Sub Topic: Reaction kinetics

SAND2017-3929C

10th U. S. National Combustion Meeting
Organized by the Eastern States Section of the Combustion Institute
April 23-26, 2017
College Park, Maryland

Inference of H_2O_2 thermal decomposition rate parameters from experimental statistics

Tiernan A. Casey⋆*, Mohammad Khalil, Habib Najm*

Combustion Research Facility, Sandia National Laboratories, Livermore CA 94551, USA
⋆*Corresponding Author Email: tcasey@sandia.gov*

Abstract: The thermal decomposition of H_2O_2 is an important process in hydrocarbon combustion playing a particularly crucial role in providing a source of radicals at high pressure where it controls the 3rd explosion limit in the H_2-O_2 system, and also as a branching reaction in intermediate-temperature hydrocarbon oxidation. As such, understanding the uncertainty in the rate expression for this reaction is crucial for predictive combustion computations. Raw experimental measurement data, and its associated noise and uncertainty, is typically unreported in most investigations of elementary reaction rates, making the direct derivation of the joint uncertainty structure of the parameters in rate expressions difficult. To overcome this, we employ a statistical inference procedure, relying on maximum entropy and approximate Bayesian computation methods, and using a two-level nested Markov Chain Monte Carlo algorithm, to arrive at a posterior density on rate parameters for a selected case of laser absorption measurements in a shock tube study, subject to the constraints imposed by the reported experimental statistics. The procedure constructs a set of H_2O_2 concentration decay profiles consistent with these reported statistics. These consistent data sets are then used to determine the joint posterior density on the rate parameters through straightforward Bayesian inference. Broadly, the method also provides a framework for the replication and comparison of missing data from different experiments, based on reported statistics, for the generation of consensus rate expressions.

Keywords: Uncertainty quantification, Bayesian inference, Arrhenius parameters

1. Introduction

Chemical kinetic mechanisms consist of uncertain specifications from a variety of sources, such as elementary rate measurements, ab initio quantum chemistry calculations, and rate-rule constructions. The effort to understand these data sources and their consequence on combustion modeling from an uncertainty quantification (UQ) perspective has gained increased recent attention[1]. Such investigations are often limited to specific temperature ranges, meaning that the global expression in a mechanism is typically a complicated aggregation over multiple individual expressions from different experiments. The effort of assigning uncertainty to the parameters of such aggregated expressions is difficult and often results in the specification of somewhat qualitative 'recommended' uncertainties. In the context of Arrhenius fits, uncertainties are sometimes presented as temperature varying uncertainty factors on the rate constant $k(T)$, but can also appear in the form of uncertainty factors attached to the pre-exponential factor only, thus providing no statement of the possibility of variation of rate constant uncertainty with temperature, or supplied as error bounds on the individual Arrhenius parameters, thus defining a bounded volume in the parameter space without any statement of the structure of the joint parameter probability density within this volume. From the point of view of UQ, an appropriate treatment of uncertainty would involve knowledge of the

data and model error in the underlying experimental or computational signals used to generate the rate constants. This underlying data is typically unavailable or unreported in investigations of elementary reactions, with only the fitting parameters of the data and perhaps some statistics on these parameters given as a measure of the uncertainty of the investigation. A method for inferring missing data subject to constraints was proposed by Berry et al. [2], where a data space is explored in a Bayesian framework with a likelihood constructed to enforce statistics on the data, resulting in a maximum entropy-based (MaxEnt [3]) solution for the posterior distribution of the data. This method of inferring unavailable or missing data has been demonstrated for model problems with a variety of constraints on processed data [4], applied to a combustion problem to determine the rate parameter correlation structure of a single-step CH_4 reaction by inferring ignition delay data [5] by comparison to data generated from a detailed model, and applied to the imperative elementary combustion reaction $H + O_2 \longrightarrow OH + O$ by inferring the real data from a shock tube experiment subject to the reported statistics (i.e. an error measure on the evaluated rates) [6]. In this work we apply the method to infer missing data from a shock tube experiment measuring the rate of H_2O_2 decomposition and use this data to reveal the structure of the posterior PDF of the Arrhenius rate parameters that fit the data. We then utilize this PDF to generate input samples for a forward model to demonstrate the necessity of knowledge of the parameter correlation structure for performing UQ investigations of combustion processes of interest.

