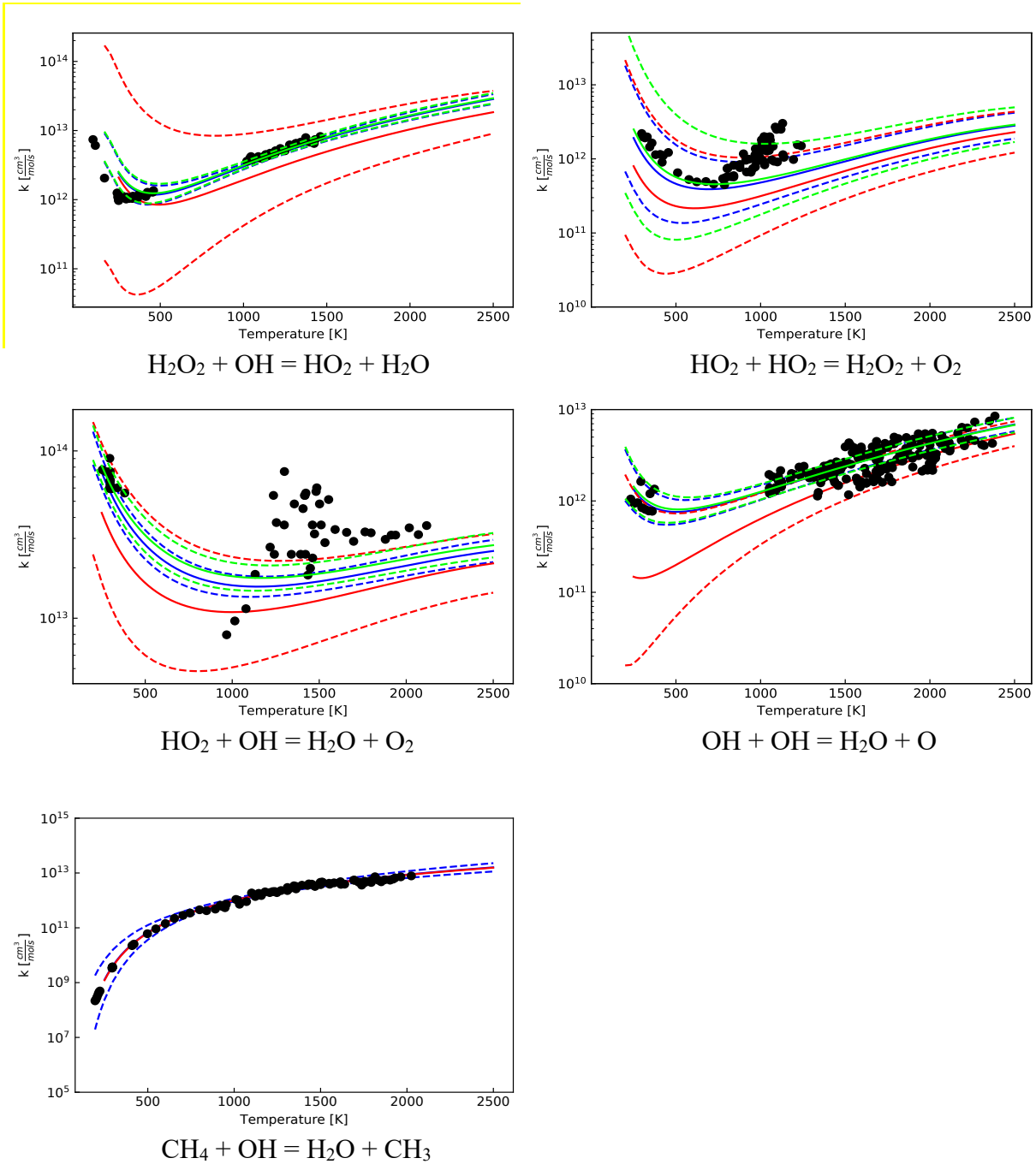


Fig. S17. Comparisons of model predictions (red lines: *a priori* model with error bars constrained by target class I only; green lines: original results from Burke et al. study[S3]; blue lines: constrained model with error bars constrained by target classes I-IV) and experimental data (symbols) for rate constants. Top left: R2 [S4, S14-S17], Top Right: R3 [S8, S18-S20], Middle Left: R4[S7-S8, S21-S28], Middle Right: R5 [S29-S33], and Bottom Left: R8 [S34-S35].

(d) Optimized Rate Constant Results

Modified Arrhenius Expression: $k = AT^n e^{-\frac{Ea}{RT}}$

	A (cm3/mol s)	n	Ea (cal/mol)
$CH_3 + HO_2 <=> CH_4 + O_2$	3.27356×10^3	2.664	-3678.322
$CH_3 + HO_2 <=> CH_3O + OH$	8.88397×10^{11}	0.288	-779.323

S1. T. Nagy, T. Turányi, Uncertainty of Arrhenius parameters, Int. J. Chem. Kinet. 43 (2011) 359-378.

S2. T. Turányi, T. Nagy, I.Gy. Zsély, et al., Determination of rate parameters based on both direct and indirect measurements, Int. J.Chem. Kinet. 44 (2012) 284–302.

S3. M.P. Burke, S.J. Klippenstein, L.B. Harding, A quantitative explanation for the apparent anomalous temperature dependence of $OH+HO_2=H_2O+O_2$ through multi-scale modeling, Proc. Combust. Inst. 34 (2013) 547-555.

S4. Z. Hong, R. D. Cook, D. F. Davidson, R. K. Hanson, A shock tube study of $OH + H_2O_2 = H_2O + HO_2$ and $HO_2 + M = 2OH + M$ using laser absorption of H_2O and OH, J. Phys. Chem. A 114 (2010) 5718–5727.

S5. Z. Hong, A. Farooq, E. A. Barbour, D. F. Davidson, R. K. Hanson, Hydrogen peroxide decomposition rate: a shock tube study using tunable laser absorption of H_2O near 2.5 micro meters, J. Phys. Chem. A 113 (2009) 12919–12925.

S6. Z. Hong, K-Y. Lam, R. Sur, S. Wang, D. Davidson, R. Hanson, On the rate constants of $OH + HO_2$ and HO_2+HO_2: A comprehensive study of H_2O_2 thermal decomposition using multispecies laser absorption, Proc. Combust. Inst. 34 (2013) 565–571.

S7. Z. Hong, S. S. Vasu, D. F. Davidson, R. K. Hanson, Experimental study of the rate of $OH + HO_2 = H_2O + O2$ at high temperatures using the reverse reaction, J. Phys. Chem. A 114 (2010) 5520–5525.

S8. C. Kappel, K. Luther, J. Troe, Shock wave study of the unimolecular dissociation of H_2O_2 in its falloff range and of its secondary reactions, Phys. Chem. Chem. Phys. 4 (2002) 4392–4398.

S9. Z. Hong, D. F. Davidson, K.-Y. Lam, R. K. Hanson, A shock tube study of the rate constants of HO_2 and CH_3 reactions, Combust. Flame 159 (2012) 3007–3013.

S10. Srinivasan, N. K., et al., Experimental and theoretical rate constants for $CH_4+ O_2 \rightarrow CH_3+HO_2$, Combust. Flame 149 (2007) 104-111.

S11. G.P. Smith, D.M. Golden, M. Frenklach, N.W. Moriatry, B. Eiteneer, M. Goldenberg, C.T. Bowman, R.K. Hanson, S. Song, W.C. Gardiner, V.V. Lissianski, and Z. Qin: http://www.me.berkeley.edu/gri_mech/.

S12. M. Sangwan, L. N. Krasnoperov, Kinetics of the gas phase reaction $CH_3 + HO_2$, J. Phys. Chem. A 117 (2013) 2916–2923.

S13. G. Smith, Y. Tao, H.Wang, Foundational Fuel Chemistry Model Version 1.0 (FFCM-1) (2016).

S14. U.C. Sridharan, B. Reimann, F. Kaufman, Kinetics of the reaction $OH + H_2O_2 \rightarrow HO_2 + H_2O$, J. Chem. Phys. 73 (1980) 1286–1293.

S15. L.F. Keyser, Absolute rate constant of the reaction $OH + H_2O_2 \rightarrow HO_2 + H_2O$ from 245 to 423 K, J. Phys. Chem. 84 (1980) 1659–1663.

S16. J.J. Lamb, L. Moiina, C.A. Smith, M.J. Molina, Rate constant of the hydroxy radical+ hydrogen peroxide→ hydroperoxo radical+ water reaction, J. Phys. Chem. 87 (1983) 4467-4470.

S17. A.B. Vakhtin, D.C. McCabe, A.R. Ravishankara, S.R. Leone, Low-temperature kinetics of the reaction of the OH radical with hydrogen peroxide, J. Phys. Chem. A 107 (2003) 10642-10647.

S18. R. Patrick, M.J. Pilling, The temperature dependence of the $HO_2 + HO_2$ reaction, Chem. Phys. Lett. 91 (1982) 343-347.

S19. P.D. Lightfoot, B. Veyret, R. Lesclaux, The rate constant for the $HO_2 + HO_2$ reaction at elevated temperatures, Chem. Phys. Lett. 150 (1988) 120-126.

S20. H. Hippler, J. Troe, J. Willner, Confirmation of a rate constant minimum near 700 K, J. Chem. Phys. 93 (1990) 1755-1760.

S21. W.B. DeMore, Reaction of hydroperoxo radicals with ozone and the effect of water vapor on hydroperoxo kinetics, J. Phys. Chem. 83 (1979) 1113-1118.

S22. W.B. DeMore, Rate constant and possible pressure dependence of the reaction hydroxyl + hydroperoxo, J. Phys. Chem. 86 (1982) 121–126.

S23. L.F. Keyser, Absolute rate constant of the reaction $OH + H_2O_2 \rightarrow HO_2 + H_2O$ from 245 to 423 K, J. Chem. Phys. 84 (1980) 1659–1663.

S24. R.R. Lii, R.A. Gorse, M.C. Sauer, S. Gordon, Temperature dependence of the gas-phase self-reaction of hydroperoxo radicals in the presence of ammonia, J. Phys. Chem. 84 (1980) 813–817.

S25. R.A. Cox, J.P. Burrows, T.J. Wallington, Rate coefficient for the reaction OH+ HO_2= H_2O+ O_2 at 1 atmosphere pressure and 308 K, Chem. Phys. Lett. 84 (1981) 217–221.

S26. M.J. Kurylo, O. Klais, A.H. Laufer, Mechanistic investigation of the hydroxyl+ hydroperoxo reaction, J. Phys. Chem. 85 (1981) 3674–3678.

S27. M. Braun, A. Hofzumahaus, F. Stuhl, Ber. Bunsenges, VUV Flash Photolysis Study of the Reaction of HO with HO_2 at 1 atm and 298 K, J. Phys. Chem. 86 (1982) 597–602.

S28. N.K. Srinivasan, M.C. Su, J.W. Sutherland, J.V.Michael, B. Ruscic, Reflected shock tube studies of high-temperature rate constants for OH+ NO_2→ HO_2+ NO and OH+ HO_2→ H_2O+ O_2, J. Phys. Chem. A 110 (2006) 6602–6607.

S29. M.K. Bahng, R.G. Macdonald, Determination of the Rate Constant for the OH (X2Π)+ OH (X2Π)→ O (3P)+ H_2O Reaction over the Temperature Range 293− 373 K, J. Phys. Chem. A 111 (2007) 3850–3861.

S30. Y. Bedjanian, G. Le Bras, G. Poulet, Kinetic study of OH+ OH and OD+ OD reactions, J. Phys. Chem. A 103 (1999) 7017–7025.

S31. M.S. Wooldridge, R.K. Hanson, C.T. Bowman, A shock tube study of the OH + OH→ H_2O + O reaction, Int. J. Chem. Kinet. 26 (1994) 389–401.

S32. J.W. Sutherland, P.M. Patterson, R.B. Klemm, Rate constants for the reaction, O (3P)+ H_2O¶ OH+ OH, over the temperature range 1053 K to 2033 K using two direct techniques, Proc. Combust. Inst. 23 (1990) 51–57.

S33. A. Lifshitz, J.V. Michael, Rate constants for the reaction, O+ H_2O→ OH+ OH, over the temperature range, 1500–2400 K, by the flash photolysis-shock tube technique: A further consideration of the back reaction, Proc. Combust. Inst. 23 (1991) 59-67.

S34. T. Gierczak, R.K. Talukdar, S.C. Herndon, G.L. Vaghjiani, A. Ravishankara, Rate coefficients for the reactions of hydroxyl radicals with methane and deuterated methanes, J. Phys. Chem. A 101 (1997) 3125-3134.

S35. J.R. Dunlop, F.P. Tully, A kinetic study of hydroxyl radical reactions with methane and perdeuterated methane, J. Phys. Chem. 97 (1993) 1148-11150.

Supporting Information

"Resolving Discrepancies Between State-of-the-art Theory and Experiment for HO_2 + HO_2 via Multiscale Informatics" *Journal of Physical Chemistry A (2022).*

Carly E. LaGrotta[†], Qinghui Meng[†], Lei Lei[†], Mark C. Barbet[†], Zekai Hong[‡], Michael P. Burke[†,¶,§]

† Department of Mechanical Engineering, Columbia University, New York, NY
‡ Aerospace Research Center, National Research Council of Canada, Ottawa, Canada
¶ Department of Chemical Engineering, Columbia University, New York, NY
§ Data Science Institute, Columbia University, New York, NY
Email: mpburke@columbia.edu

Table of Contents

(a) Optimized Values and Uncertainties

Uncertainties in the active parameters including only target classes I and IV, $\sigma_i(I, IV)$; differences between MSI values and *a priori* values, $x_{j,opt} - x_{j,nom}$; and uncertainties in the active parameters including all targets I, II, III and IV (MSI model), $\sigma_i(I, II, III, IV)$, are provided in Table S1.

Table S1: List of uncertainties in the active parameters and differences between MSI values and *a priori* values

Parameter	$\sigma_i(I, IV)$	$x_{j,opt} - x_{j,nom}$	$\sigma_i(I, II, III, IV)$
$A'_{(7)}$	N/A	-0.2035	N/A
$n_{(7)}$	N/A	0.017	N/A
$E_{(7)}/R$	N/A	0.00018	N/A
$\nu'_{B1(1)}$	0.045	0.000	0.045
$E_{B1(1)}$	0.200	-0.018	0.186
$\nu'_{W1a(1)}$	0.135	-0.005	0.134
$E_{W1a(1)}$	0.200	-0.003	0.200
$\rho'_{W1b(1)}$	1.610	-0.020	1.607
$\rho'_{W1a(1)}$	1.610	-0.175	1.439
$f'_{VRC-TST,cB1(1)}$	0.693	-0.019	0.689
$\sigma_{z'_{(1)}}$	1.100	0.150	0.916
$E_{W1b(1)}$	0.200	-0.00018	0.200
$\nu'_{img\ B2(1)}$	0.180	0.032	0.159
$\nu'_{P1(1)}$	0.045	-0.0006	0.045
$\nu'_{P2(1)}$	0.045	0.000	0.045
$\nu'_{B8(1)}$	0.135	0.002	0.135
$n_{\Delta E_d(1)}$	0.200	0.001	0.200
$A'_{\Delta E_d(1)}$	1.100	0.088	0.964
$\rho'_{B8(1)}$	1.610	0.015	1.601
$E_{P2(1)}$	0.200	0.000	0.200
$\nu'_{W1b(1)}$	0.135	-0.0006	0.135
$E_{B8(1)}$	0.200	-0.0005	0.200
$E_{B2(1)}$	0.200	0.000	0.200
$E_{P1(1)}$	0.200	0.034	0.172
$\eta'_{B2(1)}$	0.070	0.002	0.070

$\nu'_{B2(1)}$	0.045	0.002	0.044
$E_{B4(2)}$	0.400	0.002	0.395
$\nu'_{W2b(2)}$	0.225	0.000	0.225
$\sigma_{z'_{(2)}}$	1.100	-0.016	1.057
$\nu'_{B7_1(2)}$	0.045	0.000	0.045
$\nu'_{B9(2)}$	0.225	0.006	0.224
$f'_{VRC-TST,cB5(2)}$	0.693	0.052	0.670
$\nu'_{B4(2)}$	0.045	0.000	0.045
$f'_{VRC-TST,cB7_1(2)}$	0.693	0.016	0.692
$\nu'_{W2a(2)}$	0.225	0.0003	0.225
$E_{W2b(2)}$	1.000	0.000	1.000
$\nu'_{img\ B3(2)}$	0.180	0.000	0.180
$E_{B3(2)}$	2.000	0.002	2.000
$E_{B7_2(2)}$	1.000	0.0002	1.000
$E_{B9(2)}$	1.000	0.008	1.000
$E_{P4(2)}$	2.000	0.000	2.000
$\nu'_{P3(2)}$	0.045	0.000	0.045
$\rho'_{corr\ B9(2)}$	1.610	0.044	1.605
$\nu'_{B3(2)}$	0.045	0.000	0.045
$E_{B6(2)}$	2.000	-0.0005	2.000
$\nu'_{B6(2)}$	0.045	0.000	0.045
$f'_{VRC-TST,c\ B4(2)}$	0.693	0.002	0.693
$\nu'_{P5(2)}$	0.045	0.000	0.045
$E_{W3(2)}$	0.400	0.006	0.398
$\nu'_{B5(2)}$	0.075	0.001	0.075
$E_{P3(2)}$	0.2	0.000	0.200
$\nu'_{P1(2)}$	0.075	-0.0004	0.075
$\rho'_{corr\ P5(2)}$	0.690	0.000	0.693
$n_{\Delta E_d(2)}$	0.200	0.000	0.200
$A'_{\Delta E_d(2)}$	1.100	-0.018	1.053
$\nu'_{P4(2)}$	0.075	0.000	0.075
$\nu'_{W3(2)}$	0.045	0.000	0.045
$E_{P5(2)}$	0.4	0.000	0.400
$E_{B5(2)}$	1.000	0.051	0.989
$\nu'_{img\ B6(2)}$	0.180	0.000	0.180
$E_{W2a(2)}$	1.000	0.002	0.999
$E_{P1(2)}$	0.200	-0.006	0.199
$\eta'_{B7_1(2)}$	0.070	0.000	0.070

$\eta'_{B7_2(2)}$	0.070	0.000	0.070
$\eta'_{W3_1(2)}$	0.070	0.000	0.070
$\eta'_{W3_2(2)}$	0.070	0.000	0.070
$\eta'_{W3_3(2)}$	0.070	0.000	0.070
$\eta'_{P5(2)}$	0.070	0.000	0.070
$E_{B7_1(2)}$	0.400	0.006	0.400
$f'_{VRC-TST,c\ B7_2(2)}$	0.693	0.0001	0.693
$\nu'_{B7_2(2)}$	0.075	0.000	0.075
$E_{B1(3)}$	3.000	-0.654	0.693
$\nu'_{B1(3)}$	0.100	-0.079	0.044
$\eta'_{B1(3)}$	0.100	-0.006	0.099
$\nu'_{img\ B1(3)}$	0.180	-0.138	0.166
$E_{B1(4)}$	3.000	-1.364	0.430
$E_{B2(4)}$	3.000	0.000	3.000
$\nu'_{B1,B2(4)}$	0.1000	-0.020	0.072
$\nu'_{img\ B1,B2(4)}$	0.180	-0.005	0.177
$f'_{VRC-TST,c\ B1\ (5)}$	0.700	0.168	0.683
$E_{B2(5)}$	2.000	-1.107	0.599
$\nu'_{B2(5)}$	0.100	-0.018	0.065
$\eta'_{B2(5)}$	0.410	-0.011	0.403
$\nu'_{img\ B2(5)}$	0.180	0.001	0.179
$E_{B3(5)}$	2.000	0.395	1.975
$\nu'_{B3(5)}$	0.100	0.008	0.098
$\eta'_{B3(5)}$	0.410	0.032	0.409
$\nu'_{img\ B3(5)}$	0.180	-0.001	0.179
$E_{W1(5)}$	$1.00\text{x}10^{-06}$	0.000	$1.00\text{x}10^{-06}$
$\nu'_{W1(5)}$	0.010	0.000	0.010

					nominal	optimized	unit
T'_1	0.010	0.008	0.009	T'_1	1398.000	1409.661	K
P'_1	0.020	-0.005	0.020	P'_1	1.909	1.900	atm
$M'_{H_2O,o,1}$	0.050	0.001	0.048	$M'_{H_2O,o,1}$	0.123	0.124	%
$M'_{H_2O_2,o,1}$	0.050	0.005	0.042	$M'_{H_2O_2,o,1}$	0.254	0.255	%
$M'_{O_2,o,1}$	0.050	0.000	0.050	$M'_{O_2,o,1}$	0.062	0.062	%
T'_2	0.010	0.011	0.007	T'_2	1192.000	1205.347	K
P'_2	0.020	-0.002	0.020	P'_2	1.950	1.947	atm
$M'_{H_2O_2,o,2}$	0.050	0.004	0.044	$M'_{H_2O_2,o,2}$	0.222	0.222	%
$M'_{H_2O,o,2}$	0.050	-0.001	0.047	$M'_{H_2O,o,2}$	0.136	0.136	%
$M'_{O_2,o,2}$	0.050	0.000	0.050	$M'_{O_2,o,2}$	0.068	0.068	%

