

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/49829362>

# Global Sensitivity Analysis of Chemical-Kinetic Reaction Mechanisms: Construction and Deconstruction of the Probability Density Function

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY A · FEBRUARY 2011

Impact Factor: 2.69 · DOI: 10.1021/jp108017t · Source: PubMed

---

CITATIONS

21

---

READS

28

## 3 AUTHORS, INCLUDING:



Alison Tomlin

University of Leeds

143 PUBLICATIONS 2,304 CITATIONS

SEE PROFILE

# Global Sensitivity Analysis of Chemical-Kinetic Reaction Mechanisms: Construction and Deconstruction of the Probability Density Function

Michael J. Davis\*

Chemical Sciences and Engineering Division, Building 200, Argonne National Laboratory, Argonne, Illinois 60439, United States

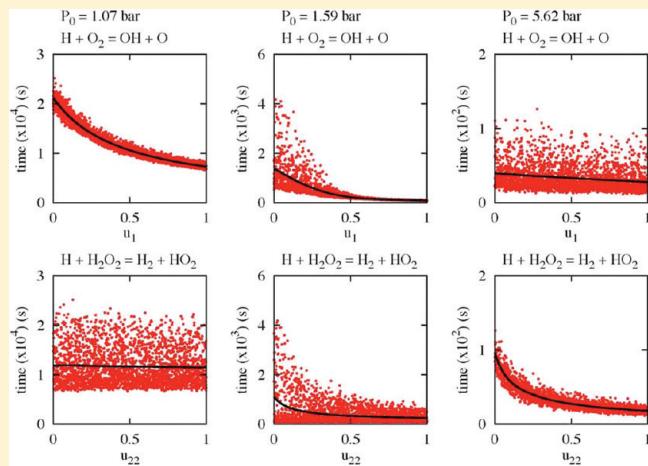
Rex T. Skodje

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309, United States

Alison S. Tomlin

School of Process, Environmental, and Materials Engineering, University of Leeds, Leeds, United Kingdom

**ABSTRACT:** This paper investigates global sensitivity analysis as applied to reaction mechanisms. It uses the HDMR (high-dimensional model representation) expansion and the features of the sensitivity indices to explore the probability density function (pdf) of predicted target outputs that results from the uncertainties in the rate coefficients. We study the autoignition of H<sub>2</sub>/O<sub>2</sub> mixtures, where the pdf describes the uncertainty in ignition delay times due to the uncertainties in the rate coefficients. The global sensitivity analysis in conjunction with the HDMR expansion allows for the deconstruction of the pdf in several different ways. These deconstructions allow the features of the pdf to be understood in terms of the constitute reactions in much finer detail than the study of a variance decomposition alone.



## I. INTRODUCTION

Chemical kinetic reaction mechanisms describing combustion processes can be as small as a few dozen reactions and as large as tens of thousands.<sup>1</sup> It is often the case that the larger mechanisms contain many reactions whose rate coefficients are not well-characterized, and even the smaller mechanisms contain many reactions whose rate coefficients can be improved significantly with further experiments and theoretical calculations. However, experiments and detailed theoretical calculations are very time-consuming. So it is important to prioritize the reactions in some manner.

The need for such prioritization has led to extensive use of uncertainty and sensitivity analysis for studying chemical-kinetic mechanisms.<sup>2</sup> Many of these applications use local versions of sensitivity analysis,<sup>3,4</sup> but there has been considerable work on global methods of sensitivity analysis<sup>5–9</sup> in chemical kinetics,<sup>10–25</sup> starting with the work of Cukier et al.<sup>10–13</sup> Applications of global sensitivity analysis are less common than their local counterparts, partly because global sensitivity analysis

is computationally expensive. However, a recent study by Ziehn et al.<sup>22</sup> has shown that it is possible to use global sensitivity analysis to study large mechanisms, which led to the work in refs 26 and 27.

References 26 and 27 used global sensitivity to study a methanol oxidation mechanism. The global sensitivity analysis indicated that two reactions contributed most to the uncertainty in ignition delay times. The two reactions were then calculated with ab initio chemical kinetics leading to their improvement and a significant reduction in the uncertainty in predicted ignition delay times for mixtures of methanol and air or oxygen. The global sensitivity analysis employed in those papers analyzed the variance of the data to decompose the overall uncertainty of the ignition delay times in terms of individual reactions,<sup>15–22</sup> but used a different algorithm than refs 15–22, one developed in ref 28.