2. Experimental data

To investigate the thermal decomposition reaction:

$$H_2O_2 + M \longrightarrow 2\,OH + M$$

we consider the reported rate data from a recent shock tube study [7]. This set of experiments employed laser absorption to measure the time decay of H_2O_2 concentration resulting from the thermal decomposition occurring in the high temperature region behind reflected shock waves in an Ar diluent in the temperature range 930-1235 K. We restrict our study to data (i.e. H_2O_2 time profiles) at four temperatures at pressures of approximately 1-2 atm. This shock tube study reports rate constants at the various temperatures and provides a best fit Arrhenius expression with symmetric error bars attached to the representation of the pre-exponential (A) and activation energy (E_a) parameters:

$$k = 10^{16.29\pm0.12}\exp\left(\frac{-21993 \pm 301}{T}\right) \tag{1}$$

We interpret these error bars as a 2 standard deviation uncertainty measure for the purposes of enforcing constraints in our inference procedure for the missing data. The data fitting procedure performed by the experimentalists to construct the rates was carried out using a calibration chemical mechanism [8] in conjunction with a specified thermochemistry database [9]. For the purposes of performing data fitting through Bayesian inference we assume an experimental data noise model in the form of standard normally distributed additive noise, ε, with a constant nominal strength σ:

$$[H_2O_{2\text{data}}]_{i,j} = [H_2O_{2\text{model}}]_{i,j} + \sigma\varepsilon_{i,j} \tag{2}$$

where the indices $\{i, j\}$ span all the data points of the experiment in temperature and temporal space respectively, and ε is independently and identically distributed across the indices $\{i, j\}$.

3. Bayesian inference

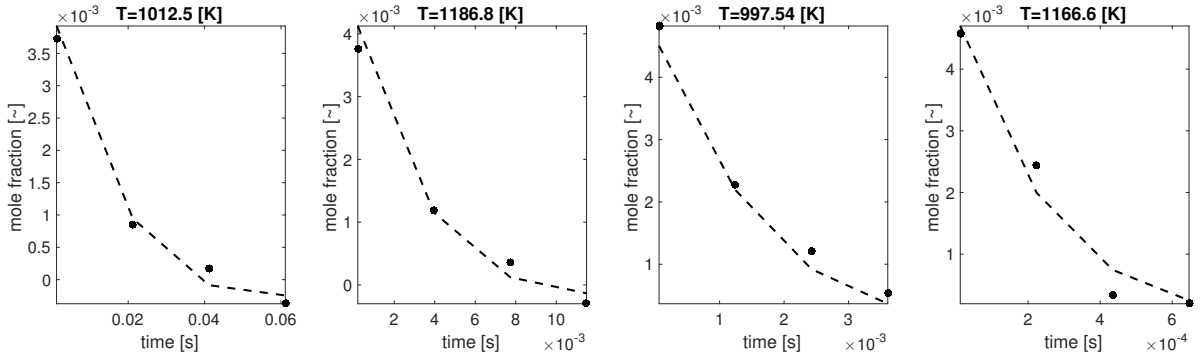


Figure 1: Sample data sets for the four experimental temperatures considered. The data are H_2O_2 concentration time profiles. The dashed lines are data synthesized using the forward model (ODE system) using the nominal reported values of the uncertain rate expression (eq. 1), while the circles are an example of the data (**D**) produced by the inference procedure.

The inference method proceeds as follows. An inference on the experimental data is constructed in a Bayesian framework (eq. 3) which proposes data, **D** (see Fig. 1), and measures its consistency with the experimentally reported statistics, S,

$$p(\mathbf{D}|S) = \frac{p(S|\mathbf{D})\pi(\mathbf{D})}{p(S)} \tag{3}$$

in the context of approximate Bayesian computation (ABC)[10]. An ABC likelihood, $p(S|\mathbf{D})$, is constructed using Gaussian error kernels (eq. 4) which penalize proposed data whose statistics ($S_{\mathbf{D}}$) do not match those of the experiment ($S_{\text{experiment}}$).