T_3'	0.010	0.001	0.009	T_3'	1057.000	1058.295	K
P_3'	0.020	0.000	0.020	P_3'	1.830	1.830	atm
$M'_{H_2O_2,o,3}$	0.050	0.004	0.047	$M'_{H_2O_2,o,3}$	0.086	0.086	%
$M'_{H_2O,o,3}$	0.050	0.004	0.047	$M'_{H_2O,o,3}$	0.066	0.067	%
$M'_{O_2,o,3}$	0.050	0.000	0.050	$M'_{O_2,o,3}$	0.033	0.033	%
T_4'	0.010	0.006	0.008	T_4'	1880.000	1891.608	K
P_4'	0.020	0.001	0.020	P_4'	1.740	1.742	atm
$M'_{H_2O,o,4}$	0.100	0.029	0.096	$M'_{H_2O,o,4}$	1.300	1.338	%
$M'_{O_2,o,4}$	0.010	0.000	0.010	$M'_{O_2,o,4}$	0.990	0.990	%
$M'_{H,o,4}$	2.303	0.480	0.298	$M'_{H,o,4}$	0.700	1.131	ppm
T_5'	0.020	0.002	0.018	T_5'	968.000	970.245	K
P_5'	0.040	-0.001	0.039	P_5'	3.590	3.585	atm
$M'_{H_2O_2,o,5}$	0.100	-0.009	0.088	$M'_{H_2O_2,o,5}$	0.314	0.311	%
T_6'	0.020	0.010	0.011	T_6'	1015.000	1025.579	K
P_6'	0.040	0.001	0.039	P_6'	3.298	3.301	atm
$M'_{H_2O_2,o,6}$	0.100	-0.004	0.089	$M'_{H_2O_2,o,6}$	0.326	0.325	%
T_7'	0.020	0.000	0.009	T_7'	1079.000	1079.434	K
P_7'	0.040	0.003	0.039	P_7'	3.442	3.452	atm
$M'_{H_2O_2,o,7}$	0.100	0.016	0.088	$M'_{H_2O_2,o,7}$	0.195	0.199	%
T_8'	0.020	-0.011	0.008	T_8'	1132.000	1119.176	K
P_8'	0.040	-0.001	0.039	P_8'	3.550	3.547	atm
$M'_{H_2O_2,o,8}$	0.100	0.006	0.089	$M'_{H_2O_2,o,8}$	0.136	0.137	%
T_9'	0.020	-0.008	0.009	T_9'	1218.000	1208.822	K
P_9'	0.040	-0.003	0.039	P_9'	3.038	3.0286	atm
$M'_{H_2O_2,o,9}$	0.100	-0.010	0.089	$M'_{H_2O_2,o,9}$	0.171	0.169	%
T'_{10}	0.010	0.000	0.008	T'_{10}	1072	1072.263	K
P'_{10}	0.020	0.000	0.020	P'_{10}	1.786	1.786	atm
$M'_{H_2O_2,o,10}$	0.050	-0.001	0.049	$M'_{H_2O_2,o,10}$	0.178	0.178	%
$M'_{H_2O,o,10}$	0.200	-0.020	0.174	$M'_{H_2O,o,10}$	0.148	0.145	%
$M'_{O_2,o,10}$	0.200	0.000	0.200	$M'_{O_2,o,10}$	0.056	0.056	%
T'_{11}	0.010	-0.001	0.008	T'_{11}	1108.000	1107.404	K
P'_{11}	0.020	0.000	0.020	P'_{11}	1.774	1.774	atm
$M'_{H_2O_2,o,11}$	0.050	0.000	0.049	$M'_{H_2O_2,o,11}$	0.161	0.161	%
$M'_{H_2O,o,11}$	0.200	0.010	0.181	$M'_{H_2O,o,11}$	0.097	0.098	%
$M'_{O_2,o,11}$	0.200	0.000	0.200	$M'_{O_2,o,11}$	0.056	0.56	%
T'_{12}	0.010	-0.003	0.008	T'_{12}	1138.000	1134.203	K
P'_{12}	0.020	0.000	0.020	P'_{12}	1.676	1.676	atm
$M'_{H_2O_2,o,12}$	0.050	-0.002	0.049	$M'_{H_2O_2,o,12}$	0.200	0.197	%
$M'_{H_2O,o,12}$	0.200	-0.010	0.183	$M'_{H_2O,o,12}$	0.113	0.113	%

$M'_{O_2,o,12}$	0.20	0.000	0.200	$M'_{O_2,o,12}$	0.056	0.056	%
T'_{13}	0.010	-0.001	0.009	T'_{13}	1182.000	1181.391	K
P'_{13}	0.020	0.000	0.020	P'_{13}	1.676	1.676	atm
$M'_{H_2O_2,o,13}$	0.050	0.000	0.049	$M'_{H_2O_2,o,13}$	0.205	0.205	%
$M'_{H_2O,o,13}$	0.200	-0.009	0.186	$M'_{H_2O,o,13}$	0.114	0.113	%
$M'_{O_2,o,13}$	0.200	0.000	0.200	$M'_{O_2,o,13}$	0.056	0.056	%
T'_{14}	0.010	-0.001	0.009	T'_{14}	1200.000	1198.689	K
P'_{14}	0.020	0.000	0.020	P'_{14}	1.641	1.641	atm
$M'_{H_2O_2,o,14}$	0.050	0.001	0.049	$M'_{H_2O_2,o,14}$	0.232	0.233	%
$M'_{H_2O,o,14}$	0.20	-0.004	0.188	$M'_{H_2O,o,14}$	0.106	0.106	%
$M'_{O_2,o,14}$	0.200	0.000	0.200	$M'_{O_2,o,14}$	0.056	0.56	%
T'_{15}	0.010	-0.003	0.009	T'_{15}	1194.000	1190.700	K
P'_{15}	0.020	0.000	0.020	P'_{15}	1.548	1.549	atm
$M'_{H_2O_2,o,15}$	0.050	-0.001	0.049	$M'_{H_2O_2,o,15}$	0.191	0.190	%
$M'_{H_2O,o,15}$	0.200	-0.009	0.182	$M'_{H_2O,o,15}$	0.133	0.132	%
$M'_{O_2,o,15}$	0.20	0.000	0.200	$M'_{O_2,o,15}$	0.056	0.056	%
T'_{16}	0.010	-0.002	0.009	T'_{16}	1227.000	1224.271	K
P'_{16}	0.020	0.000	0.020	P'_{16}	1.558	1.559	atm
$M'_{H_2O_2,o,16}$	0.050	0.000	0.049	$M'_{H_2O_2,o,16}$	0.170	0.170	%
$M'_{H_2O,o,16}$	0.200	0.005	0.183	$M'_{H_2O,o,16}$	0.112	0.112	%
$M'_{O_2,o,16}$	0.200	0.000	0.200	$M'_{O_2,o,16}$	0.056	0.056	%
T'_{17}	0.010	-0.004	0.009	T'_{17}	1186.000	1181.242	K
P'_{17}	0.020	0.001	0.020	P'_{17}	1.706	1.707	atm
$M'_{H_2O_2,o,17}$	0.050	-0.002	0.049	$M'_{H_2O_2,o,17}$	0.161	0.160	%
$M'_{H_2O,o,17}$	0.200	-0.011	0.181	$M'_{H_2O,o,17}$	0.112	0.110	%
$M'_{O_2,o,17}$	0.200	0.000	0.200	$M'_{O_2,o,17}$	0.056	0.056	%
T'_{18}	0.010	-0.002	0.008	T'_{18}	1135.000	1132.991	K
P'_{18}	0.020	0.000	0.020	P'_{18}	1.665	1.665	atm
$M'_{H_2O_2,o,18}$	0.050	-0.002	0.049	$M'_{H_2O_2,o,18}$	0.148	0.147	%
$M'_{H_2O,o,18}$	0.200	-0.004	0.175	$M'_{H_2O,o,18}$	0.128	0.128	%
$M'_{O_2,o,18}$	0.200	0.000	0.200	$M'_{O_2,o,18}$	0.056	0.056	%
T'_{19}	0.010	-0.002	0.009	T'_{19}	1210.000	1207.421	K
P'_{19}	0.020	0.002	0.020	P'_{19}	1.668	1.671	atm
$M'_{H_2O_2,o,19}$	0.050	-0.004	0.049	$M'_{H_2O_2,o,19}$	0.240	0.239	%
$M'_{H_2O,o,19}$	0.200	-0.018	0.175	$M'_{H_2O,o,19}$	0.248	0.243	%
$M'_{O_2,o,19}$	0.200	0.000	0.200	$M'_{O_2,o,19}$	0.056	0.056	%
T'_{20}	0.010	-0.001	0.009	T'_{20}	1245.000	1243.574	K
P'_{20}	0.020	0.001	0.020	P'_{20}	1.644	1.646	atm
$M'_{H_2O_2,o,20}$	0.050	-0.002	0.049	$M'_{H_2O_2,o,20}$	0.206	0.206	%

$M'_{H_2O,o,20}$	0.200	-0.022	0.183	$M'_{H_2O,o,20}$	0.138	0.134	%
$M'_{O_2,o,20}$	0.20	-0.001	0.200	$M'_{O_2,o,20}$	0.056	0.056	%
T'_{21}	0.010	-0.002	0.009	T'_{21}	1268.000	1265.789	K
P'_{21}	0.020	0.001	0.020	P'_{21}	1.661	1.663	atm
$M'_{H_2O_2,o,21}$	0.050	-0.004	0.049	$M'_{H_2O_2,o,21}$	0.274	0.272	%
$M'_{H_2O,o,21}$	0.200	-0.024	0.184	$M'_{H_2O,o,21}$	0.165	0.161	%
$M'_{O_2,o,21}$	0.200	-0.002	0.200	$M'_{O_2,o,21}$	0.056	0.056	%
T'_{22}	0.010	0.001	0.008	T'_{22}	1091.000	1092.303	K
P'_{22}	0.020	0.000	0.020	P'_{22}	1.754	1.754	atm
$M'_{H_2O_2,o,22}$	0.050	0.000	0.049	$M'_{H_2O_2,o,22}$	0.257	0.257	%
$M'_{H_2O,o,22}$	0.20	-0.004	0.182	$M'_{H_2O,o,22}$	0.160	0.160	%
$M'_{O_2,o,22}$	0.20	0.000	0.200	$M'_{O_2,o,22}$	0.056	0.056	%
T'_{23}	0.010	-0.002	0.009	T'_{23}	1149.000	1146.771	K
P'_{23}	0.020	0.000	0.020	P'_{23}	1.654	1.654	atm
$M'_{H_2O_2,o,23}$	0.050	-0.001	0.049	$M'_{H_2O_2,o,23}$	0.308	0.308	%
$M'_{H_2O,o,23}$	0.200	0.010	0.190	$M'_{H_2O,o,23}$	0.122	0.123	%
$M'_{O_2,o,23}$	0.200	0.000	0.200	$M'_{O_2,o,23}$	0.056	0.056	%
T'_{24}	0.010	-0.001	0.009	T'_{24}	1244.000	1242.195	K
P'_{24}	0.020	0.001	0.020	P'_{24}	1.546	1.547	atm
$M'_{H_2O_2,o,24}$	0.050	-0.004	0.046	$M'_{H_2O_2,o,24}$	0.304	0.302	%
$M'_{H_2O,o,24}$	0.200	-0.042	0.116	$M'_{H_2O,o,24}$	0.163	0.156	%
$M'_{O_2,o,24}$	0.200	0.000	0.200	$M'_{O_2,o,24}$	0.056	0.056	%
T'_{25}	0.010	0.000	0.009	T'_{25}	1283.000	1282.743	K
P'_{25}	0.020	0.000	0.020	P'_{25}	1.635	1.635	atm
$M'_{H_2O_2,o,25}$	0.050	0.002	0.049	$M'_{H_2O_2,o,25}$	0.309	0.310	%
$M'_{H_2O,o,25}$	0.200	0.008	0.193	$M'_{H_2O,o,25}$	0.111	0.112	%
$M'_{O_2,o,25}$	0.20	0.000	0.200	$M'_{O_2,o,25}$	0.056	0.056	%
T'_{26}	0.010	0.000	0.010	T'_{26}	240.000	240.000	K
P'_{26}	0.020	0.000	0.020	P'_{26}	0.105	0.105	atm
$M'_{Cl,o,26}$	0.100	-0.002	0.090	$M'_{Cl,o,26}$	16.500	16.47	ppm
$M'_{Cl_2,o,26}$	0.050	0.000	0.050	$M'_{Cl_2,o,26}$	0.259	0.259	%
$M'_{CH_3OH,o,26}$	0.050	0.000	0.050	$M'_{CH_3OH,o,2}$	0.094	0.094	%
$M'_{O_2,o,26}$	0.050	0.000	0.050	$M'_{O_2,o,26}$	8.620	8.620	%
T'_{27}	0.010	0.000	0.010	T'_{27}	268.000	268.000	K
P'_{27}	0.020	0.000	0.020	P'_{27}	0.119	0.119	atm
$M'_{Cl,o,27}$	0.100	0.002	0.090	$M'_{Cl,o,27}$	16.100	16.126	ppm
$M'_{Cl_2,o,27}$	0.050	0.000	0.050	$M'_{Cl_2,o,27}$	0.253	0.253	%
$M'_{CH_3OH,o,27}$	0.050	0.000	0.050	$M'_{CH_3OH,o,2}$	0.092	0.092	%
$M'_{O_2,o,27}$	0.050	0.000	0.050	$M'_{O_2,o,27}$	8.440	8.440	%

T'_{28}	0.010	0.000	0.010	T'_{28}	298.000	298.000	K
P'_{28}	0.020	0.000	0.020	P'_{28}	0.132	0.132	atm
$M'_{Cl,o,28}$	0.100	-0.001	0.098	$M'_{Cl,o,28}$	16.200	16.100	ppm
$M'_{Cl_2,o,28}$	0.050	0.000	0.050	$M'_{Cl_2,o,28}$	0.255	0.255	%
$M'_{CH_3OH,o,28}$	0.050	0.000	0.050	$M'_{CH_3OH,o,2}$	0.093	0.093	%
$M'_{O_2,o,28}$	0.050	0.000	0.050	$M'_{O_2,o,28}$	8.480	8.480	%
T'_{29}	0.010	0.000	0.010	T'_{29}	333.000	333.000	K
P'_{29}	0.020	0.000	0.020	P'_{29}	0.147	0.147	atm
$M'_{Cl,o,29}$	0.100	0.001	0.097	$M'_{Cl,o,29}$	16.300	16.310	ppm
$M'_{Cl_2,o,29}$	0.050	0.000	0.050	$M'_{Cl_2,o,29}$	0.255	0.255	%
$M'_{CH_3OH,o,29}$	0.050	0.000	0.050	$M'_{CH_3OH,o,2}$	0.093	0.093	%
$M'_{O_2,o,29}$	0.050	0.000	0.050	$M'_{O_2,o,29}$	8.500	8.500	%
T'_{30}	0.010	0.000	0.010	T'_{30}	417.000	417.000	K
P'_{30}	0.020	0.000	0.020	P'_{30}	0.184	0.183	atm
$M'_{Cl,o,30}$	0.100	-0.001	0.097	$M'_{Cl,o,30}$	16.300	16.292	ppm
$M'_{Cl_2,o,30}$	0.050	0.000	0.050	$M'_{Cl_2,o,30}$	0.255	0.255	%
$M'_{CH_3OH,o,30}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.093	0.092	%
$M'_{O_2,o,30}$	0.050	0.000	0.050	$M'_{O_2,o,30}$	8.505	8.505	%
T'_{31}	0.010	0.000	0.010	T'_{31}	240.000	240.000	K
P'_{31}	0.020	0.000	0.020	P'_{31}	0.329	0.329	atm
$M'_{Cl,o,31}$	0.010	-0.001	0.099	$M'_{Cl,o,31}$	5.23	5.22	ppm
$M'_{Cl_2,o,31}$	0.050	0.000	0.050	$M'_{Cl_2,o,31}$	0.082	0.082	%
$M'_{CH_3OH,o,31}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.030	0.030	%
$M'_{O_2,o,31}$	0.050	0.000	0.050	$M'_{O_2,o,31}$	2.733	2.733	%
T'_{32}	0.010	0.000	0.010	T'_{32}	268.000	268.005	K
P'_{32}	0.020	0.000	0.020	P'_{32}	0.658	0.658	atm
$M'_{Cl,o,32}$	0.010	0.003	0.098	$M'_{Cl,o,32}$	2.920	3.001	ppm
$M'_{Cl_2,o,32}$	0.050	0.000	0.050	$M'_{Cl_2,o,32}$	0.046	0.046	%
$M'_{CH_3OH,o,32}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.017	0.017	%
$M'_{O_2,o,32}$	0.050	0.000	0.050	$M'_{O_2,o,32}$	1.530	1.530	%
T'_{33}	0.01	0.000	0.010	T'_{33}	298.000	298.000	K
P'_{33}	0.02	0.000	0.020	P'_{33}	0.461	0.461	atm
$M'_{Cl,o,33}$	0.100	-0.002	0.098	$M'_{Cl,o,33}$	4.640	4.554	ppm
$M'_{Cl_2,o,33}$	0.050	0.000	0.050	$M'_{Cl_2,o,33}$	0.073	0.073	%
$M'_{CH_3OH,o,33}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.027	0.024	%
$M'_{O_2,o,33}$	0.050	0.000	0.050	$M'_{O_2,o,33}$	2.242	2.424	%
T'_{34}	0.010	0.000	0.010	T'_{34}	333.000	333.000	K
P'_{34}	0.020	0.000	0.020	P'_{34}	0.526	0.526	atm
$M'_{Cl,o,34}$	0.100	0.025	0.092	$M'_{Cl,o,34}$	4.530	4.643	ppm

$M'_{Cl_2,o,34}$	0.050	0.000	0.050	$M'_{Cl_2,o,34}$	0.071	0.071	%
$M'_{CH_3OH,o,34}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.026	0.026	%
$M'_{O_2,o,34}$	0.050	0.000	0.050	$M'_{O_2,o,34}$	2.370	2.370	%
T'_{35}	0.010	0.000	0.010	T'_{35}	417.000	417.000	K
P'_{35}	0.020	0.000	0.020	P'_{35}	0.658	0.658	atm
$M'_{Cl,o,35}$	0.100	-0.003	0.097	$M'_{Cl,o,35}$	4.54	4.525	ppm
$M'_{Cl_2,o,35}$	0.050	0.000	0.050	$M'_{Cl_2,o,35}$	0.071	0.071	%
$M'_{CH_3OH,o,35}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.026	0.026	%
$M'_{O_2,o,35}$	0.050	0.000	0.050	$M'_{O_2,o,35}$	2.375	2.375	%
T'_{36}	0.010	0.000	0.010	T'_{36}	240.000	240.000	K
P'_{36}	0.020	0.000	0.020	P'_{36}	0.731	0.731	atm
$M'_{Cl,o,36}$	0.100	0.001	0.099	$M'_{Cl,o,36}$	2.350	2.351	ppm
$M'_{Cl_2,o,36}$	0.050	0.000	0.050	$M'_{Cl_2,o,36}$	0.037	0.037	%
$M'_{CH_3OH,o,36}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.013	0.013	%
$M'_{O_2,o,36}$	0.050	0.001	0.048	$M'_{O_2,o,36}$	1.230	1.231	%
T'_{37}	0.010	0.000	0.010	T'_{37}	268.000	268.000	K
P'_{37}	0.020	0.000	0.020	P'_{37}	0.921	0.921	atm
$M'_{Cl,o,37}$	0.100	0.001	0.096	$M'_{Cl,o,37}$	2.090	2.092	ppm
$M'_{Cl_2,o,37}$	0.050	0.000	0.050	$M'_{Cl_2,o,37}$	0.033	0.033	%
$M'_{CH_3OH,o,37}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.012	0.012	%
$M'_{O_2,o,37}$	0.050	0.000	0.050	$M'_{O_2,o,37}$	1.090	1.090	%
T'_{38}	0.010	0.000	0.010	T'_{38}	298.000	298.000	K
P'_{38}	0.020	0.000	0.020	P'_{38}	0.921	0.921	atm
$M'_{Cl,o,38}$	0.100	-0.003	0.097	$M'_{Cl,o,38}$	2.320	2.314	ppm
$M'_{Cl_2,o,38}$	0.050	0.000	0.050	$M'_{Cl_2,o,38}$	0.036	0.036	%
$M'_{CH_3OH,o,38}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.013	0.013	%
$M'_{O_2,o,38}$	0.050	0.000	0.050	$M'_{O_2,o,38}$	1.210	1.210	%
T'_{39}	0.010	0.000	0.010	T'_{39}	333.000	333.000	K
P'_{39}	0.020	0.000	0.020	P'_{39}	1.028	1.028	atm
$M'_{Cl,o,39}$	0.100	0.001	0.098	$M'_{Cl,o,39}$	2.320	2.321	ppm
$M'_{Cl_2,o,39}$	0.050	0.000	0.050	$M'_{Cl_2,o,39}$	0.036	0.036	%
$M'_{CH_3OH,o,39}$	0.050	0.000	0.050	$M'_{CH_3OH,o,3}$	0.013	0.013	%
$M'_{O_2,o,39}$	0.050	0.000	0.050	$M'_{O_2,o,39}$	1.214	1.214	%
T'_{40}	0.010	0.000	0.010	T'_{40}	417.000	417.000	K
P'_{40}	0.020	0.000	0.020	P'_{40}	1.306	1.306	atm
$M'_{Cl,o,40}$	0.100	-0.001	0.097	$M'_{Cl,o,40}$	2.290	2.288	ppm
$M'_{Cl_2,o,40}$	0.050	0.000	0.050	$M'_{Cl_2,o,40}$	0.036	0.036	%
$M'_{CH_3OH,o,40}$	0.050	0.000	0.050	$M'_{CH_3OH,o,4}$	0.013	0.013	%