**Received:** August 24, 2010

**Revised:** December 11, 2010

**Published:** February 11, 2011

Our continuing work on mechanism improvement aims to extend the efforts of refs 26 and 27 in two directions. One is the extension to much larger reaction mechanisms, for example, butanol.<sup>29</sup> The second extension is epitomized in this study. The purpose of this work is to study some of the features of the methods of global sensitivity analysis used in refs 26 and 27, and to extend the analysis of those papers. In this paper, we will study in detail the probability density function that describes the uncertainty of the ignition delay times induced by the uncertainties in the rate coefficients, which will be referred to here as the “ignition-time pdf”. Because the ignition delay time is the observable, the details of the ignition-time pdf reflect the uncertainty of the observable. In addition, the features of the ignition-time pdf are very sensitive to individual reactions, as will be shown in the paper, so it is an ideal function to study.

There are several reasons why the H<sub>2</sub>/O<sub>2</sub> system was chosen for study. First, it is a submechanism for all hydrocarbon combustion mechanisms and improvements of it may contribute to the improvement of many other mechanisms. The second is that it is a well-studied system and there is much already known about it, for example, that there are distinct transitions in behavior as pressure, temperature, and stoichiometry are varied. These transitions are referred to as “explosion limits”.<sup>30</sup> However, even though it is a well-studied system, there are still significant discrepancies between theoretical modeling and experiments for the H<sub>2</sub>/O<sub>2</sub> mechanism and additional study may prove useful for resolving the discrepancies.<sup>31</sup>

The hydrogen mechanism provides a very useful test case because of the transitions in its behavior as pressure, temperature, and stoichiometry are varied. The transitions are reflected in changes in the ignition-time pdf and the global sensitivity indices and therefore provide a wealth of interesting cases to study. Finally, because features of the ignition-time pdf are very sensitive to the uncertainties in the reaction rates, it is useful to have a system that is computationally tractable for large sample sizes. The H<sub>2</sub>/O<sub>2</sub> mechanism used here is relatively small (25 reactions) and satisfies this condition. Our intention is not to provide a complete analysis of the H<sub>2</sub>/O<sub>2</sub> mechanism under all conditions but rather to use it as a realistic test case to evaluate appropriate global sensitivity methods.

The outline of this paper is as follows. Section II will provide background information on the collection of data, the H<sub>2</sub>/O<sub>2</sub> system, and the equations that describe the constant-volume adiabatic ignition examples in this paper. Global sensitivity analysis is applied in section III, where it is also demonstrated that the characteristics of this analysis change significantly as the initial pressure is changed for a fixed constant initial temperature. The uncertainty in the reaction rates leads to a distribution for the ignition delay times. References 26 and 27, as well as a number of earlier papers,<sup>2,32–34</sup> have used histograms to generate the ignition-time pdf, and section IV describes a more sophisticated procedure to generate the ignition-time pdf that is designed to give maximal information. Section V proposes several ways that the ignition-time pdf may be deconstructed in terms of the individual terms of the HDMR expansion. This deconstruction allows a fine-detailed understanding of the features of the ignition-time pdf. Section VI extends the analysis of the ignition-time pdf, giving a geometrical interpretation of the ignition-time pdf based on one of the deconstructions proposed in section V. Section VII has a discussion and conclusion. The appendices give further information. Appendix A presents the uncertainties used in this paper, Appendix B shows tests of

the convergence of the algorithm, and Appendix C expands the information about the construction of the ignition-time pdf described in section IV. Appendix D expands the discussion of the geometric interpretation of section VI with a simple model that describes the significant features of the ignition-time pdf.

## II. BACKGROUND

**A. H<sub>2</sub>/O<sub>2</sub> Mechanism.** The mechanism of Li et al.<sup>35</sup> is used here. It is a submechanism of the methanol/formaldehyde mechanism developed in ref 36 that was studied previously in refs 26 and 27. The reactions for the H<sub>2</sub>/O<sub>2</sub> mechanism are enumerated in Appendix A. The rate parameters are listed in ref 35. Associated with all reactions are uncertainties, which are listed and described in Appendix A.

The current paper investigates in detail some of the procedures implemented in refs 26 and 27. Because the H<sub>2</sub>/O<sub>2</sub> system is relatively small, these procedures can be studied in much greater detail than for methanol. Even though the H<sub>2</sub>/O<sub>2</sub> mechanism is well characterized experimentally and theoretically, there are still experiments that cannot be modeled accurately with existing mechanisms<sup>37</sup> and there are many different mechanisms that may lead to different results.<sup>38</sup>

In the present paper it is shown that the “sensitivity spectrum” changes as the initial pressure is changed. In particular, at low pressure the most sensitive reaction is H + O<sub>2</sub> = OH + O and at high pressure it is H + H<sub>2</sub>O<sub>2</sub> = H<sub>2</sub> + HO<sub>2</sub>. This transition is analogous to the transition from the second to third explosion limit that is well-known in hydrogen combustion.<sup>30</sup> It is shown below that as this transition occurs, complexities arise in the global sensitivity analysis.