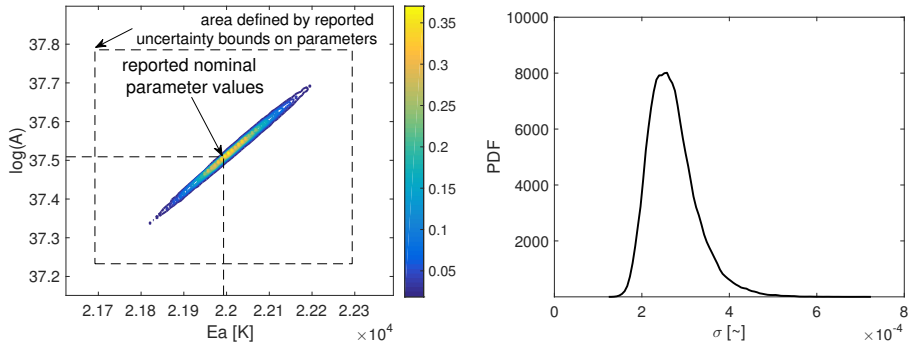
$$p(S|\mathbf{D}) = \prod_{i=1}^{4} \exp\left[-\delta_i \left(\frac{S_{i,\mathbf{D}} - S_{i,\text{experiment}}}{S_{i,\text{experiment}}}\right)\right] \tag{4}$$

The δ_i values in the ABC likelihood expression represent the degree to which consistency is demanded for each statistic. The ABC likelihood requires the determination of statistics on the Arrhenius parameters of the proposed data, which are produced by performing a second inner Bayesian inference (eq. 5)

$$p(\lambda|\mathbf{D}) = \frac{p(\mathbf{D}|\lambda)\pi(\lambda)}{p(\mathbf{D})} \tag{5}$$

on the Arrhenius parameters using the assumed error model described in eq. 2. λ in eq. 5 is the vector of inferred parameters consisting of the Arrhenius parameters of the rate constant and the data error strength parameter (i.e. error standard deviation), σ. Equation 6 shows the parameter likelihood construction:

$$p(\mathbf{D}|\lambda) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\left([\mathrm{H_2O_2}_{\text{data}}]_{i,j} - [\mathrm{H_2O_2}_{\text{model}}]_{i,j}\right)^2}{2\sigma^2}\right) \tag{6}$$



(a) (b)

Figure 2: (a) Posterior parameter PDF indicating the correlated joint density between the pre-exponential and activation energy, compared to the nominal values and reported uncertainty bounds from the experiment, (b) marginal PDF of the standard deviation of the data error.

The nested inferences are performed using Markov Chain Monte Carlo (MCMC) sampling, with single-site Metropolis sampling employed for the outer data chain and adaptive Metropolis sampling used for the inner parameter chain. For a data vector $\mathbf{D}$ (representing the vector of all experimental data points across all temperatures), a set of constraints S is applied to the Arrhenius parameters associated with the data. In this case S is the mean and standard deviation of the Arrhenius parameters (i.e. four statistics). $\pi(\mathbf{D})$ is a prior PDF on the data vector.

As each inner parameter chain must be run long enough to generate sufficient samples to calculate converged statistics on those parameters, and the evaluation of the parameter likelihood at each point on the inner MCMC chain involves the solution of an ODE system, considerable computational effort is required. This is further compounded by the fact that this effort constitutes only a single evaluation of the outer data chain likelihood. To accelerate the evaluation of the ODE system, a surrogate model is constructed using a polynomial chaos expansion (PCE) to represent the model output (i.e. H_2O_2 concentration as a function of time, see Figs. 3a, 3b) with a stochastic dimension reflecting the uncertain rate constant of interest, using a separate PCE to construct a surrogate for each experimental temperature considered.

On completion of the outer MCMC data chain, the accepted data sets that are consistent with the statistics are pooled and a final parameter inference is performed to compute samples from the pooled posterior parameter PDF. Figure 2a shows the results of this sampling procedure, indicating a strong degree of correlation between the Arrhenius parameters (presented as log(A) and E_a) in the joint posterior density, in stark contrast to the uniform uncertainty space defined by the symmetric error bars applied to the parameters from the reported experimental fit. The marginal PDF of the data error standard deviation is also shown which represents the most likely noise amplitude (*iid* Gaussian) consistent with the reported statistics.