$M'_{O_2,o,40}$	0.050	0.000	0.050	$M'_{O_2,o,40}$	1.196	1.196	%
σ'_{1,H_2O_2}	0.700	0.095	0.133	σ'_{1,H_2O_2}	96000.000	105546.376	cm^2mol^{-1}
σ'_{2,H_2O_2}	0.300	-0.005	0.299	σ'_{2,H_2O_2}	-19.000	-18.908	$cm^2mol^{-1}K^{-1}$
σ'_{1,HO_2}	0.700	-0.119	0.301	σ'_{1,HO_2}	1155000.00	1025196.65	cm^2mol^{-1}
σ'_{2,HO_2}	0.300	-0.012	0.293	σ'_{2,HO_2}	1299.000	1283.071	K
$\widehat{\sigma'}_{1,H_2O_2}$	0.7	0.130	0.371	$\widehat{\sigma'}_{1,H_2O_2}$	130178.000	148198.394	cm^2mol^{-1}
$\widehat{\sigma'}_{2,H_2O_2}$	0.7	-0.006	0.699	$\widehat{\sigma'}_{2,H_2O_2}$	-4.315	-4.290	$cm^2mol^{-1}K^{-1}$
$\widehat{\sigma'}_{1,HO_2}$	0.700	0.0193	0.253	$\widehat{\sigma'}_{1,HO_2}$	2000000.00	2038933.23	cm^2mol^{-1}
$\widehat{\sigma'}_{2,HO_2}$	0.700	-0.087	0.625	$\widehat{\sigma'}_{2,HO_2}$	-639.580	-586.442	$cm^2mol^{-1}K^{-1}$

(b) Additional Figures

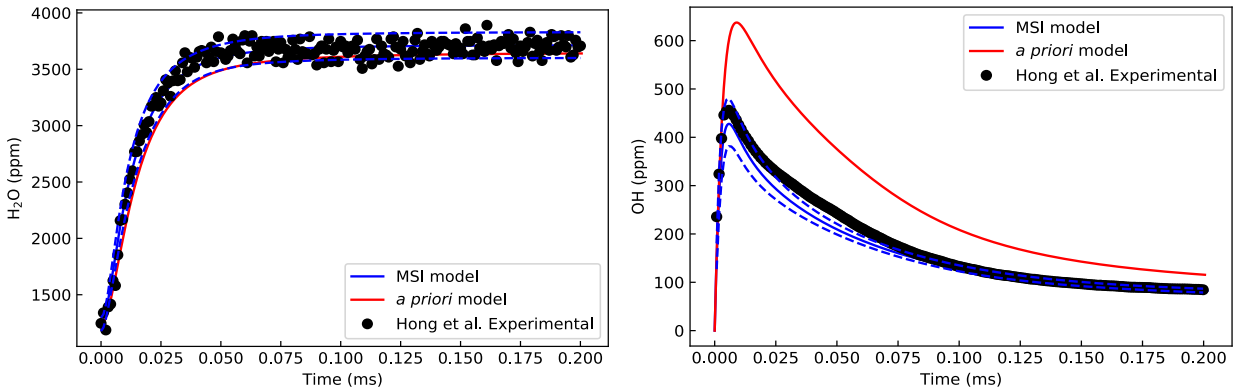


Fig. S1. H_2O (left) and OH (right) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1398 K and 1.909 atm. Symbols represent experimental data from Hong et al. [S1]; lines model predictions.

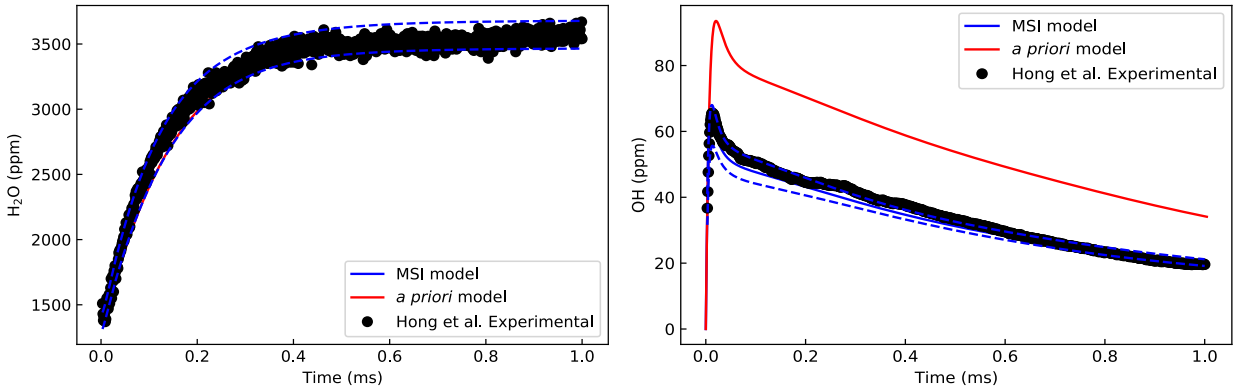


Fig. S2. H_2O (left) and OH (right) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1192 K and 1.950 atm. Symbols represent experimental data from Hong et al. [S1]; lines model predictions.

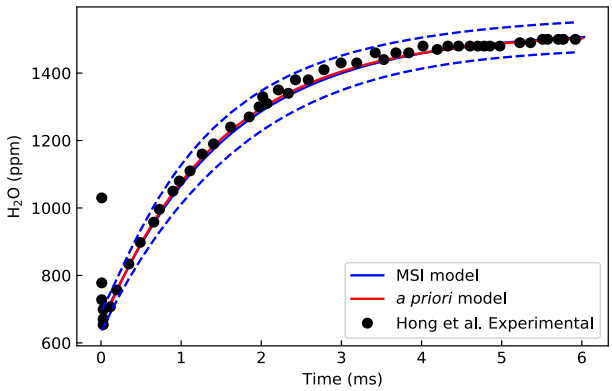


Fig. S3. H_2O time-history in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1057 K and 1.830 atm. Symbols represent experimental data from Hong et al. [S2]; lines model predictions.

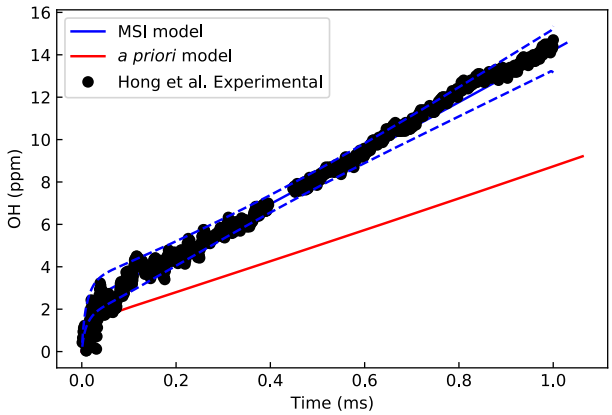


Fig. S4. OH time-history in shock heated $H_2O/O_2/H/Ar$ mixture near 1880 K and 1.740 atm. Symbols represent experimental data from Hong et al. [S3]; lines model predictions.

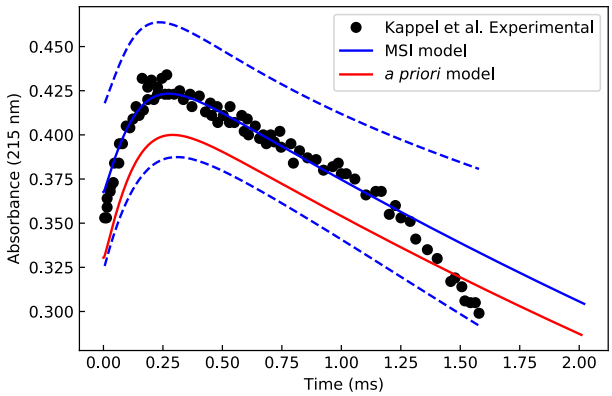


Fig. S5. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 968 K and 3.590 atm. Symbols represent experimental data from Kappel et al. [S4]; lines model predictions.

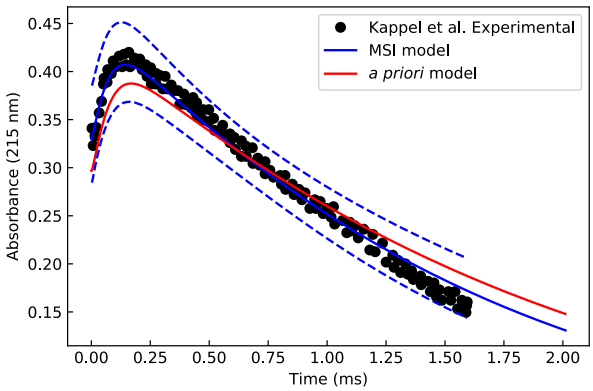


Fig. S6. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1015 K and 3.298 atm. Symbols represent experimental data from Kappel et al. [S4]; lines model predictions.

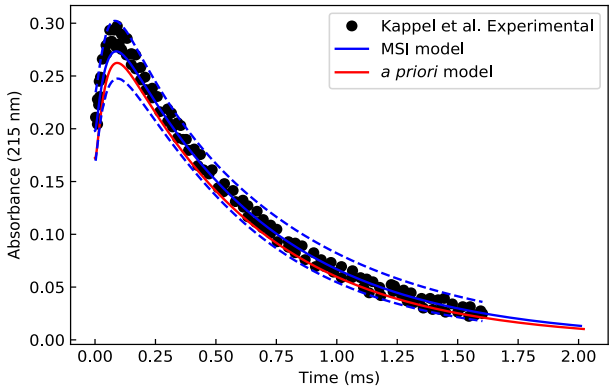


Fig. S7. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1079 K and 3.441 atm. Symbols represent experimental data from Kappel et al. [S4]; lines model predictions.

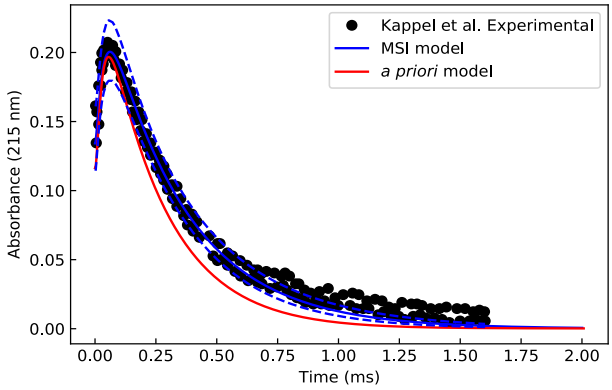


Fig. S8. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1132 K and 3.550 atm. Symbols represent experimental data from Kappel et al. [S4]; lines model predictions.

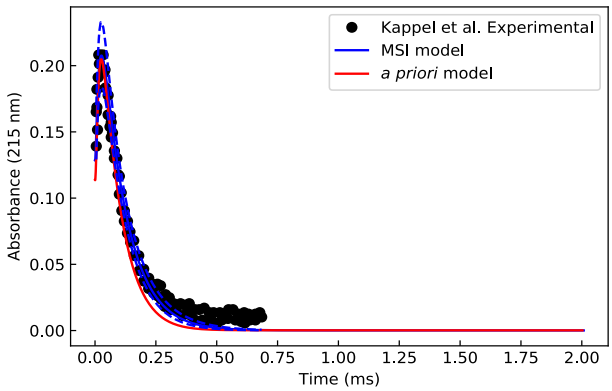


Fig. S9. Absorbance 215 nm time-history in shock heated H_2O_2/Ar mixture near 1218 K and 3.038 atm. Symbols represent experimental data from Kappel et al. [S4]; lines model predictions.

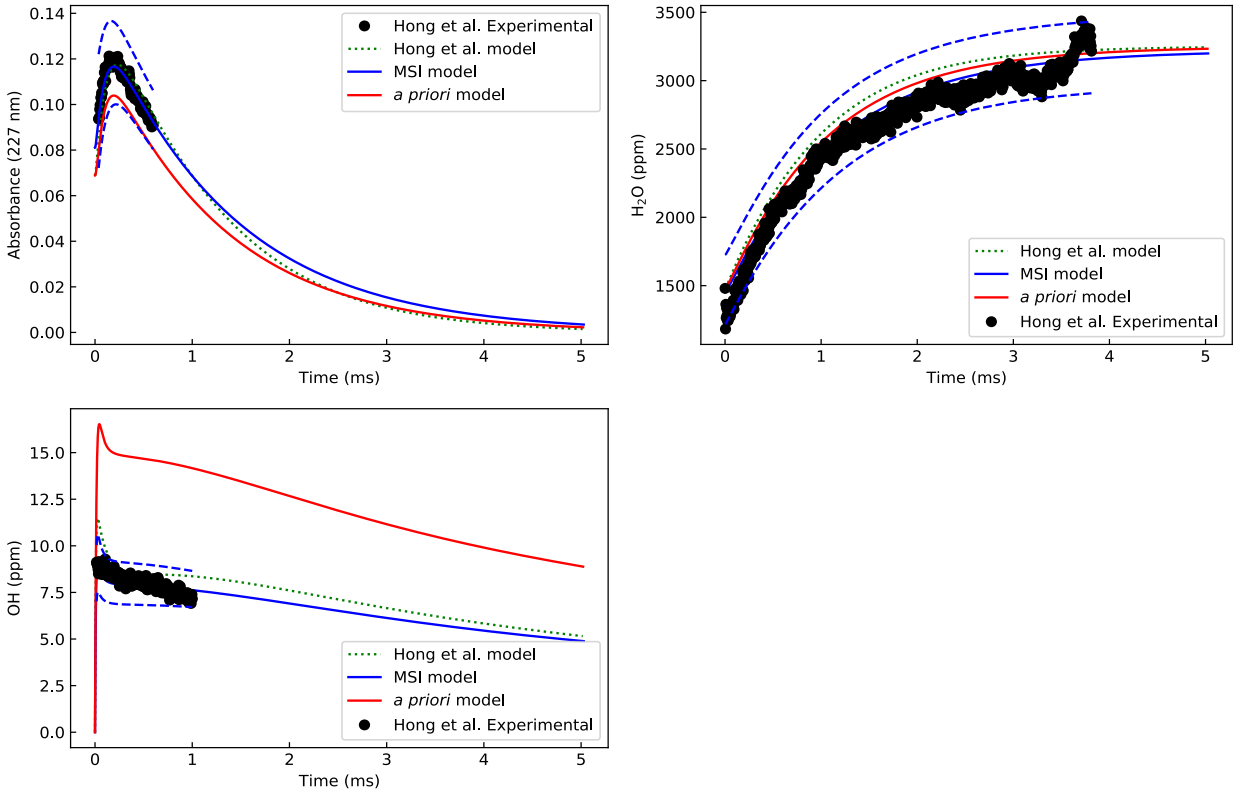


Fig. S10. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2$/Ar mixture near 1072 K and 1.786 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

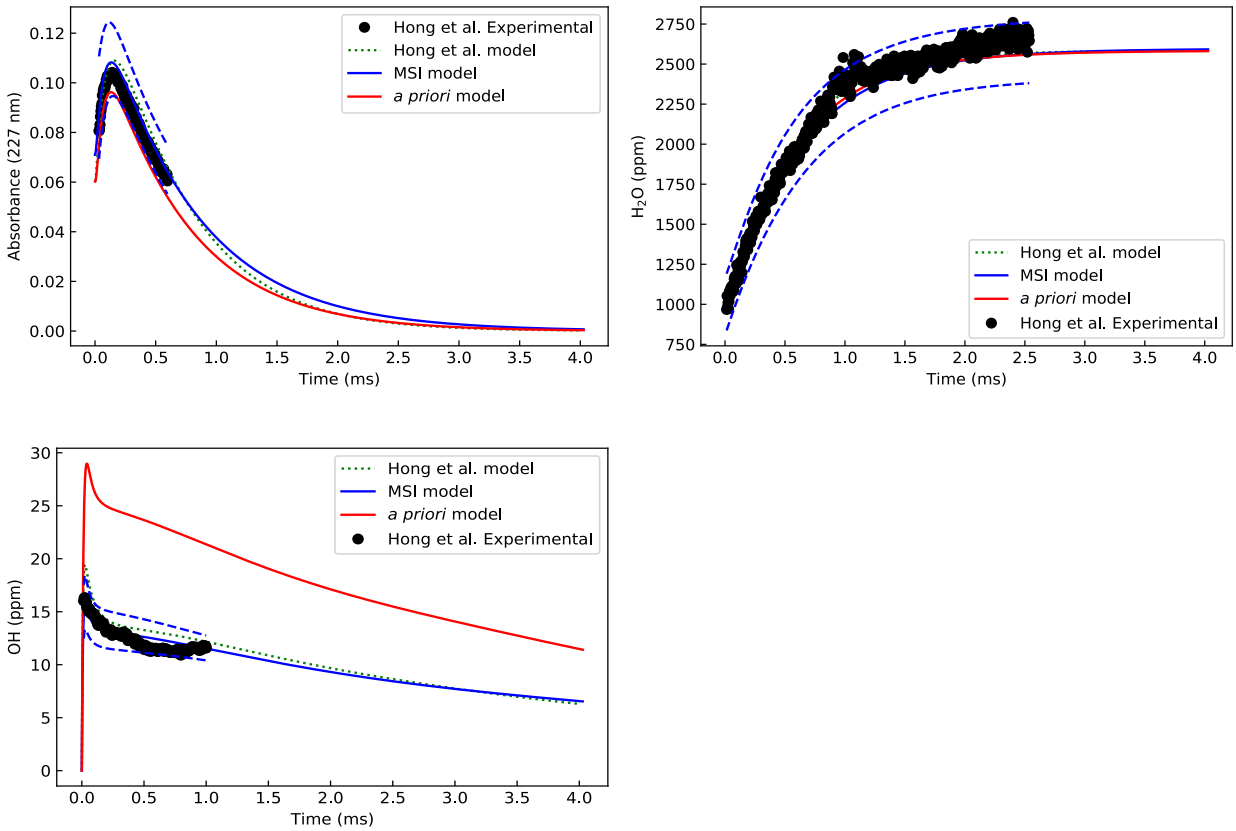
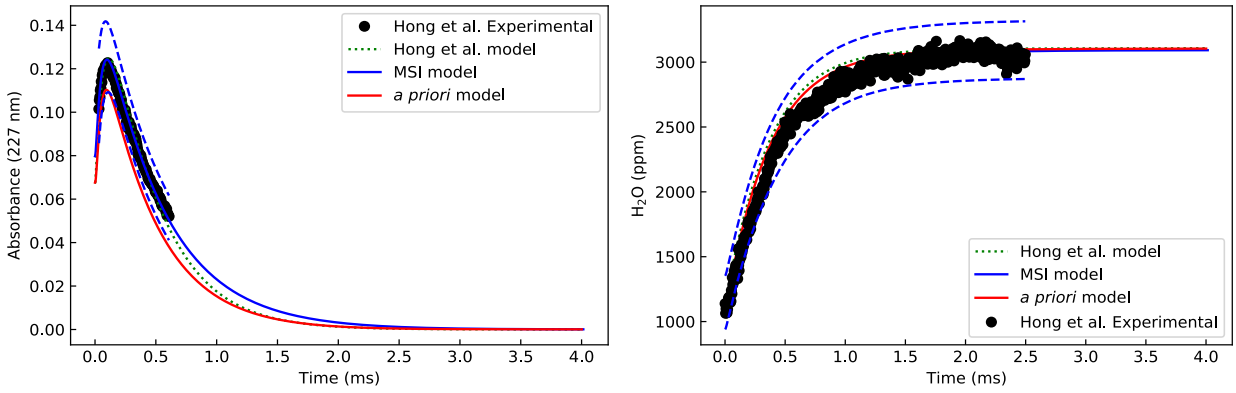


Fig. S11. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1108 K and 1.774 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.