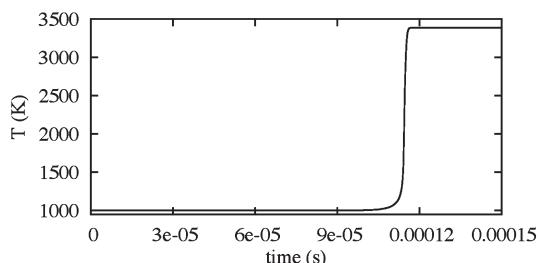
**B. Constant-Volume Adiabatic Ignition.** All results presented in this paper are run under constant volume adiabatic conditions, one of the ways of testing the sensitivity of a mechanism to the rate parameters.<sup>39</sup> It is also a common way to model many chemical-kinetics experiments, such as shock tubes and rapid compression machines.<sup>40</sup> Under these conditions, the time development of each species is followed by integration of the following set of equations for the species and temperature:<sup>39</sup>

$$\frac{dY_k}{dt} = \frac{W_k \dot{\omega}_k}{\rho} \quad (2.1a)$$

$$\frac{dT}{dt} = -\frac{1}{\rho c_v} \sum_k W_k \dot{\omega}_k e_k \quad (2.1b)$$

In eq 2.1a  $Y_k$  refers to the mass fraction of a species,  $W_k$  its molecular weight,  $\rho$  is the bulk density, and  $\dot{\omega}_k$  refers to the production rate of the  $k$ th species, as calculated from the law of mass action. In eq 2.1b  $c_v$  refers to the bulk constant volume heat capacity in mass units,  $e_k$  is the internal energy of the  $k$ th species, with the other quantities the same as they are in eq 2.1a. The system of equations is closed with the ideal gas law. These equations describe combustion in a closed vessel, assuming no heat exchange with the outside and no significant wall reactions.

A typical example of the results of these equations is shown in Figure 1, for a stoichiometric H<sub>2</sub>/O<sub>2</sub> mixture started at  $T_0 = 1000$  K and  $P_0 = 1.07$  bar, for the system described in Appendix A. The ignition is characterized by a very steep increase in the temperature starting at about  $t = 0.11$  ms. In the rest of the paper the ignition delay time  $\tau$  is defined as the time it takes to reach a temperature of 1800 K.



**Figure 1.** Trajectory run under constant volume adiabatic conditions (eqs 2.1a and 2.1b). It is run for the system in Appendix A. It is initially a stoichiometric mixture of H<sub>2</sub> and O<sub>2</sub> with  $P_0 = 1.07$  bar and  $T_0 = 1000$  K.

### III. GLOBAL SENSITIVITY ANALYSIS

Reference 27 has a detailed description of our implementation of global sensitivity for ignition delay times. The purpose of this section is to present results specifically related to H<sub>2</sub>/O<sub>2</sub>, and it provides enough information about our implementation of global sensitivity analysis to make the present paper self-contained. Reference 27 should be consulted for further details not contained here and in other sections of this paper.

**A. Setup.** The result presented in Figure 1 is for the nominal values of the rate coefficients presented in ref 35. Section II A discusses the uncertainties associated with the reaction rates, as enumerated in Appendix A. In this paper and in previous ones,<sup>26,27</sup> the rate coefficients are varied by changing the A-factors in the generalized Arrhenius form of the rate expression (see, for example, ref 3) over the range of their uncertainties:

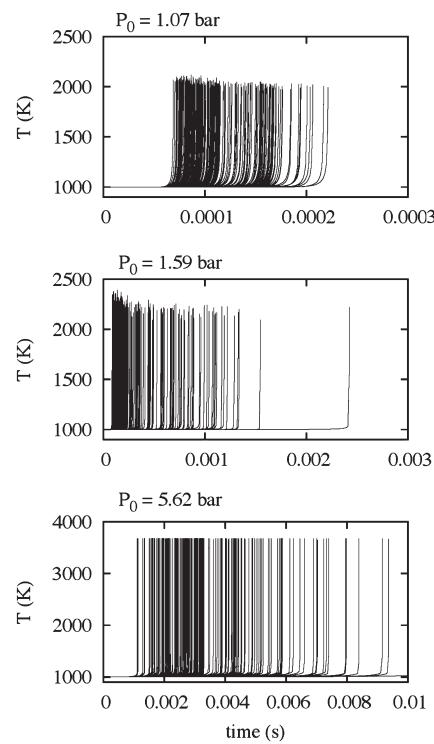
$$A_c = \beta A_0 \quad (3.1a)$$

where  $A_c$  is the nominal value of the A-factor,<sup>35</sup> and  $A_c$  is the current value of the A-factor used in the runs. It is also convenient to define a progress variable for the uncertainty range:

$$\beta_i = \frac{1}{f_i} + u_i \left( f_i - \frac{1}{f_i} \right), \quad 0 \leq u_i \leq 1, \quad (3.1b)$$

where  $f$  refers to an uncertainty factor (Appendix A). When the uncertainties are incorporated into the analysis of the ignition properties as described in the previous subsection for Figure 1, the following picture emerges as demonstrated in Figure 2. Once again these are results for initial stoichiometric mixtures of H<sub>2</sub> and O<sub>2</sub> at an initial temperature of  $T = 1000$  K. Ignition results for a random sample of 200 realizations of the rate coefficients over their uncertainty ranges are shown in all the panels of this figure. The initial pressure of the mixture is shown on each panel. As noted in the discussion of Figure 1, the ignition delay time is defined as the time it takes to reach the temperature  $T_{ign} = 1800$  K. Unlike Figure 1, the simulations are not run to completion, but until shortly after the ignition temperature has been reached.

The uncertainty factor “ $f$ ” for each reaction is listed in Appendix A. A uniform sampling of the range in eq 3.1b is used.<sup>27</sup> The 25 A-factors for each of the 200 runs in Figure 2 were generated by taking a point along the interval in eq 3.1b, whose location is chosen from a random number. This means that each of the 200 runs involved the sampling of 25 random numbers to choose the rate coefficients. Figure 2 demonstrates that there is a range of ignition delay times as the rate coefficients are varied.



**Figure 2.** Series of ignition results for stoichiometric mixtures of H<sub>2</sub> and O<sub>2</sub> which are started at  $T = 1000$  K and three different pressures, as marked on each panel.

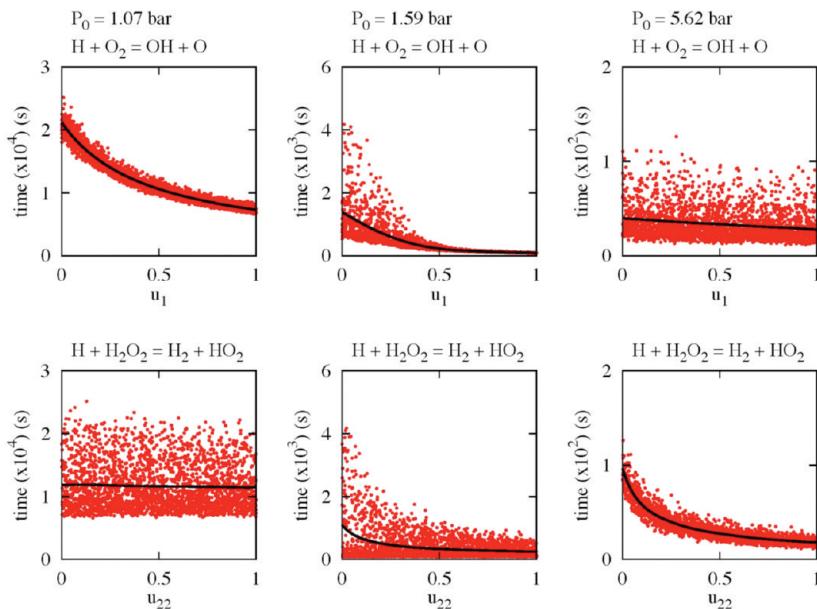
Although there is interesting information in Figure 2 related to how the uncertainties of the rate coefficients are manifest in the ignition delay time, there is no specific information about how individual reactions affect the uncertainty evident in Figure 2. To investigate such correlations, a further transformation of the information in Figure 2 is used and this is illustrated in Figure 3.

Figure 3 shows projections onto the uncertainty ranges of two of the reactions of the hydrogen mechanism over their uncertainty ranges. These plots were generated from a sample size of 250 000 randomly chosen realizations, with every 100th point plotted in Figure 3. Each column shows results for a different pressure, those already presented in Figure 2. The data points shown in each panel of a column are the same. The ordinate of each panel shows the ignition delay time and the abscissa shows the point along the uncertainty range as defined in eq 3.1b. For example, the value  $u_i = 0.0$  has a rate coefficient defined by an A-factor that is  $1/f$  of the nominal A-factor and  $u_i = 1.0$  shows results for ignition delay times generated for an A-factor which is a factor of “ $f$ ” greater than the nominal A-factor. In addition to the scatter of points showing the actual ignition delay times for the sample of 250 000 ignitions, there are black curves drawn on each panel which show a one-dimensional, eighth-order polynomial fit to the full sample of 250 000 points:

$$\tau(u_i) = \sum_{k=0}^8 a_{ik} u_i^k \quad (3.2a)$$

where  $u_i$  is a scaled progress variable defined in eq 3.1b. Each fit is done separately and the designation “ $a_{ik}$ ” indicates that the regression coefficients are distinct for each reaction.