4. Application to forward UQ in combustion modeling

To demonstrate the importance of having knowledge of the joint parameter distribution we perform an uncertainty study of a system relevant to this reaction, namely the 3rd explosion limit of H_2-O_2 mixtures. At high pressures, the decomposition of the H_2O_2 originating from the reaction of an increasing concentration of relatively inactive HO_2 with abundant H_2 provides a sufficient radical population (i.e. OH) to compensate for radical destruction at solid walls to render mixtures

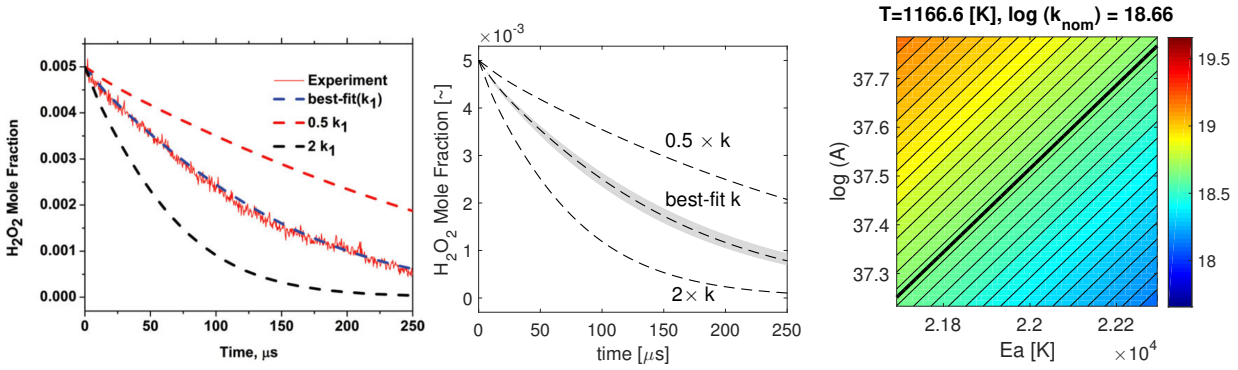


Figure 3: H_2O_2 decay profiles for (T = 1166.6K), comparing the results of the experiment (a) (reproduced from [7]) to the inference (b). The shaded area in (b) is the 97% confidence interval of the inferred data. (c) Contours of the log of the associated rate constant, with the iso-contour of the nominal rate constant (k_{nom}) shown with a solid black line.

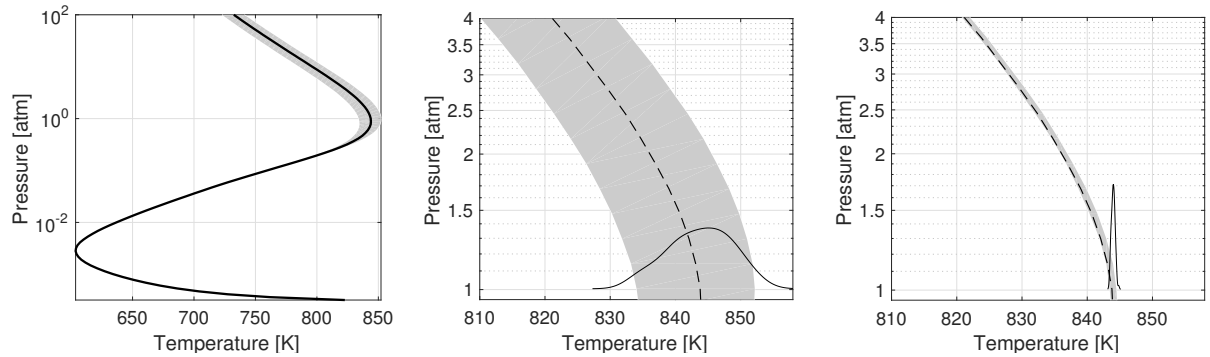


Figure 4: (a) Explosion limit Z-curve for stoichiometric H_2-O_2 mixtures. The solid black curve uses nominal rate parameter values and the shaded area is the 95% confidence interval using uniform sampling within the reported uncertainty bounds, (b) 1-4 atm pressure range with uniform sampling within reported bounds, (c) 1-4 atm pressure range using samples from the correlated joint posterior resulting from the inference. The dashed lines in (b) and (c) are predictions using the nominal rate parameters and the shaded areas represent the 95% confidence interval of the samples. 1D-PDFs of the explosion temperature are overlayed for 1 atm pressure conditions in both (b) and (c).