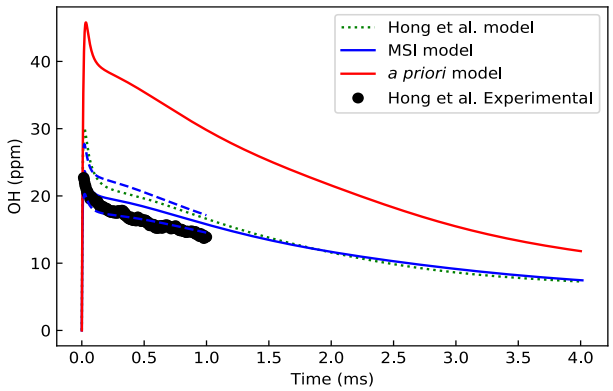


Fig. S12. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1138 K and 1.676 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

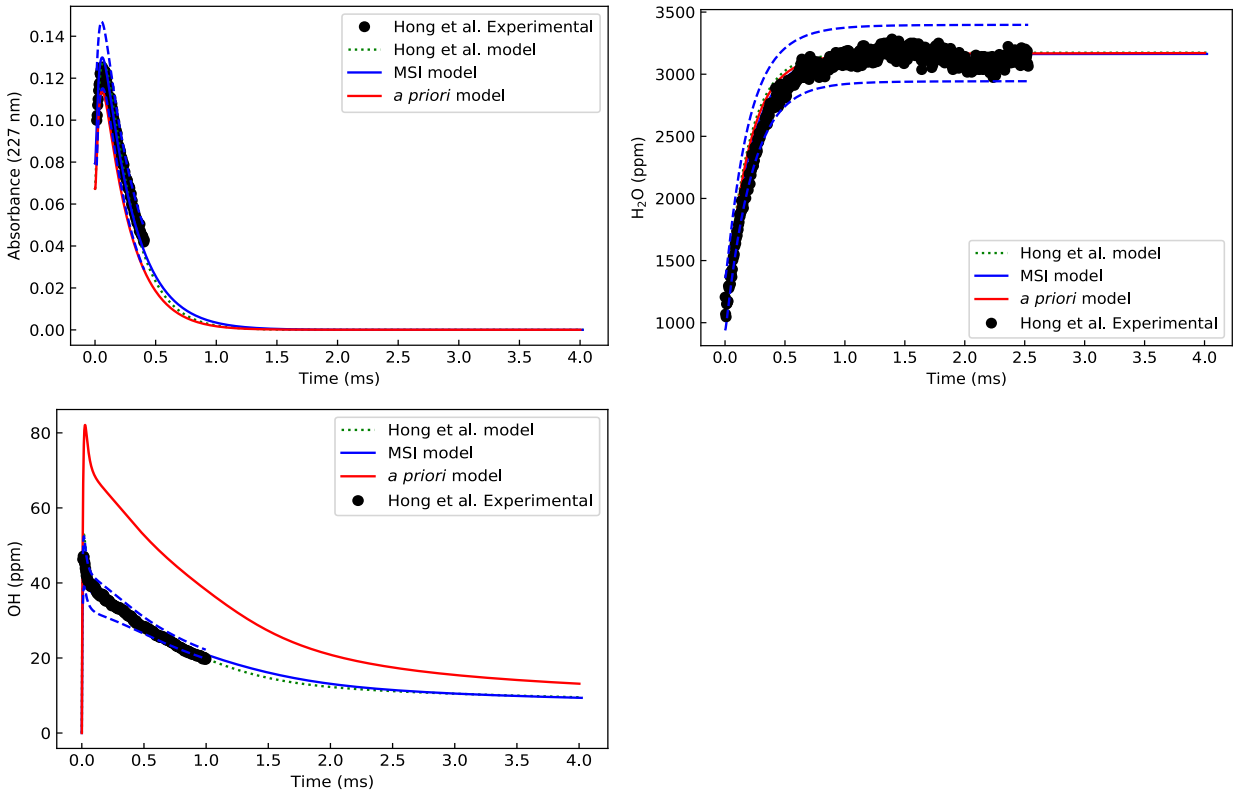


Fig. S13. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1182 K and 1.676 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

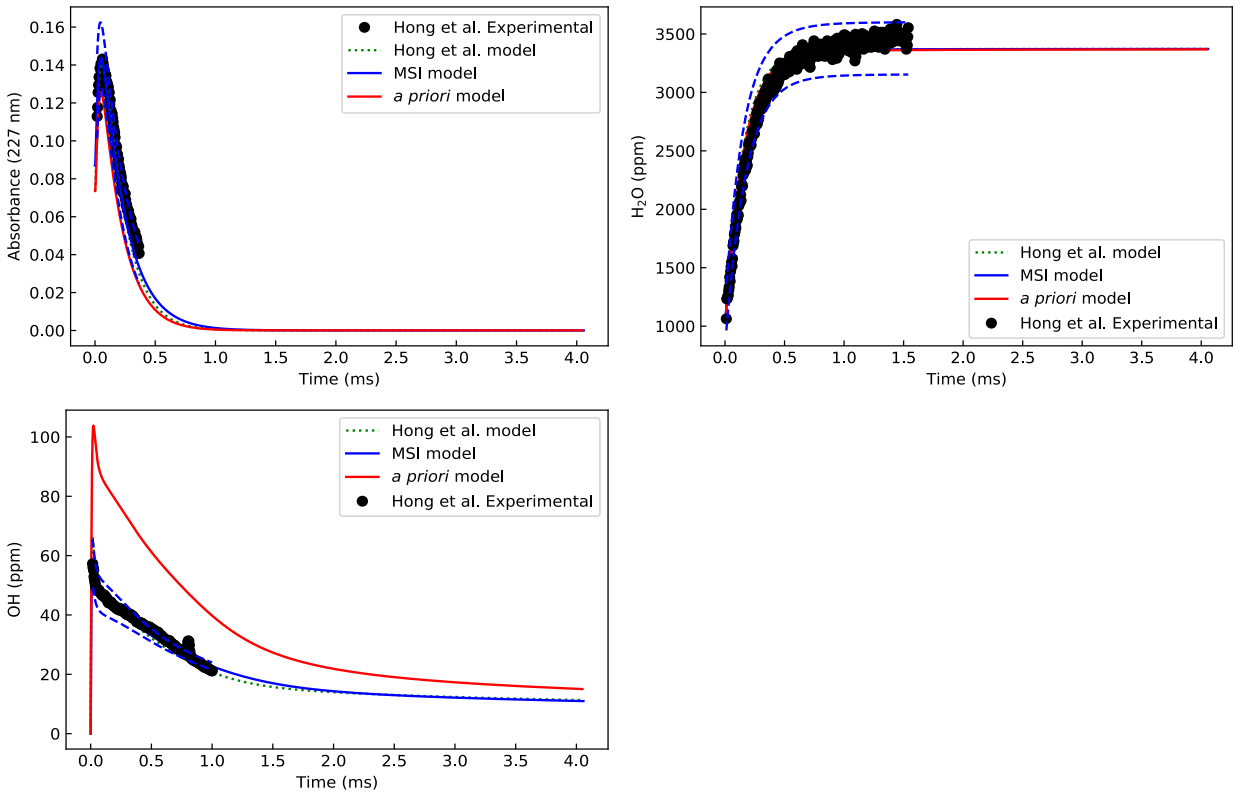


Fig. S14. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1200 K and 1.641 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

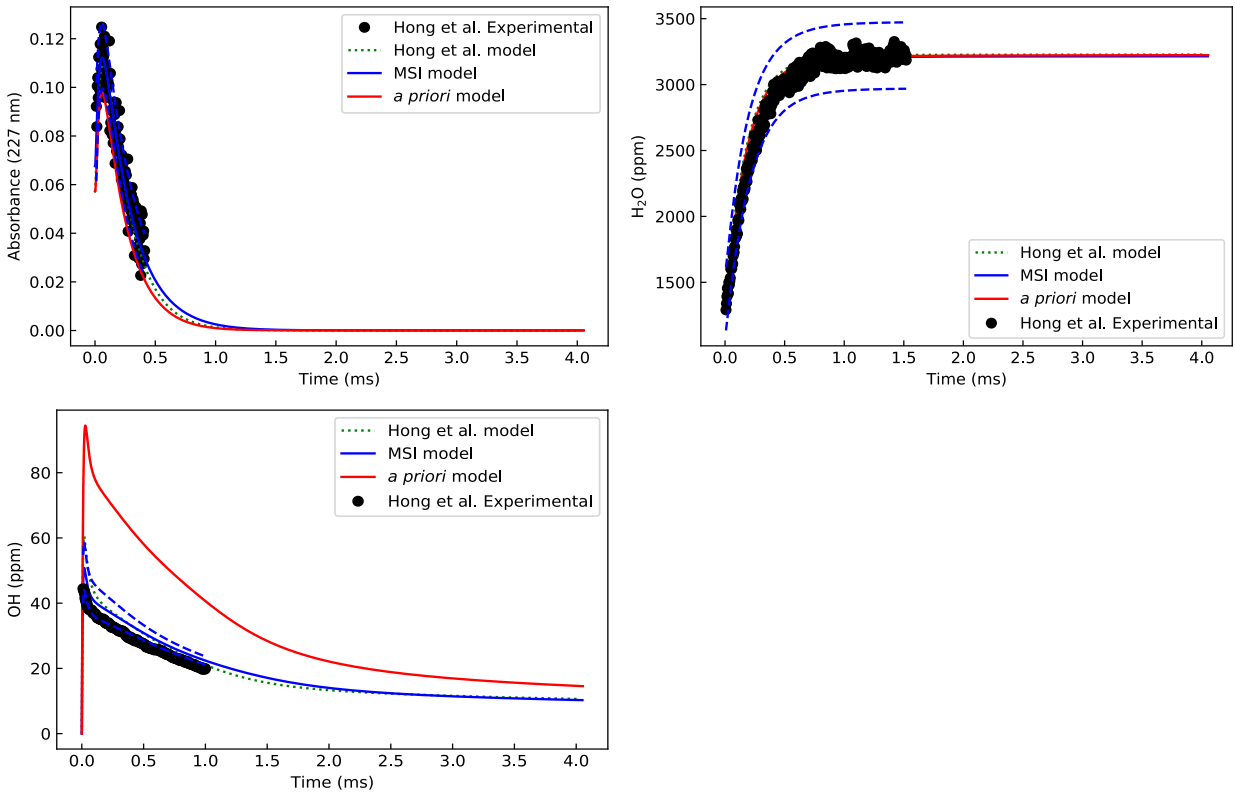


Fig. S15. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1194 K and 1.548 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

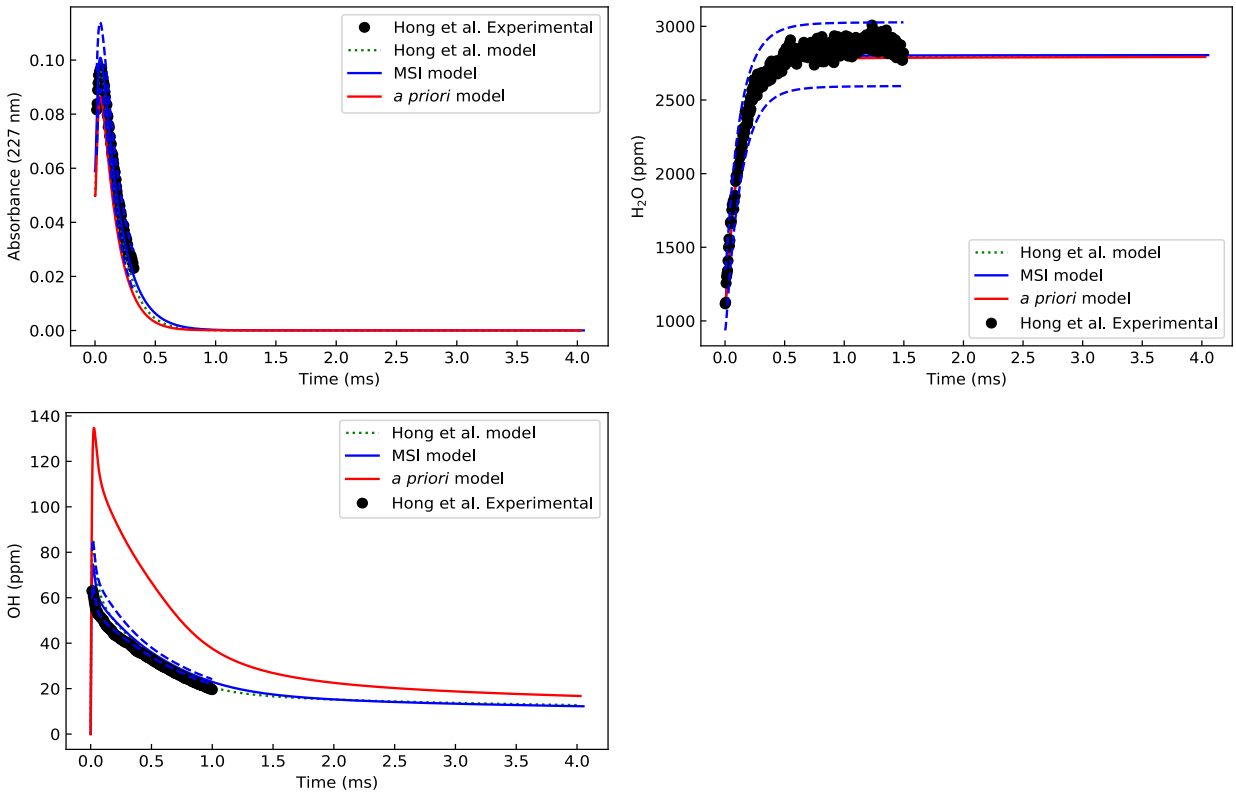


Fig. S16. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1227 K and 1.558 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

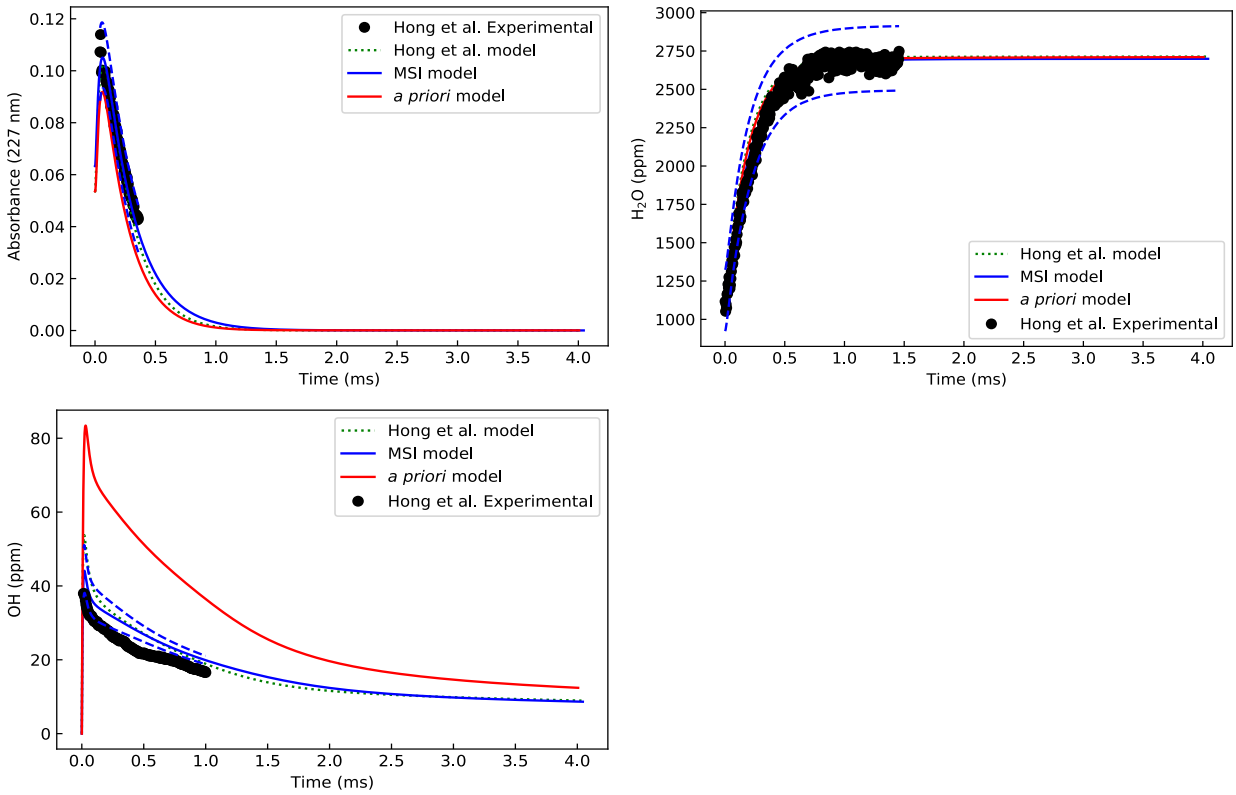


Fig. S17. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1186 K and 1.706 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

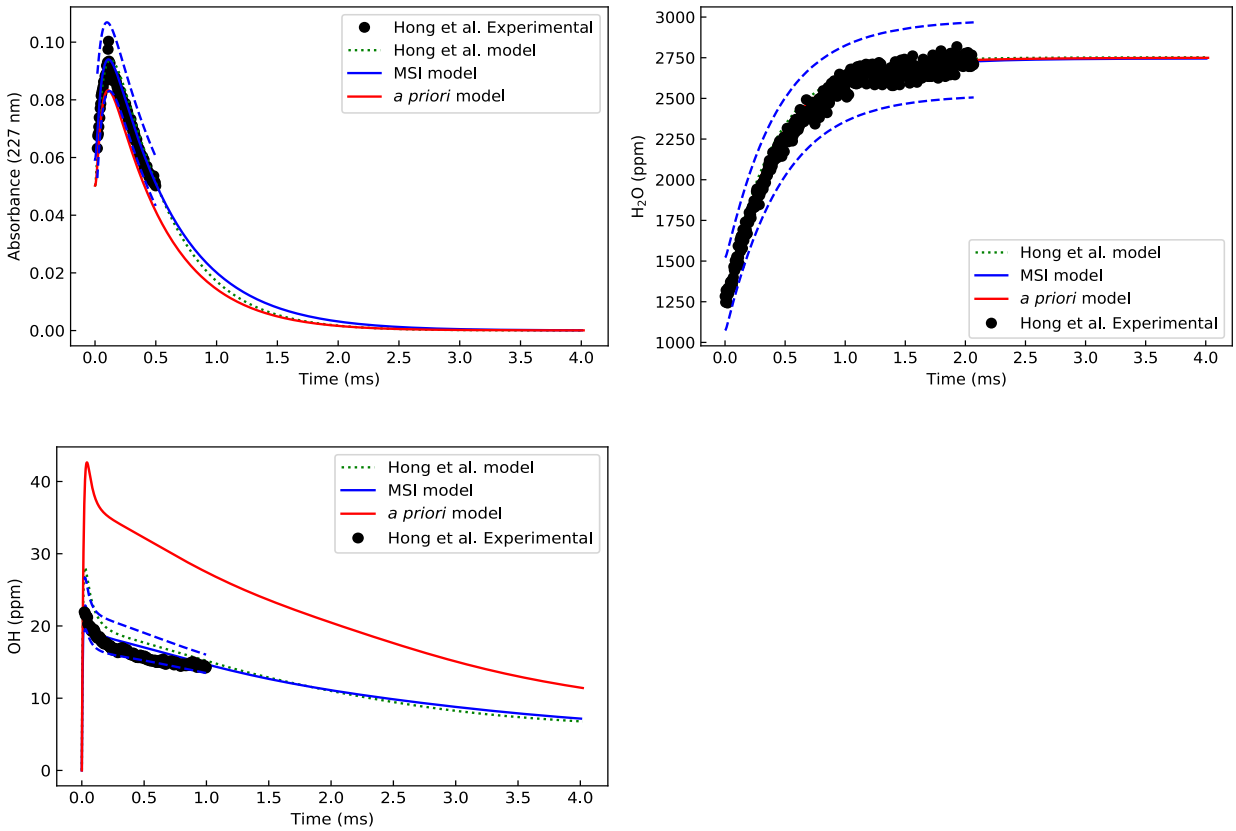


Fig. S18. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1135 K and 1.665 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