It is important to emphasize that the scatter observed in Figure 3 is not due to random dynamical behavior. The Monte



**Figure 3.** Projections onto two of the reactions generated from random samples of 250 000 realizations for stoichiometric mixtures with  $T_0 = 1000$  K and three different pressures. The black lines are eighth-order polynomial fits used to estimate the first-order sensitivity coefficients.

Carlo sampling of rate coefficients leads to a set of results that lie on a 25-dimensional hyper-surface (25 rate coefficients) and the scatter results from projecting a 25-dimensional object onto a two-dimensional plane.

Figure 3 demonstrates that there can be strong correlations between the uncertainty and the ignition delay time for some reactions compared to others. It also shows that the strong correlation can switch from one reaction to another as a parameter is changed, which in this case is the initial pressure of the hydrogen–oxygen mixture. In the left two plots for  $P_0 = 1.07$  bar there is a strong correlation between the ignition delay time and the uncertainty of the  $H + O_2$  reaction, with little correlation between the uncertainty of the  $H + H_2O_2$  reaction and the ignition delay time. In the panels on the far right for  $P_0 = 5.62$  bar the strong correlation has switched to the  $H + H_2O_2$  reaction, with little correlation for the  $H + O_2$  reaction.

The change in behavior from the two panels on the far left to those on the far right are most likely associated with the types of transitions that occur in the  $H_2/O_2$  systems that are typically referred to as the second and third explosion limits.<sup>30</sup> Making a more explicit connection between the transition in Figure 3 and the explosion limits would be interesting but is beyond the scope of this paper.

The two panels at  $P_0 = 1.59$  bar show a transitional case. They show correlations in the uncertainty for both rate parameters, with the top one showing a distinctly stronger correlation than the bottom one. As with the other two pressures, the correlation is quantified with the polynomial fit (eq 3.2a), which is shown as a thick black line on each panel. The transitional behavior observed in the middle panels of Figure 3 is a key reason for studying the three pressures pictured there and in previous and subsequent figures in the paper.

Although the black curves drawn on the samples at  $P_0 = 1.59$  bar quantify the correlations, it is clear from the scatter of points that they do not describe the data sufficiently. The width of the scatter of points clearly changes as the uncertainty changes for these panels, with the change most obvious for the panel at the top of the middle column of Figure 3. At low values of uncertainty

in this plot the scatter of points is much broader than it is at higher values of the uncertainty factor. This behavior is different than what is observed at  $P_0 = 1.07$  bar and  $P_0 = 5.62$  bar, as pictured on the left and right columns, respectively. A more complete description of the uncertainty is clearly warranted on the basis of the behavior observed in the middle column of Figure 3 and global sensitivity analysis addresses this, with an HDMR (high-dimensional model representation) expansion being our method of choice, as described in the next subsection.

**B. HDMR Expansion.** The HDMR expansion is described in Chapters 4 and 5 of ref 6. This expansion is generally attributed to Sobol<sup>5</sup> and Rabitz and co-workers,<sup>15–18</sup> although such expansions have a long history, as discussed in ref 6. Important numerical techniques for its implementation were developed in a series of papers by Rabitz and co-workers.<sup>15–18,23</sup> A series of applications were undertaken using the HDMR expansion by one of us in collaboration with Ziehn,<sup>19–22</sup> reviewed in ref 24. A computer program, GUI-HDMR, was developed to implement the HDMR expansion using the language MATLAB.<sup>20</sup> The current study and refs 26 and 27 use a different algorithm than ref 20, one developed in ref 28. References 26 and 27 studied first-order indices for the most part and in this paper the new algorithm is extended to second-order and higher indices, something already incorporated into the computer program described in ref 20. We use the new algorithm as part of a set of Fortran computer codes that we have developed that are interfaced with the Chemkin package.<sup>41</sup> The global sensitivity analysis presented here is generated either in-line with the calculation of the ignition properties or in postprocessing, with a series of Fortran programs used for the analysis of the ignition data.