explosive. To generate the characteristic Z-shaped explosion limit boundary, we perform homogeneous ignition calculations for a range of pressures and employ a binary classification approach to determine the temperature boundary that separates explosion from non-explosion. The classification procedure iterates until the limit boundary is isolated to within an arbitrary threshold of 0.1 K. Figure 4a shows the explosion limit curve calculated using the nominal rate parameters as well as the 95% confidence interval for variations of the parameters within the reported uncertainty bounds, indicating that the H_2O_2 decomposition reaction begins to influence the curve at pressures above 0.2 atm. Comparing evaluations of the explosion limit model using input samples taken from the uniform space defined by the uncertainty bounds (Fig. 4b) versus sampling from the correlated joint density (Fig. 4c) demonstrates the drastic over-prediction of the uncertainty in the location of the explosion limit boundary in temperature space using uniform sampling. This arises from the over and under-prediction of the value of the rate of the H_2O_2 decomposition reaction when uniform sampling chooses Arrhenius parameter value pairs in directions transverse to the direction of correlation, which correspond to the largest departures from the nominal rate value (Fig. 3c).

5. Conclusion

We present the application of a data inference procedure to generate missing experimental data from a shock tube experiment consistent with the statistics reported in the experiment in the form of error bounds on Arrhenius parameters. The joint posterior PDF on the Arrhenius parameters conditioned on the inferred data indicates strong correlation between log(A) and E_a. Samples from this correlated joint PDF are propagated through a forward model for predicting the explosion limit boundary in the H_2-O_2 system, and compared to the result of sampling uniformly from the bounded uncertainty region defined by the reported error bars, highlighting the necessity of the correlated density for forward UQ. Besides delivering the joint PDF on the parameters of a reaction, the data inference procedure presents the possibility of generating improved (i.e. more likely supported by data) nominal Arrhenius parameters from the maximum likelihood estimate of the posterior. A future direction of effort is aggregation of inferred data from multiple experiments spanning different temperature ranges to generate consensus rate expressions for use in chemical mechanisms by performing inference on the data sets pooled from different experiments, in contrast to the current practice of aggregating the rate expressions themselves.

6. Acknowledgements

This work was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences. Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

References

[1] H. Wang, D. A. Sheen, Progress in Energy and Combustion Science 47 (2015) 1–31.

[2] R. Berry, H. Najm, B. Debusschere, H. Adalsteinsson, Y. Marzouk, Journal of Computational Physics 231 (2012) 2180–2198.

[3] E. Jaynes, Probability Theory: The Logic of Science, G.L. Bretthorst, Ed, Cambridge University Press, Cambridge, UK, 2003.

[4] K. Chowdhary, H. N. Najm, Statistics and Computing 26 (2016) 149–169.

[5] H. N. Najm, R. D. Berry, C. Safta, K. Sargsyan, B. J. Debusschere, Int. J. Uncertainty Quantification 4 (2014) 111–132.

[6] M. Khalil, K. Chowdhary, C. Safta, K. Sargsyan, H. Najm, Proceedings of the Combustion Institute 36 (2017) 699 – 708.

[7] M. B. Sajid, E. Es-sebbar, T. Javed, C. Fittschen, A. Farooq, International Journal of Chemical Kinetics 46 (2014) 275–284.

[8] Z. Hong, D. F. Davidson, R. K. Hanson, Combustion and Flame 158 (2011) 633–644.

[9] R. Kee, F. Rupley, J. Miller, UC-4, Sandia National Laboratories (1987).

[10] M. A. Beaumont, W. Zhang, D. J. Balding, Genetics 162 (2002) 2025–2035.