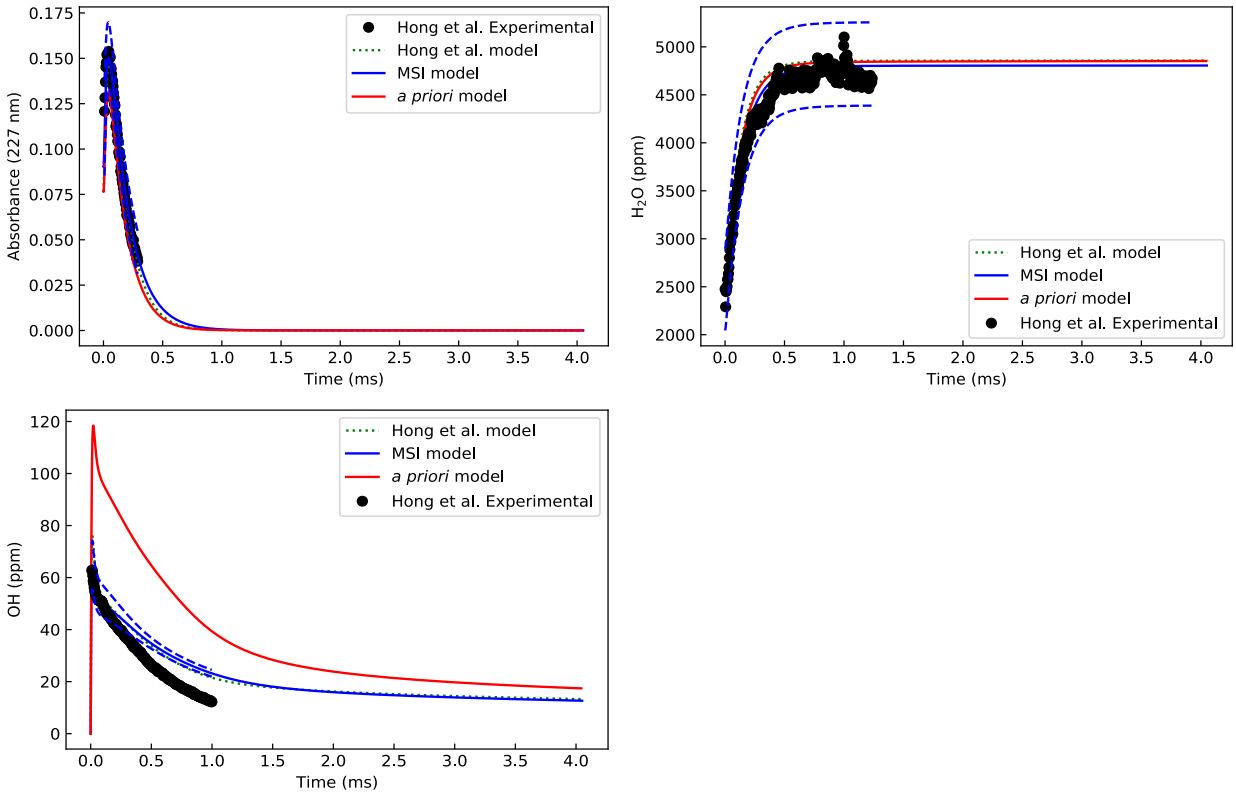


Fig. S19. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1210 K and 1.668 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

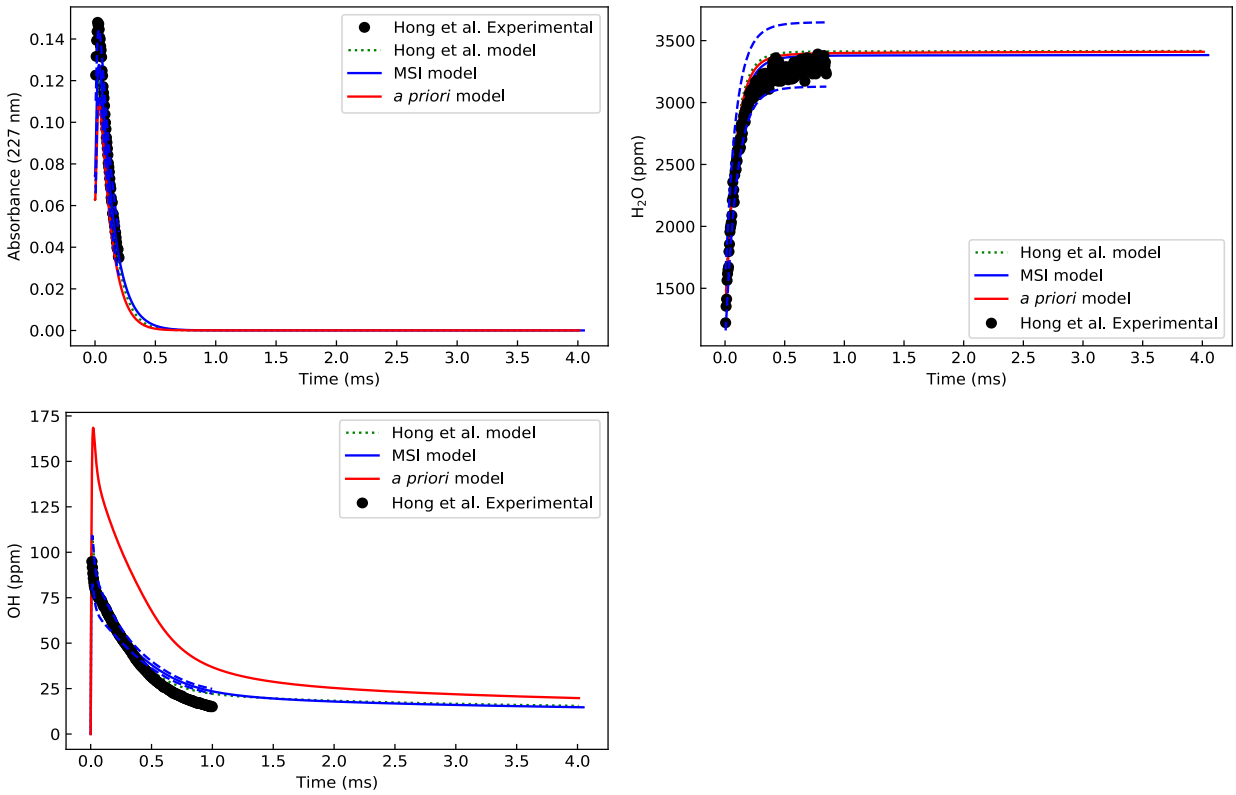


Fig. S20. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1245 K and 1.644 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

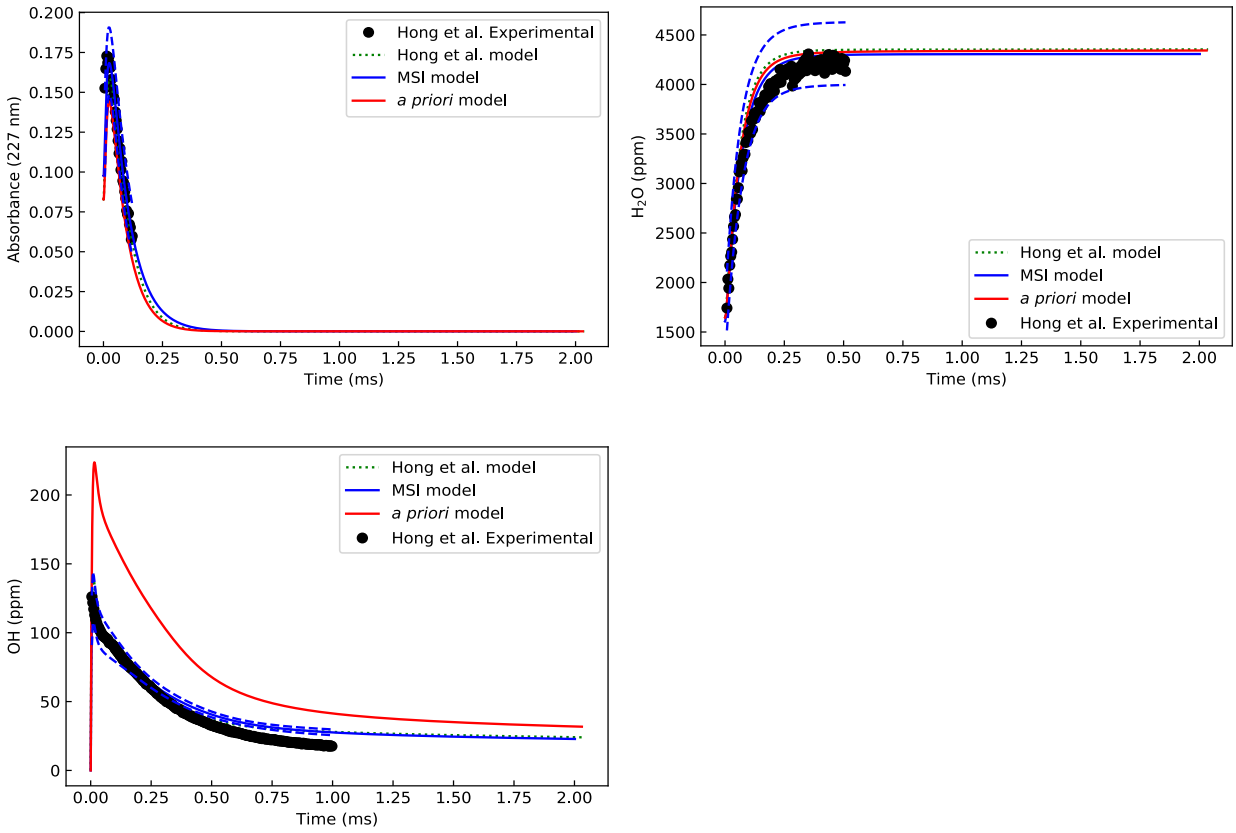


Fig. S21. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1268 K and 1.661 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

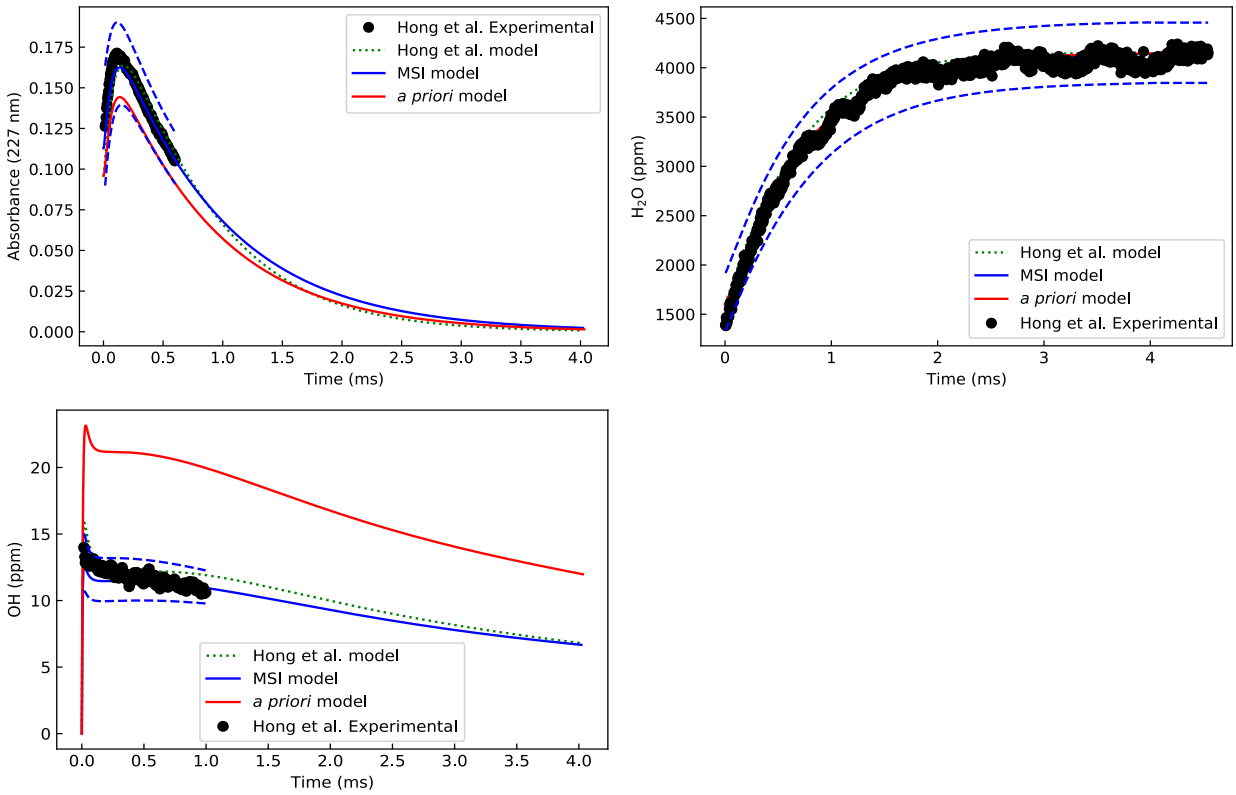


Fig. S22. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1091 K and 1.754 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

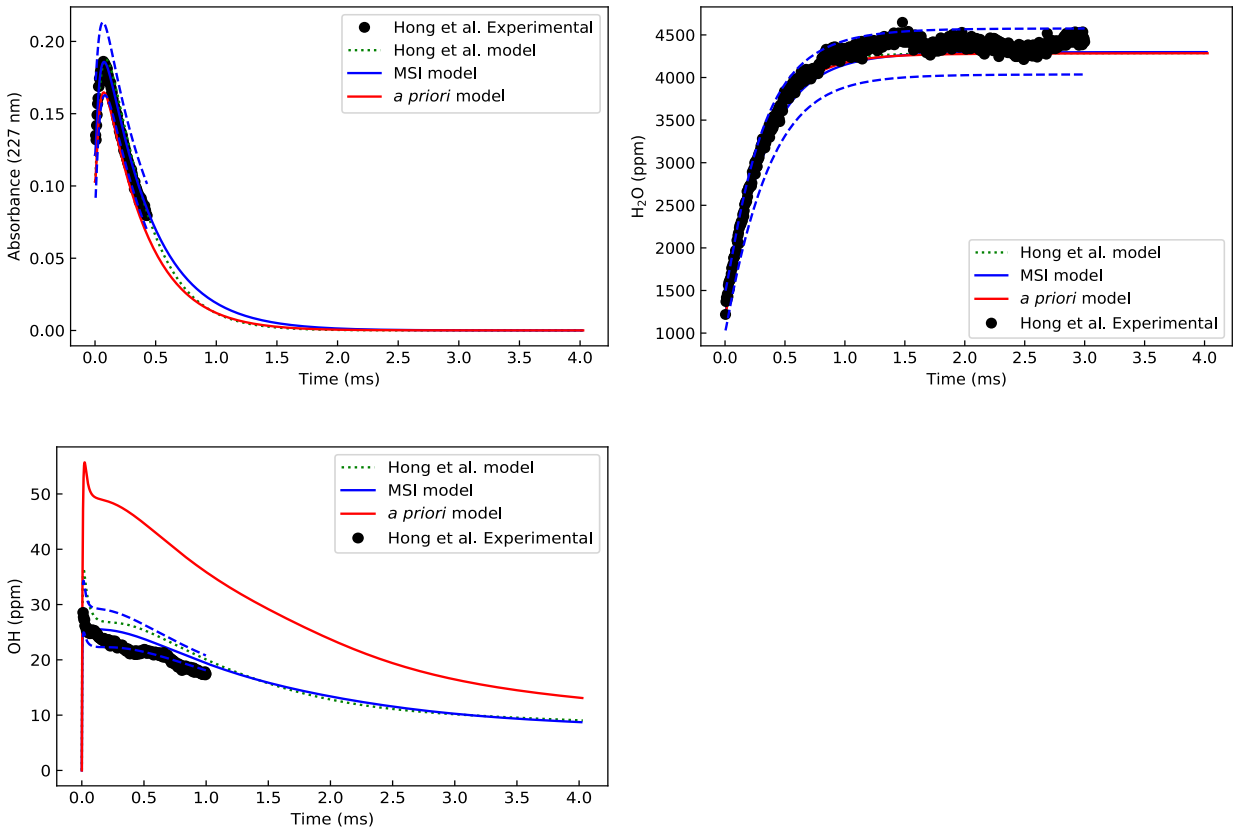


Fig. S23. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1149 K and 1.654 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

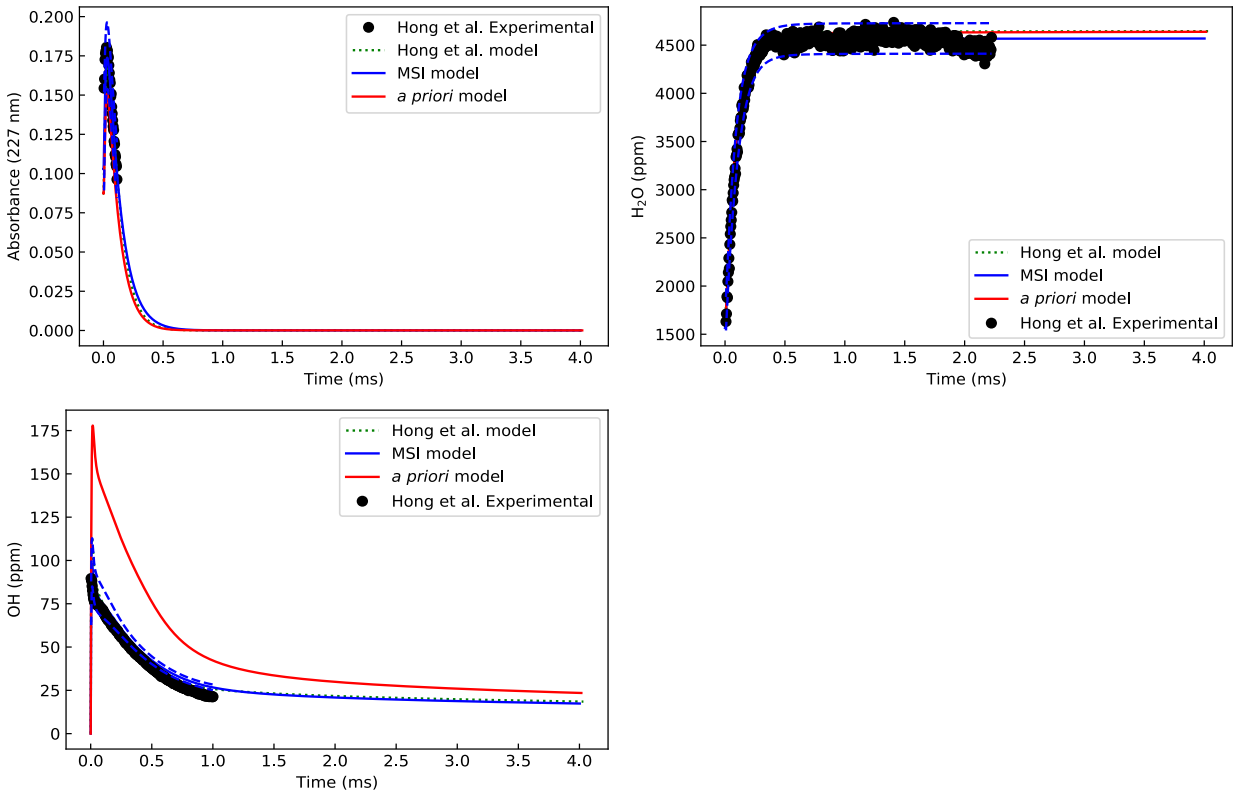


Fig. S24. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1244 K and 1.546 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