The HDMR expansion of the ignition delay time data is written as<sup>16</sup>

$$\tau(\{u_i\}) = \tau_0 + \sum A_i(u_i) + \sum \sum_{i < j} B_{ij}(u_i, u_j) + \dots \quad (3.3)$$

with the  $u_i$  defined in eq 3.1b and  $\tau$  refers to the ignition delay time. The HDMR thus expands the ignition delay time in terms

of a mean ( $\tau_0$ ) and a set of one-reaction terms ( $A_i$ ), two-reaction terms ( $B_{ij}$ ), and higher (not shown). Using the expansion in eq 3.3, it is possible, under well-defined assumptions,<sup>5</sup> to expand the variance of the data as a set of terms involving an increasing number of reactions:

$$V = \sum V_i + \sum \sum_{i < j} V_{ij} + \dots \quad (3.4a)$$

It is common to normalize the partial variances ( $V_i$ ,  $V_{ij}$ , etc) in terms of the overall variance of the data:

$$S_i = \frac{V_i}{V} \quad S_{ij} = \frac{V_{ij}}{V} \quad \text{etc.} \quad (3.4b)$$

The  $S_i$ 's are generally referred to as the first-order sensitivity indices and the  $S_{ij}$ 's as the second-order sensitivity indices. The variance,  $V$ , is defined in the usual manner:

$$V = \frac{1}{n} \sum_{k=1}^n (\tau - \tau_0)^2 = \langle \tau^2 \rangle - \tau_0^2 \quad (3.5)$$

and the overall sample size is  $n$  (here 250 00). Note that the sum of all sensitivity indices is thus

$$\sum S_i + \sum_{i < j} \sum_j S_{ij} + \dots = 1 \quad (3.6)$$

The first-order partial variances ( $V_i$ 's) are calculated from the eighth-order polynomial fits of eq 3.2a drawn as thick lines on Figure 3. The average of  $\tau$  and  $\tau^2$  is calculated by integration over the unit interval defining the uncertainty range, leading to the following estimates:

$$\bar{\tau}_i^f = \sum_{k=0}^8 \frac{a_{ik}}{k+1} \quad (3.7a)$$

$$\langle \tau_i^2 \rangle_f = \sum_{k=0}^8 \sum_{m=0}^8 \frac{a_{ik} a_{im}}{k+m+1} \quad (3.7b)$$

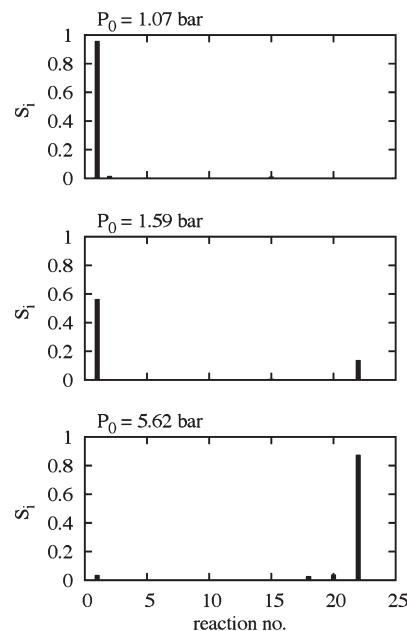
and the first-order sensitivity index is thus estimated as

$$S_i = \frac{V_i}{V} = \frac{\langle \tau_i^2 \rangle_f - (\bar{\tau}_i^f)^2}{V} \quad (3.7c)$$

where the superscript/subscript "f" indicates that the estimate of the ignition delay time is a result of the fit.

The first-order indices in eq 3.7c define a "sensitivity spectrum", which we use here and in earlier papers to identify key reactions whose rate coefficients should be re-evaluated to reduce the uncertainty in the ignition delay time.<sup>26,27</sup> The values of the first-order indices lead to the screening and the theoretical validation of chemical-kinetic reaction mechanisms.<sup>27</sup>

The result of the first-order analysis of the three pressures shown in Figure 3 is presented in Figure 4, where the sensitivity indices of all 25 reactions are shown. Figure 4 demonstrates that the transition evident in Figure 3 can be observed in the first-order sensitivity indices. Figure 4 also demonstrates that the sum of the first-order indices are far from 1.0 for the middle panel, something that might be anticipated from the discussion of the scatter plots of Figure 3 above. The sums of the first-order indices for the three plots in Figure 4 are (from top to bottom) 0.99, 0.71, and 0.98. The rather simple form we have used to present the first-order sensitivities in Figure 4 and our previous work<sup>26,27</sup> is



**Figure 4.** First-order sensitivity indices for the three pressures in Figure 3. These demonstrate that  $\text{H} + \text{O}_2 = \text{OH} + \text{O}$  (reaction 1) has the highest first-order sensitivity at  $P = 1.07$  and 1.59 bar, while reaction 22,  $\text{H} + \text{H}_2\text{O}_2 = \text{H}_2 + \text{HO}_2$ , is highest at  $P = 5.62$ . It is clear that the sum of the first-order indices is much less than 1 in the middle plot at  $P = 1.59$  bar.

meant to convey quickly whether ignition is sensitive to many or a few reactions and to also demonstrate whether the sum of the first-order sensitivities is close to one. There are many more sophisticated and possibly more informative ways to present this material as demonstrated, for example, in ref 33.

We have also investigated the same set of conditions in Figure 4 with local sensitivity analysis and found that the local sensitivity indices do not show the transition. The local indices are similar to the global indices for the top two panels but differ significantly at  $P_0 = 5.62$  bar. The global index for reaction 1 is very small in the bottom panel of Figure 4, but it is the largest local index by far, thus missing the transition.

The second-order indices are calculated starting with an eighth-order expansion for pairs of uncertainties:

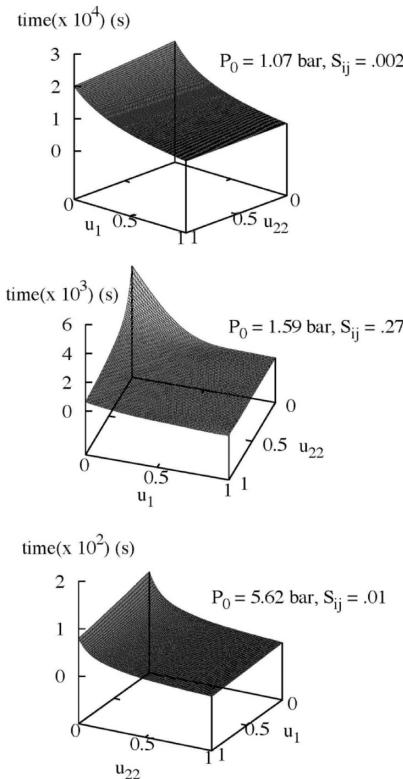
$$\tau(u_i, u_j) = \sum_{n=0}^8 \sum_{k+m=n} a_{ijkm} u_i^k u_j^m \quad (3.8)$$

Figure 5 shows examples of these functions for the three pressures studied here.

The second-order variances are calculated from these expansions in a similar manner as they are for the first-order variances. For example, the average of the fit to the ignition delay time is

$$\bar{\tau}_{ij}^f = \sum_{n=0}^8 \sum_{k+m=n} \frac{a_{ijkm}}{(k+1)(m+1)} \quad (3.9)$$

An expansion for these sums and the average of the squares can be written explicitly as it was for the first-order variances, but it is not presented in this paper. The second-order variance calculated in this manner is labeled  $V_{ij}$  and the corresponding second-order



**Figure 5.** Second-order fits for the reaction pairs 1, 22 (Appendix A) over their uncertainty ranges at three pressures, for stoichiometric mixtures with  $T = 1000$  K. Each panel also includes second-order indices, which is large at  $P = 1.59$  bar and small for the other two.

sensitivity is labeled as

$$s_{ij} = \frac{V_{ij}}{V} \quad (3.10)$$

The second-order sensitivity index (eqs 3.4a and 3.4b) is calculated:

$$S_{ij} = s_{ij} - S_i - S_j \quad (3.11a)$$

The partial variance,  $V_{ij}$ , in eq 3.10 is part of the variance decomposition formula in eq 3.4a and explicitly can be written as<sup>5,42–44</sup>

$$V_{ij} = \text{Var}[E(\tau|u_i=u_i^*, u_j=u_j^*)] - \text{Var}[E(\tau|u_i=u_i^*)] - \text{Var}[E(\tau|u_j=u_j^*)] \quad (3.11b)$$

where  $[E(\tau|u_i=u_i^*)]$  refers to the expectation of the ignition delay time conditioned on a specific value of the uncertainty of reaction  $i$  and the variances are averages over these designated uncertainties. Equation 3.11a is derived from eq 3.11b by dividing all terms by the total variance,  $V$ . Each of the three terms in eq 3.11b corresponds to the terms in eq 3.11a, in the same order. As outlined above, each term in eqs 3.11a and 3.11b is derived from linear regression with eighth-order polynomials.

Each panel of Figure 5 also shows the second-order index for the reaction pair which are the reactions 1 and 22 (Appendix A),  $\text{H} + \text{O}_2 = \text{OH} + \text{O}$  and  $\text{H} + \text{H}_2\text{O}_2 = \text{H}_2 + \text{HO}_2$ , respectively. The second-order sensitivity indices of these pairs are the highest at each pressure. The correlation in the scatter plots in the middle column of Figure 3 is evident in the middle plot of Figure 5,

where there is a distinct rise in the value of the function when the values of  $u_1$  and  $u_{22}$  are both small, but not when either is small with the other being large.

It is also common to study the “component functions” from the HDMR expansion, for example the component  $B_{1,22}$  in eq 3.3 is calculated:

$$B_{1,22}(u_1, u_{22}) = \tau(u_1, u_{22}) - \tau_0 - A_1(u_1) - A_{22}(u_{22}) \quad (3.12)$$

In the rest of the paper we will refer to such functions as component functions and functions such as  $A_1(u_1)$  and  $\tau(u_1, u_{22})$  as first-order and second-order functions, respectively.