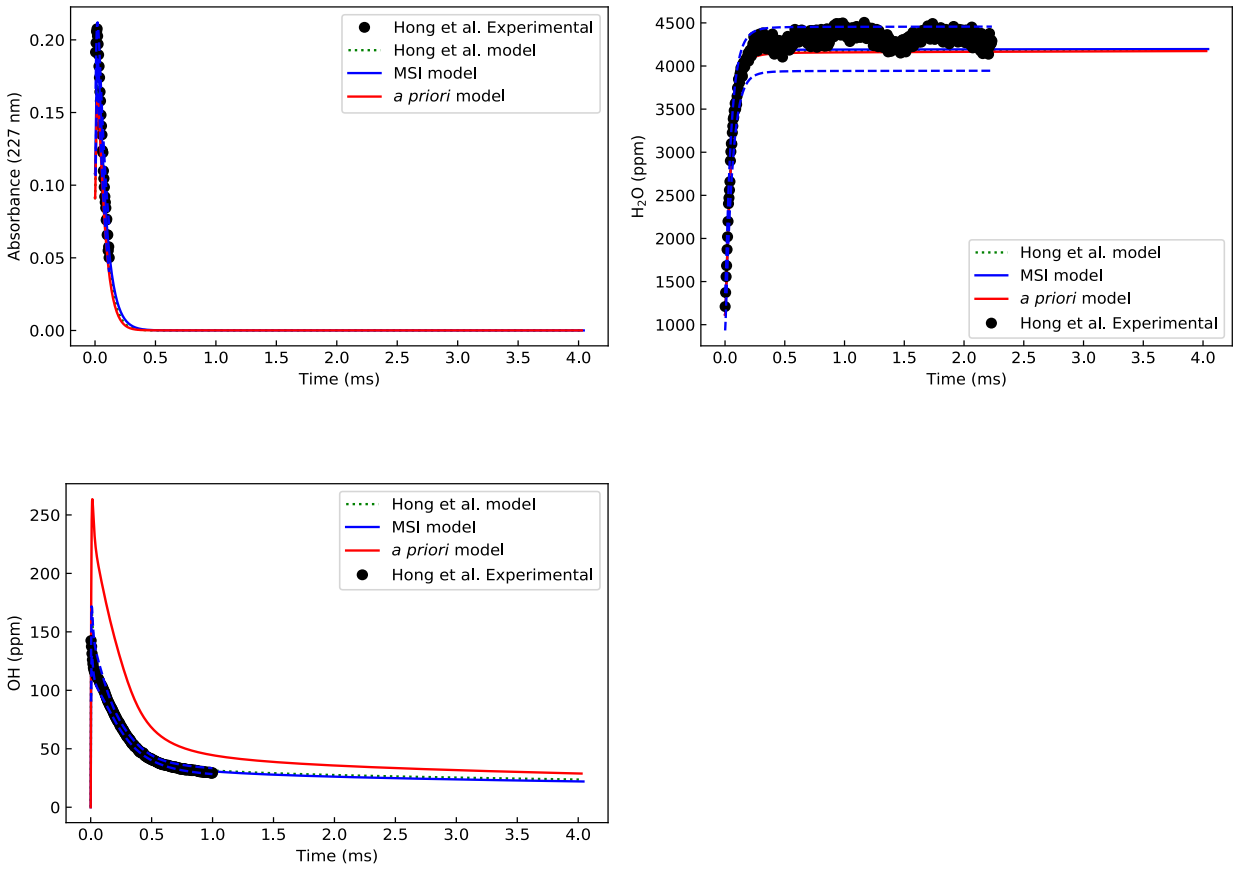


Fig. S25. Absorbance 227 nm (top left), H_2O (top right) and OH (bottom left) time-histories in shock heated $H_2O_2/H_2O/O_2/Ar$ mixture near 1283 K and 1.635 atm. Symbols represent experimental data from Hong et al. [S5]; lines model predictions.

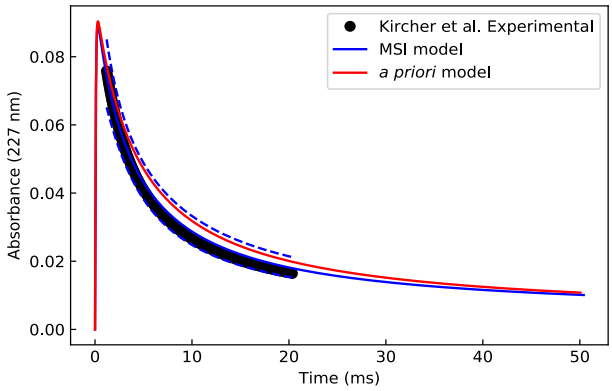


Fig. S26. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 240 K and 0.105 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

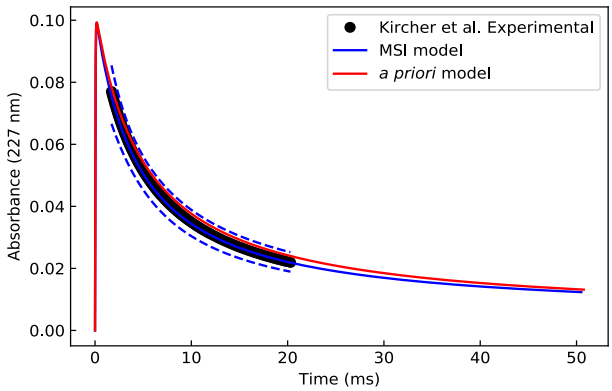


Fig. S27. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 268 K and 0.119 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

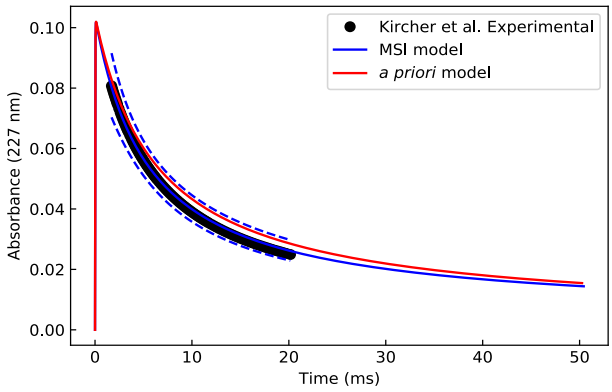


Fig. S28. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 298 K and 0.132 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

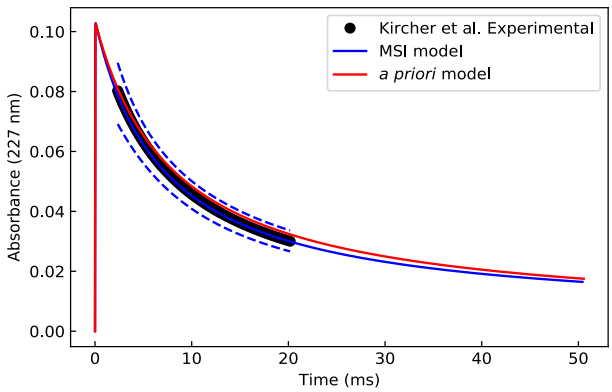


Fig. S29. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 333 K and 0.147 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

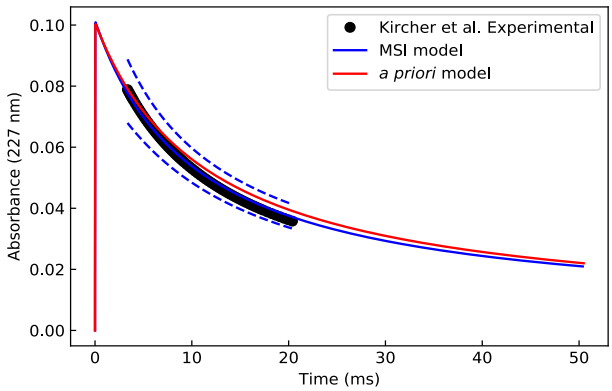


Fig. S30. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 417 K and 0.184 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

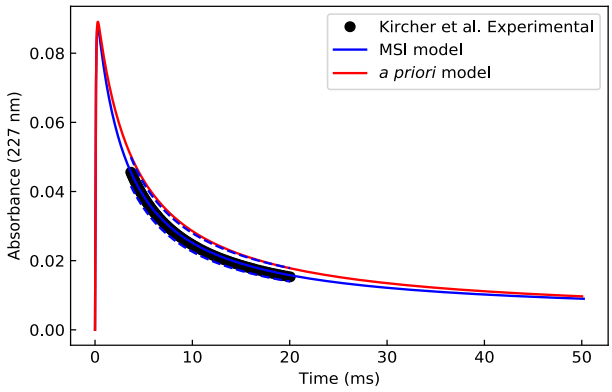


Fig. S31. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 240 K and 0.329 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

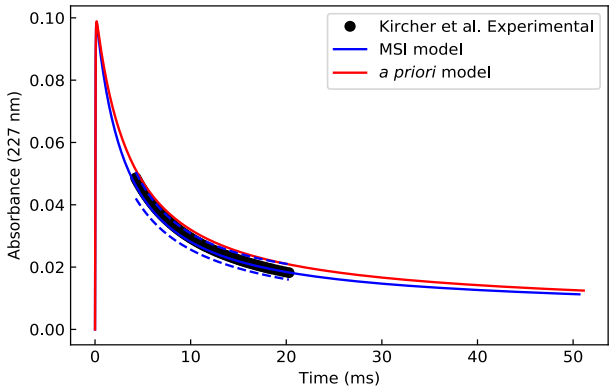


Fig. S32. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 268 K and 0.657 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

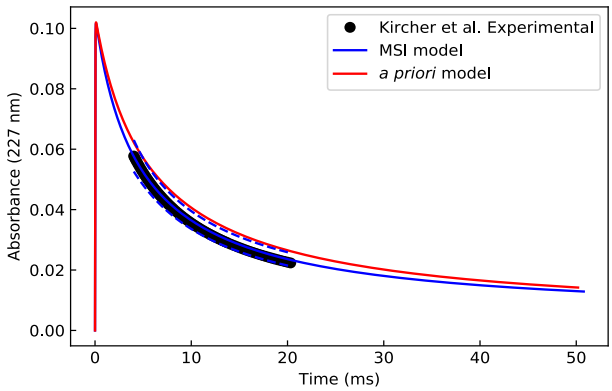


Fig. S33. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 298 K and 0.461 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

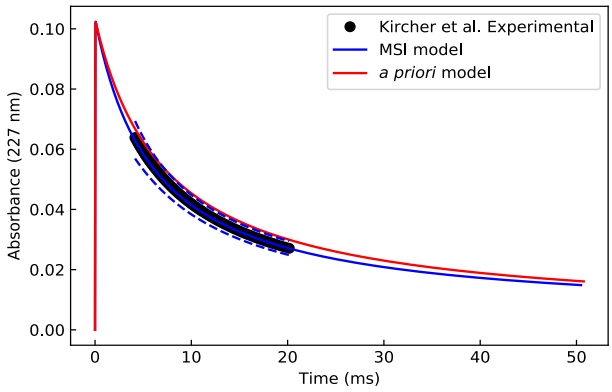


Fig. S34. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 333 K and 0.526 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

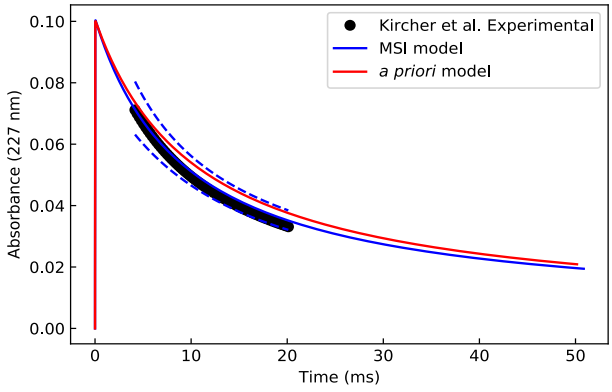


Fig. S35. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 417 K and 0.658 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

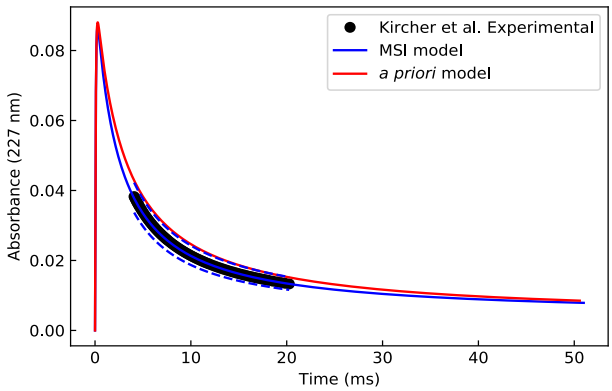


Fig. S36. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 240 K and 0.731 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

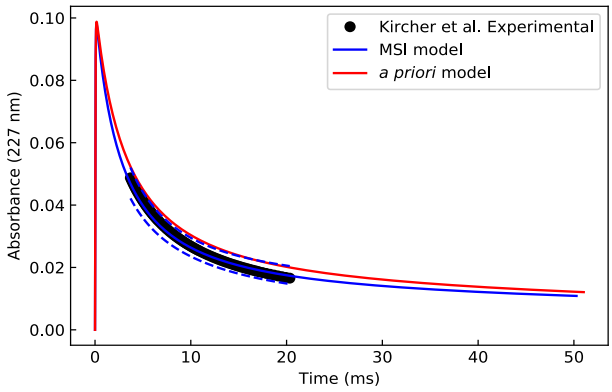


Fig. S37. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 268 K and 0.921 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

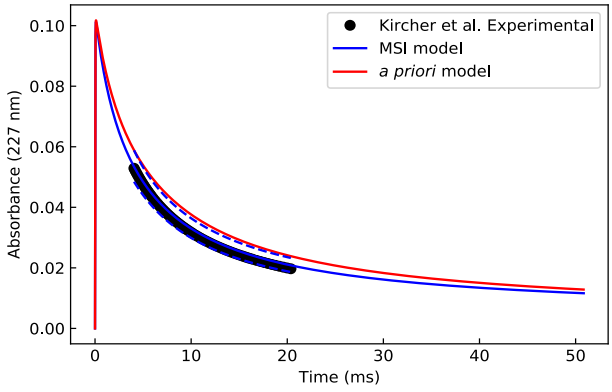


Fig. S38. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 298 K and 0.921 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

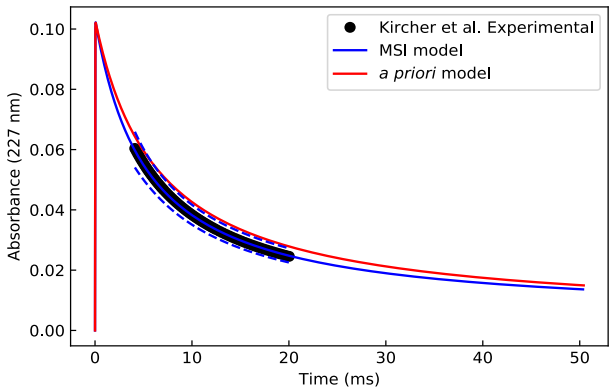


Fig. S39. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 333 K and 1.028 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

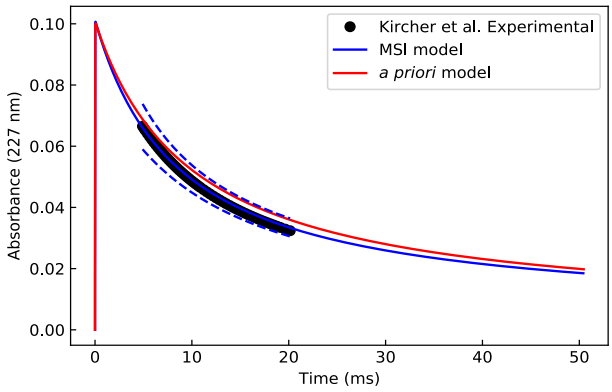


Fig. S40. Absorbance 227 nm time-history in shock heated $Cl/Cl_2/CH_3OH/O_2/Ar$ mixture near 417 K and 1.305 atm. Symbols represent generated experimental data from Kircher et al. [S6]; lines model predictions.

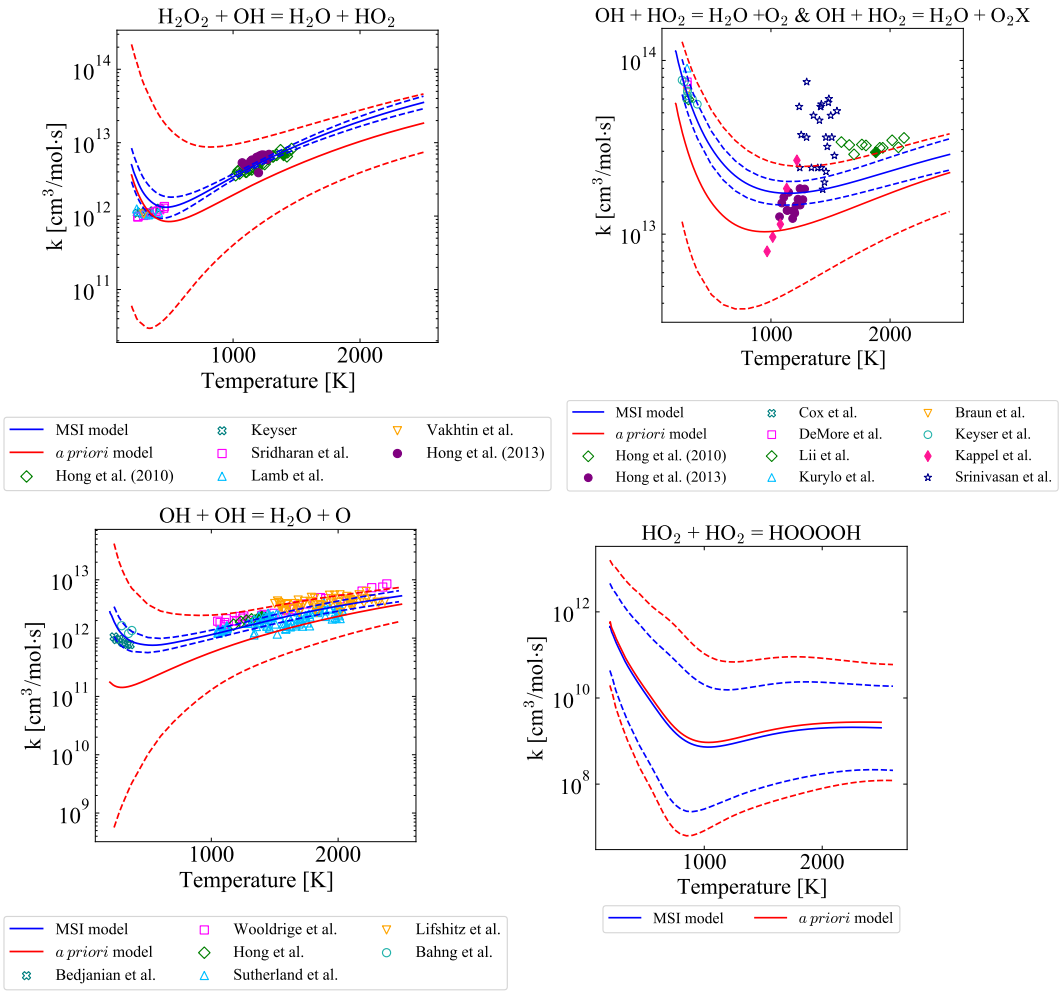


Fig. S41. Comparisons of model predictions (red lines: *a priori* model with error bars constrained by target class I only; blue lines: MSI model with error bars constrained by target classes I-IV) and experimental data (symbols) for rate constants at 1 atm. Top left: R3 [S1, S5, S7-S20], Top Right: R5, R6 [S5, S3-S4, S11-S18], Bottom Left: R4 [S19-S23]. Bottom Right: R2b.

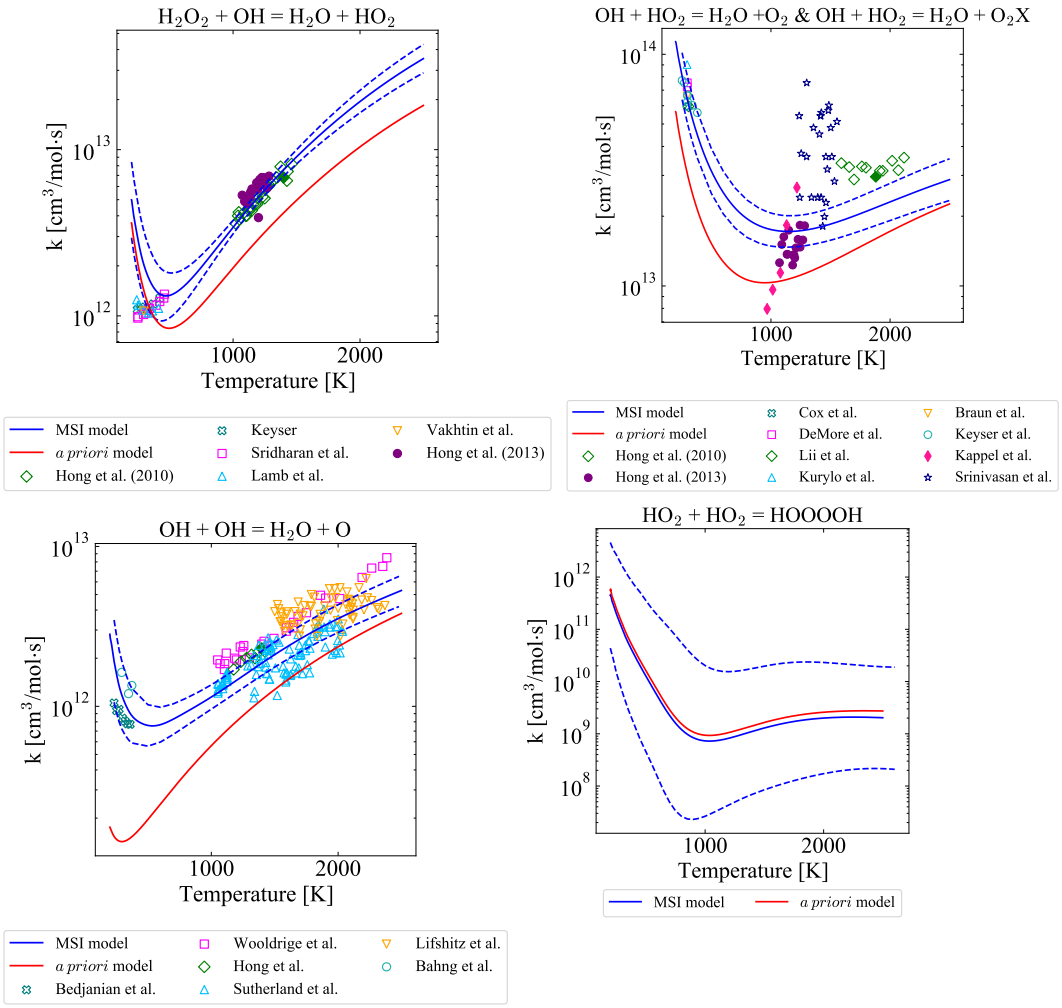


Fig. S42. Duplicate of S41 without prior error bars for higher resolution on the Y-axis.

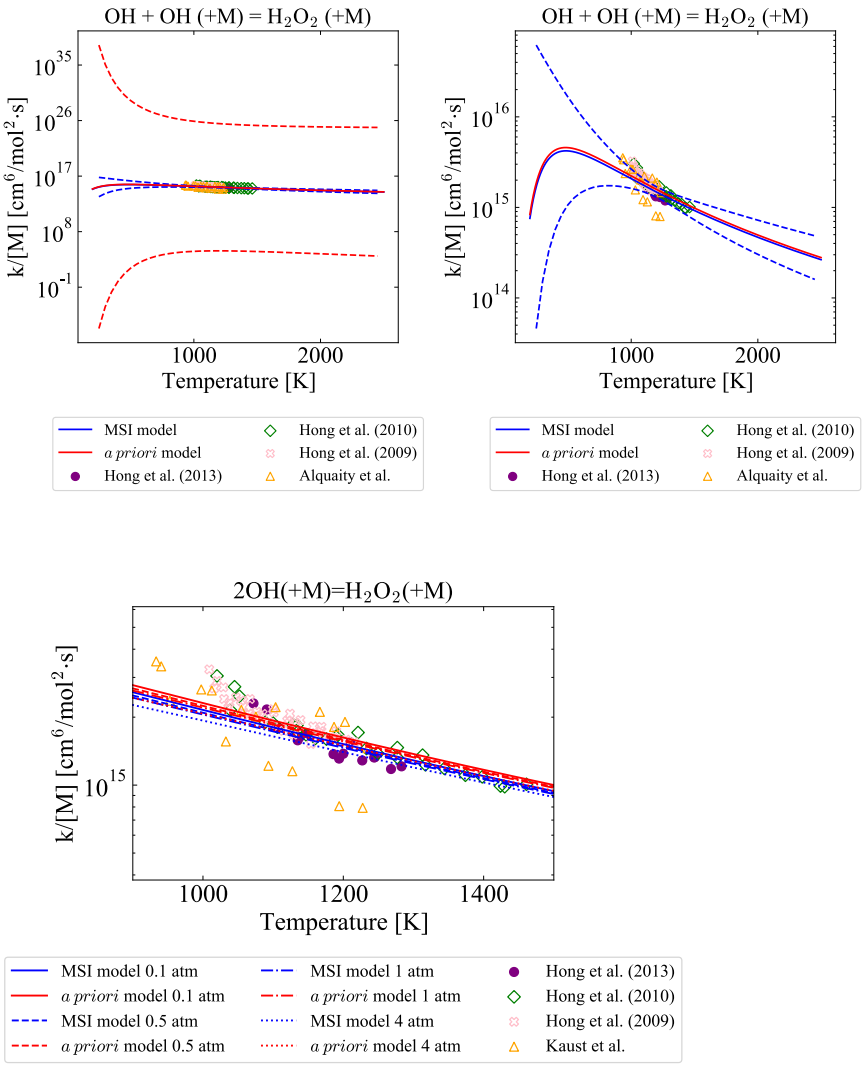


Fig. S43. Comparisons of model predictions (red lines: *a priori* model with error bars constrained by target class I only; blue lines: MSI model with error bars constrained by target classes I-IV) and experimental data (symbols) for rate constants at 1 atm (top two figures) and 0.1-4.0 atm (bottom figure). R7 [S1, S2, S5, S24].

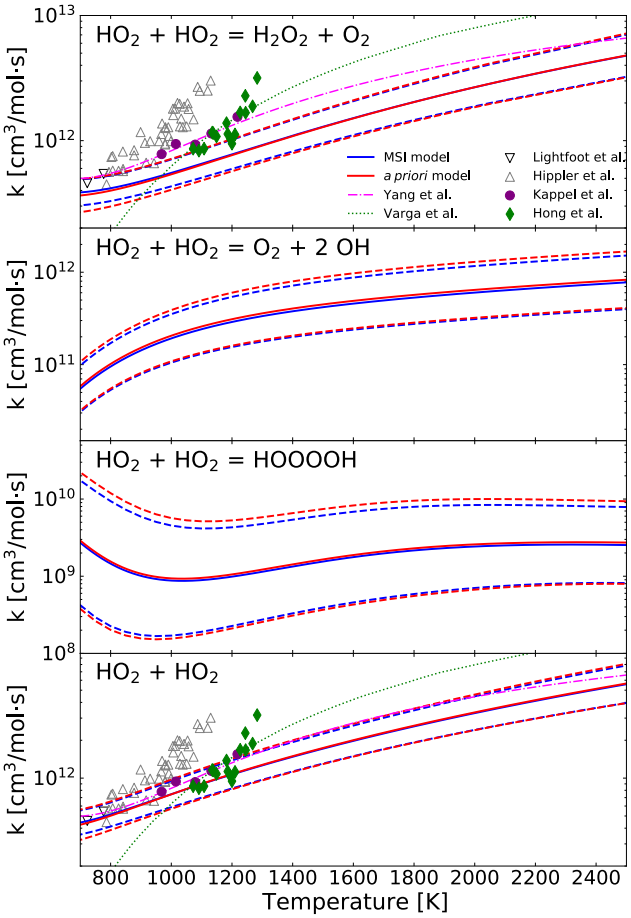


Fig. S44. Duplicate of Fig. 14 in main text with addition of the rate constants obtained by data analysis studies of Yang et al. and Varga et al. [S26, S27].

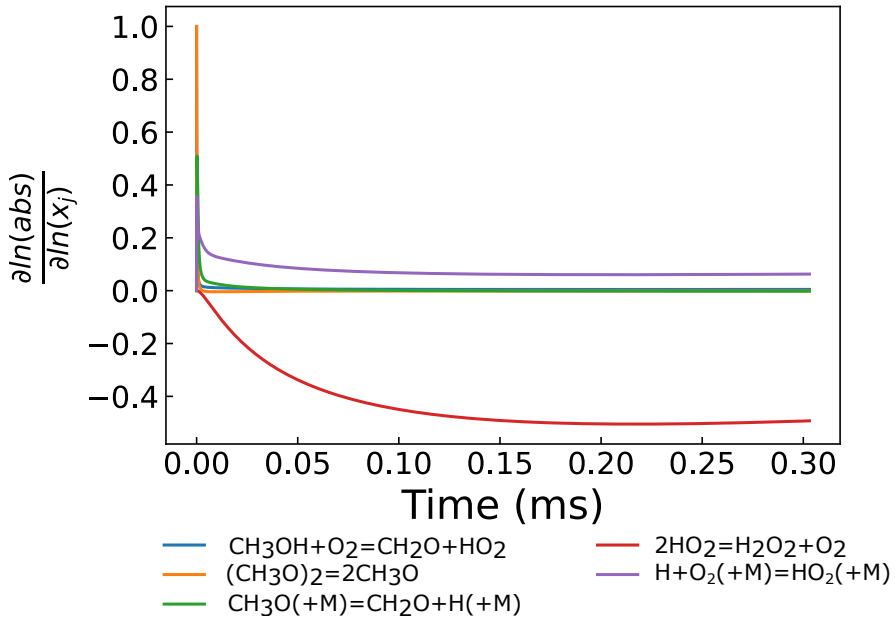


Fig. S45. Kinetic sensitivity analysis of absorbance at 230 nm for shock heated $(CH_3O)_2/O_2/Ar$ mixture near 970 K and 6.420 atm. Initial conditions from Hippler et al. [S24].

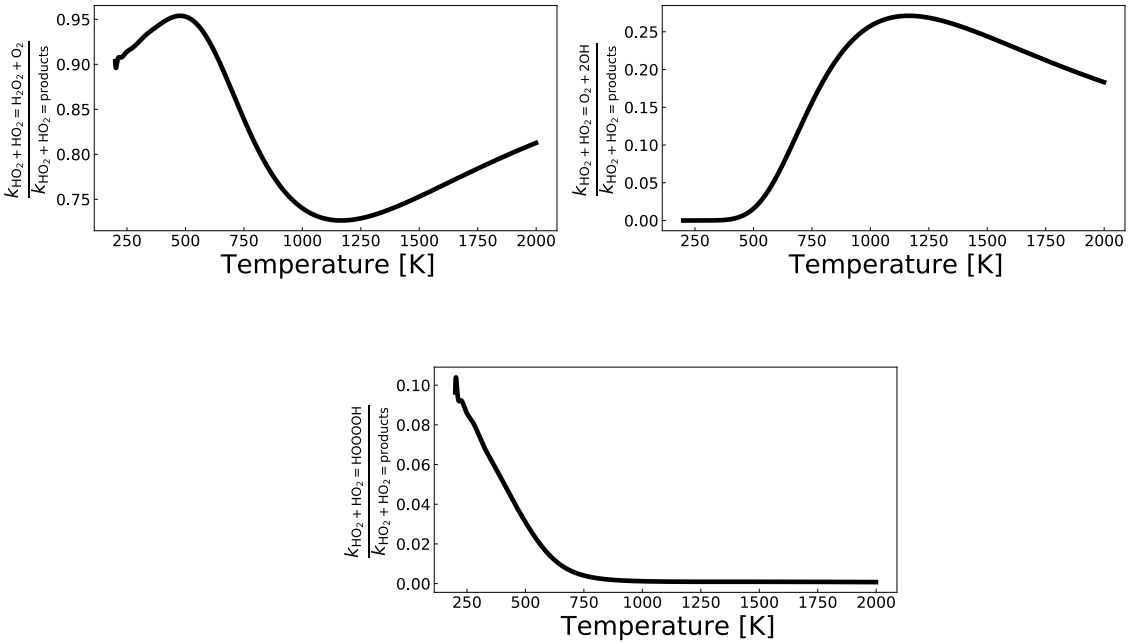


Fig. S46 Branching ratio plots at 1.0 atm. Rate constants for the channels of $HO_2 + HO_2 = H_2O_2 + O_2$ (top left), $HO_2 + HO_2 = O_2 + 2OH$ (top right) and $HO_2 + HO_2 =$ HOOOOH (bottom) normalized by the rate constant of $HO_2 + HO_2 =$ all products.

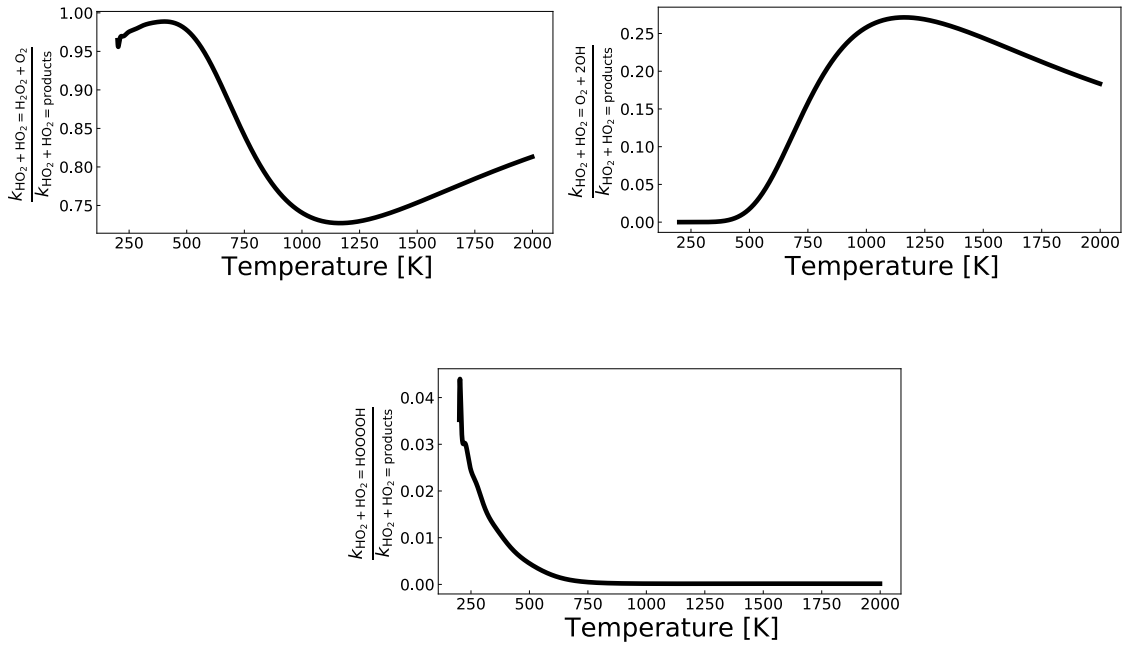


Fig. S47 Branching ratio plots at 0.10 atm. Rate constants for the channels of $HO_2 + HO_2 = H_2O_2 + O_2$ (top left), $HO_2 + HO_2 = O_2 + 2OH$ (top right) and $HO_2 + HO_2 =$ HOOOOH (bottom) normalized by the rate constant of $HO_2 + HO_2 =$ all products.

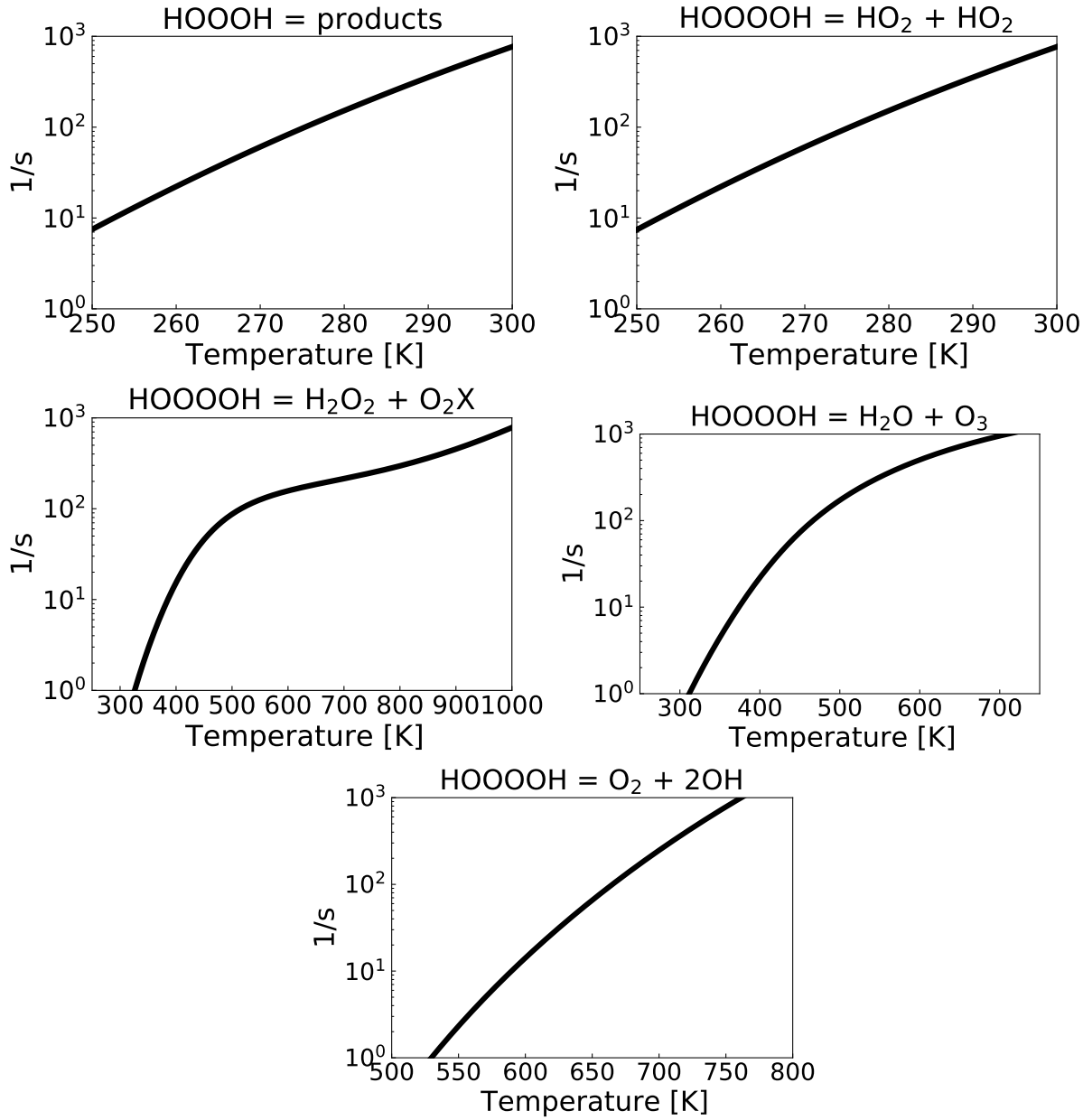


Fig. S48 Rate constants for HOOOOH at 1 atm.

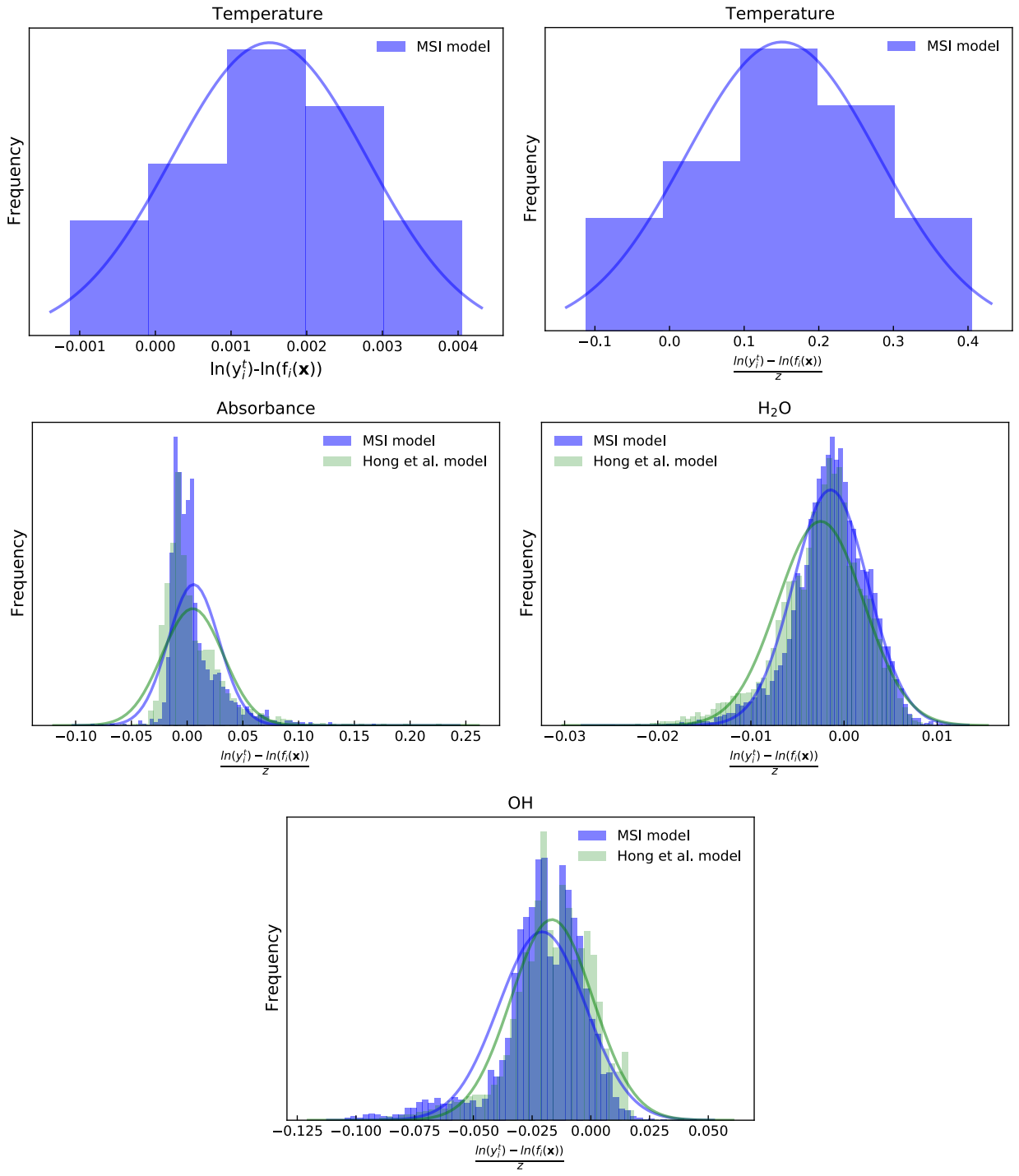


Fig. S49 Distribution of ln difference between experimental data and model for temperature (weighted by uncertainty and not weighted by uncertainty) OH, H_2O, and absorbance (all weighted by uncertainty) for the experimental profiles from Hong et al. [S23].

(c) Optimized Rate Constant Results
Chebyshev Fits

Chebyshev Expression: $\log k(T,P) = \Sigma_{t=1}^{N_T} \Sigma_{p=1}^{N_P} \alpha_{tp} \phi_t(\tilde{T}) \phi_p(\tilde{P})$

$$\tilde{T} = \frac{2T^{-1} - T_{min}^{-1} - T_{max}^{-1}}{T_{max}^{-1} - T_{min}^{-1}}$$

$$\tilde{P} = \frac{2logP - logP_{min} - logP_{max}}{log\ P_{max} - log\ P_{min}}$$

Chebyshev fits are for 0.1-4 atm. These are the preferred fits for implementation.
Units are cm, s, mol.

chebyshev_reaction("HO2 + HO2 <=> H2O2 + O2',
Tmin=200.0, Tmax=3000.0,
Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
coeffs=[[12.37, -0.1174, 0.2525],
[-0.3931, 0.1334, -0.3123],
[0.4154, 0.023, 0.0101],
[0.3355, -0.0543, 0.0657],
[0.1126, 0.0080, -0.0032],
[0.0234, 0.0116, -0.0149],
[0.0101, -0.0037, -0.0010],
[0.0059, -0.0024, 0.0028],
[0.0018, 0.0021, 0.0004],
[0.0008, 0.0015, -0.0005],
[0.0016, -0.0019, 0.0004],
[0.0006, -0.0006, -7.0167e-05],
[0.0004, -0.0003, 0.0001],
[-0.0006, 0.00159, -0.0004],
[0.0001, 2.966e-06, 0.0001]])

```
chebyshev_reaction("HO2 + HO2 <=> O2 + 2 OH",
                   Tmin=200.0, Tmax=3000.0,
                   Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
                   coeffs=[[8.6217, -0.0023, 0.0018],
                           [3.6249, 0.0019, -0.0014],
                           [-0.2055, 0.0018, -0.0014],
                           [-0.0827, 0.0005, -0.0004],
                           [-0.0156, 0.0003, -0.0003],
                           [0.0070, 0.0001, -0.0001],
                           [0.0104, 0.0002, -0.0002],
                           [0.0079, 0.0001, -0.0001],
                           [0.0049, 1.6173e-05, -6.1260e-05],
                           [0.0025, 6.8107e-05, -9.9014e-05],
                           [0.0012, -0.0001, 4.7730e-05],
                           [0.0004, 5.8690e-06, -4.15601e-05],
                           [0.0001, -0.0002, 0.0001],
                           [2.5659e-05, -9.2687e-06, -8.8667e-06],
                           [-2.9400e-05, -0.0002, 0.0001]])

chebyshev_reaction("HOOOOH => 2 HO2",
                   Tmin=200.0, Tmax=3000.0,
                   Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
                   coeffs=[[2.6929, 2.3508, -0.2621],
                           [6.1780, -0.2502, 0.1562],
                           [-0.3134, -0.2973, -0.0018],
                           [0.1076, -0.1369, -0.0175],
                           [0.1684, -0.0857, -0.0168],
                           [0.0824, -0.0566, 0.0137],
                           [0.0018, -0.0529, 0.0211],
                           [-0.0376, -0.0481, 0.0258],
                           [-0.031, -0.0437, 0.0259],
                           [0.0081, -0.0518, 0.0129],
                           [0.0262, -0.0435, 0.0225],
                           [0.0435, -0.0477, 0.0097],
                           [0.0228, -0.0214, 0.0234],
                           [0.0192, -0.0165, 0.0054],
                           [-0.0025, 0.0048, 0.0136]])
```

```
chebyshev_reaction("HO2 + HO2 => HOOOOH",
        Tmin=200.0, Tmax=3000.0,
        Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
        Coeffs= [[8.997, 2.352, -0.2627],
            [-1.2054, -0.2474, 0.15595],
            [0.1252, -0.2951, -0.0021],
            [0.2933, -0.1349, -0.0179],
            [0.2389, -0.0838, -0.0171],
            [0.1090, -0.0553, 0.0135],
            [0.0121, -0.0521, 0.0210],
            [-0.0332, -0.0476, 0.0258],
            [-0.02867, -0.0436, 0.0259],
            [0.0090, -0.0517, 0.0129],
            [0.0269, -0.0437, 0.0226],
            [0.0438, -0.0477, 0.0097],
            [0.0232, -0.0217, 0.0236],
            [0.0193, -0.0165, 0.0053],
            [-0.0022, 0.0047, 0.0137]])

chebyshev_reaction("HO2 + OH <=> H2O + O2",
        Tmin=200.0, Tmax=3000.0,
        Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
        coeffs=[[13.6583],
            [-0.4066],
            [0.0383],
            [0.0818],
            [0.0573],
            [0.0299],
            [0.0129],
            [0.0039],
            [-0.0001],
            [-0.0009],
            [-0.0017],
            [-0.0008],
            [-0.0013],
            [-0.0002],
            [-0.0008]])
```

```
chebyshev_reaction("HO2 + OH <=> H2O + O2X",
        Tmin=200.0, Tmax=3000.0,
        Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
        coeffs=[[11.3250],
            [0.6841],
            [0.4131],
            [0.1552],
            [0.0633],
            [0.0285],
            [0.0130],
            [0.0058],
            [0.0025],
            [0.0011],
            [0.0004],
            [0.0002],
            [1.4968e-05],
            [6.2679e-05],
            [-9.8156e-05]])

chebyshev_reaction("OH + OH <=> H2O + O",
        Tmin=200.0, Tmax=3000.0,
        Pmin=(0.01, "atm"), Pmax=(100.0, "atm"),
        coeffs=[[12.2347],
            [0.0173],
            [0.3487],
            [0.1588],
            [0.0646],
            [0.0335],
            [0.0109],
            [0.0056],
            [0.0019],
            [-8.7600e-05],
            [0.0003],
            [-0.0008],
            [0.0018],
            [0.0006],
            [0.0028]])
```

Arrhenius Fits

Arrhenius fits are for 1 atm, maximum error of Arrhenius fits is ~10%.
Units are cm, s, mol.

reaction("HO2 + HO2 <=> H2O2 + O2" , [7.93143247e+02, 2.85418458e+00, -9.229686e+02] , options = 'duplicate')
reaction("HO2 + HO2 <=> H2O2 + O2", [1.40287606e+15, -1.46599397e+00, -8.735761e+02] , options = 'duplicate')

reaction("HO2 + HO2 <=> O2 + 2 OH", [2.64940286e+17, -1.42708413e+00, 8.4919345e+03])

200 - 800 K
reaction("HOOOOH => 2 HO2", [1.82077302e+38, -8.77288922e+00, 1.86827733e+04])

8000 - 2000 K
reaction("HOOOOH => 2 HO2", [3.32757750e+03, 1.97317861e+00, 6.84335587e+03], options= 'duplicate')
reaction("HOOOOH => 2 HO2", [.32760075e+03, 1.97320171e+00, 6.85308080e+03], options= "duplicate")
2000 - 2500 K
reaction("HOOOOH => 2 HO2" , [2.21586601e+37, -7.16781883e+00, 3.77423714e+04])

200 - 800 K
reaction("2 HO2 => HOOOOH", [6.21428758e+26, -5.95332823e+00, 1.29950489e+03])
800 - 2000 K
reaction("2 HO2 => HOOOOH", [1.71944917e+23 -3.57732688e+00, 1.88652135e+04], options= "duplicate")
reaction("2 HO2 => HOOOOH", [2.78869216e+7, 1.04062375e-04, -6.20644076e+03]], options= "duplicate")
2000 - 2500 K
reaction("2 HO2 => HOOOOH", [5.67514847e+28 -5.09187197e+00 2.34279998e+04])

reaction("HO2 + OH <=> H2O + O2", [1.23646897e+09, 1.22686493e+00, -2.90970210e+03] , options = "duplicate')
reaction("HO2 + OH <=> H2O + O2", [-5.75989229e+07, 1.42858415e+00, -3.771099e+03] , options = "duplicate")

reaction("HO2 + OH <=> H2O + O2X", [2111.33923045, 2.67589343, -1297.31029924])

reaction("2 OH <=> H2O + O", [8.81776798e+03, 2.51974969e+00, -2.57540034e+03])

falloff_reaction("2 OH (+M) <=> H2O2 (+M) ",
kf = [3.49105E+12, 0.2048623456878356, 1503.0003607276622],
kf0 = [3.93967E+25, -3.2111376543121644, 2713.0003607276617],
efficiencies = "CO2:3.0 H2O:10.0 HE:0.4 N2:2.0",
falloff = Troe(A = 0.55, T3 = 1e-30, T1 = 1.0000000000000002e+30, T2 = 1e+30))

(d) Additional Reactions Included in the Mechanism

Reactions From IUPAC:

1. $Cl + CH_3OH = HCl + CH_2OH$
2. $Cl + H_2O_2 = HCl + HO_2$
3. $Cl + CH_2O = HCl + HCO$
4. $Cl + O_2 (+M) = ClOO (+M)$
5. $Cl + ClOO = Cl_2 + O_2$
6. $Cl + ClOO = ClO + ClO$
7. $Cl + CH_2OH = HCl + CH_2O$
8. $OH + Cl_2 = HOCl + Cl$
9. $Cl + HO_2 = ClO + OH$
10. $Cl + HO_2 = HCl + O_2$
11. $HO_2 + ClO = HOCl + O_2$
12. $OH + ClO = Cl + HO_2$
13. $OH + ClO = HCl + O_2$
14. $OH + HCl = H_2O + Cl$
15. $ClO + ClO = 2Cl + O_2$
16. $ClO + ClO = ClOO + Cl$
17. $ClO + ClO (+M) = Cl_2O_2 (+M)$

Reactions From Konnov:

1. $O_3 + O = O_2X + O_2$
2. $O + O + M = O_2X + M$
3. $O_2X + M = O_2 + M$
4. $O_2X + O = O_2 + O$
5. $O_2X + H = O_2 + H$
6. $O_2X + O + M = O + O_2 + M$
7. $O_2X + O_3 = O_2 + O_2 + O$
8. $OX + O_2X = O + O_2$
9. $OX + O_2 = O + O_2X$
10. $OX + O_2 = O + O_2$
11. $OX + M = O + M$
12. $OX + N_2 = O + N_2$
13. $OX + O_3 = O_2 + O + O$
14. $OX + O_3 = O_2 + O_2$
15. $H_2 + O_2X = H + HO_2$

16. $H + O_2X = OH + O$
17. $H + O_2X + M = HO_2 + M$
18. $HO_2 + OH = H_2O + O_2X$
19. $OH + O_2X = O + HO_2$
20. $O_3 + H = OH + O_2$
21. $O_3 + OH = HO_2 + O_2$
22. $O_3 + HO_2 = OH + O_2 + O_2$
23. $H + HO_2 = H_2O + OX$
24. $OX + H_2 = OH + H$
25. $OX + H_2O = OH + OH$
26. $O + H + M = OHX + M$
27. $OHX + O_2 = OH + O_2$
28. $OHX + N_2 = OH + N_2$
29. $OHX + H_2O = OH + H_2O$
30. $OHX + H_2 = OH + H_2$
31. $OHX + OH = OH + OH$
32. $OHX + H = OH + H$
33. $OHX + O = OH + O$
34. $OHX + AR = OH + AR$
35. $OHX + H_2 = H_2O + H$
36. $OHX + O_2 = O_3 + H$
37. $OHX + O_2 = HO_2 + O$
38. $OHX + H_2O = H_2O_2 + H$
39. $OHX = OH + hv$

(e) References

S1. Z. Hong, R. D. Cook, D. F. Davidson, R. K. Hanson, A shock tube study of OH + H_2O_2 = H_2O + HO_2 and HO_2 + M = 2OH + M using laser absorption of H_2O and OH, J. Phys. Chem. A 114 (2010) 5718–5727.

S2. Z. Hong, A. Farooq, E. A. Barbour, D. F. Davidson, R. K. Hanson, Hydrogen peroxide decomposition rate: a shock tube study using tunable laser absorption of H_2O near 2.5 micro meters, J. Phys. Chem. A 113 (2009) 12919–12925.

S3. Z. Hong, S. S. Vasu, D. F. Davidson, R. K. Hanson, Experimental study of the rate of OH + HO_2 = H_2O + O_2 at high temperatures using the reverse reaction, J. Phys. Chem. A 114 (2010) 5520–5525.

S4. C. Kappel, K. Luther, J. Troe, Shock wave study of the unimolecular dissociation of H_2O_2 in its falloff range and of its secondary reactions, Phys. Chem. Chem. Phys. 4 (2002) 4392–4398.

S5. Z. Hong, K-Y. Lam, R. Sur, S. Wang, D. Davidson, R. Hanson, On the rate constants of OH + HO_2 and HO_2 + HO_2: A comprehensive study of H_2O_2 thermal decomposition using multispecies laser absorption, Proc. Combust. Inst. 34 (2013) 565–571.

S6. C. C. Kircher, S. P. Sander, Kinetics and the mechanism of HO_2 and DO_2 disproportionations, J. Phys. Chem. 88 (1984) 2082-2091.

S7. U.C. Sridharan, B. Reimann, F. Kaufman, Kinetics of the reaction OH + H_2O_2 = HO_2+ H_2O, J. Chem. Phys. 73 (1980) 1286–1293.

S8. L.F. Keyser, Absolute rate constant of the reaction OH + H_2O_2 = HO_2 + H_2O from 245 to 423 K, J. Phys. Chem. 84 (1980) 1659–1663.

S9. J.J. Lamb, L. Moiina, C.A. Smith, M.J. Molina, Rate constant of the OH + H_2O_2 = HO_2 + H_2O Reaction, J. Phys. Chem. 87 (1983) 4467-4470.

S10. A.B. Vakhtin, D.C. McCabe, A.R. Ravishankara, S.R. Leone, Low-temperature kinetics of the reaction of the OH radical with hydrogen peroxide, J. Phys. Chem. A 107 (2003) 10642-10647.

S11. W.B. DeMore, Reaction of HO_2 radicals with O_3 and the effect of water vapor on HO_2 kinetics, J. Phys. Chem. 83 (1979) 1113-1118.

S12. W.B. DeMore, Rate constant and possible pressure dependence of the reaction OH + HO_2, J. Phys. Chem. 86 (1982) 121–126.

S13. L.F. Keyser, Absolute rate constant of the reaction OH + H_2O_2 = HO_2 + H_2O from 245 to 423 K, J. Chem. Phys. 84 (1980) 1659–1663.

S14. R.R. Lii, R.A. Gorse, M.C. Sauer, S. Gordon, Temperature dependence of the gas-phase self-reaction of hydroperoxo radicals in the presence of ammonia, J. Phys. Chem. 84 (1980) 813–817.

S15. R.A. Cox, J.P. Burrows, T.J. Wallington, Rate coefficient for the reaction OH + HO_2= H_2O + O_2 at 1 atmosphere pressure and 308 K, Chem. Phys. Lett. 84 (1981) 217–221.

S16. M.J. Kurylo, O. Klais, A.H. Laufer, Mechanistic investigation of the OH + HO_2 reaction, J. Phys. Chem. 85 (1981) 3674–3678.

S17. M. Braun, A. Hofzumahaus, F. Stuhl, Ber. Bunsenges, VUV Flash Photolysis Study of the Reaction of HO with HO_2 at 1 atm and 298 K, J. Phys. Chem. 86 (1982) 597–602.

S18. N.K. Srinivasan, M.C. Su, J.W. Sutherland, J.V.Michael, B. Ruscic, Reflected shock tube studies of high-temperature rate constants for OH + NO_2 = HO_2 + NO and OH + HO_2 = H_2O + O_2, J. Phys. Chem. A 110 (2006) 6602–6607.

S19. M.K. Bahng, R.G. Macdonald, Determination of the Rate Constant for the OH (X2Π) + OH (X2Π) = O (3P) + H_2O Reaction over the Temperature Range 293− 373 K, J. Phys. Chem. A 111 (2007) 3850–3861.

S20. Y. Bedjanian, G. Le Bras, G. Poulet, Kinetic study of OH + OH and OD + OD reactions., J. Phys. Chem. A 103 (1999) 7017–7025.

S21. M.S. Wooldridge, R.K. Hanson, C.T. Bowman, A shock tube study of the OH + OH = H_2O + O reaction, Int. J. Chem. Kinet. 26 (1994) 389–401.

S22. J.W. Sutherland, P.M. Patterson, R.B. Klemm, Rate constants for the reaction, O (3P) + H_2O¶ OH + OH, over the temperature range 1053 K to 2033 K using two direct techniques, Proc. Combust. Inst. 23 (1990) 51–57.

S23. A. Lifshitz, J.V. Michael, Rate constants for the reaction, O + H_2O = OH + OH, over the temperature range, 1500–2400 K, by the flash photolysis-shock tube technique: A further consideration of the back reaction, Proc. Combust. Inst. 23 (1991) 59-67.

S24. H. Hippler, J. Troe, J. Willner, Shock wave study of the reaction, HO_2 + HO_2 = H_2O_2 + O_2, Confirmation of a rate constant minimum near 700 K, J. Chem. Phys. 93 (1990) 1755-1760.

S25. M.B. Sajid, Et. Es-Seebar, T. Javed, C. Fittschen, A. Farooq, Measurement of the Rate of Peroxide Thermal decomposition in a Shock Tube Using Quantum Cascade Laser Absorption Near 7.7 micro meters, Int. J. Chem. Kinet. 46 (2014) 275-284.

S26. X. Yang, X. Shen, P. Zhao, C.K. Law, Statistical Analysis on Rate Parameters of the H_2-O_2 Reaction System, J. Phys. Chem. A. 125 (2021) 10223-10234.

S27. T Varga, T. Nagy, C. Olm, I.Gy. Zsely, R. Palvolgyi, E. Valko, G. Vincze, M. Cserhati, H.J. Curran, T. Turanyi, Optimization of a hydrogen combustion mechanism using both direct and indirect measurements, Proc. Combust. Inst. 35 (2015) 589-596.