Most of the results presented in this paper have been for three initial pressures,  $P_0 = 1.07$ , 1.59, and 5.62 bar for an initial stoichiometric  $\text{H}_2/\text{O}_2$  mixture with initial temperature of 1000 K. These are the pressures that (1) maximize the first-order sensitivity index for  $\text{H} + \text{O}_2 = \text{OH} + \text{O}$  ( $P_0 = 1.07$  bar), (2) maximize the first-order index for  $\text{H} + \text{H}_2\text{O}_2 = \text{H}_2 + \text{HO}_2$  ( $P_0 = 5.62$  bar), and (3) minimize the sum of the first-order terms (eq 3.3) ( $P_0 = 1.59$  bar).

#### IV. CONSTRUCTING THE IGNITION-TIME PDF

In this work the accuracy of the HDMR has been ascertained by confirming that it reproduces the full output variance (i.e., the sensitivity indices sum to close to 1). However, the full uncertainty of the ignition process should be captured by the HDMR through the entire distribution of predictions as expressed by the probability density function (pdf). The ability of the HDMR to capture the shape of the pdf is important if we wish to assess the influence of all significant parameters and their interactions. In our experience, such interactions can often reveal key chemical pathways that drive the predictions of chemical-kinetic models.

In this section we will demonstrate how the HDMR expansion can be used to construct and study the probability distribution of the ignition delay times, referred to in this paper as the “ignition-time pdf”. As noted, this construction serves two purposes: (1) gauge the accuracy of the expansions and (2) gain further insight into the kinetic processes that lead to the shape of the ignition-time pdf. In this section, two ways of constructing the ignition-time pdf with the HDMR are studied.

The uncertainty in the rate coefficients leads to an uncertainty in the ignition delay times. As in previous work,<sup>24,45</sup> the uncertainty is studied in this paper using the HDMR expansion to generate two different functions. The top panel of Figure 6 shows a cumulative distribution function (cdf), and the bottom two panels show an estimate of the ignition delay time probability density function (pdf) generated in two ways from the cumulative distribution function. These plots describe the same set of data as the middle column of Figure 3.

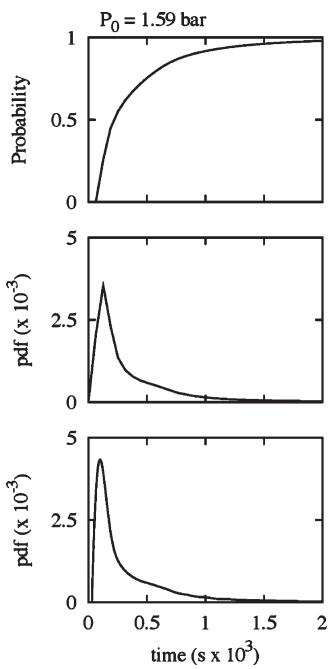
The cumulative distribution function (cdf) is generated in the usual manner:

$$P(\tau) = \frac{N(\tau_{\text{ign}} < \tau)}{N} \quad (4.1)$$

where  $P$  refers to probability and  $\tau$  refers to the ignition delay time. A direct count of the 250 000 realizations of the ignition process are used to calculate  $N(\tau_{\text{ign}} < \tau)$  in eq 4.1.

The probability density function is calculated from the cumulative distribution function:

$$p(\tau) = \frac{dP(\tau)}{d\tau} \quad (4.2)$$



**Figure 6.** Cumulative distribution function (top plot) and pdfs generated from it by differentiation (bottom two plots). Derivatives are generated in the middle plot by finite-difference and the bottom plot from analytic differentiation of a Gaussian process interpolation of the cdf in the top panel. The initial pressure and temperature are the same as they were for the middle panel of Figure 4.

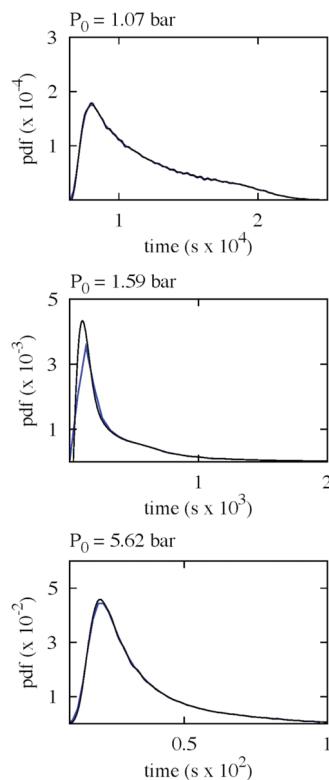
The derivative is calculated in two ways in this paper. The first is with a finite difference approximation to the derivative. At the  $k$ th point along the cumulative distribution the finite difference is calculated:

$$\left. \frac{dP(\tau)}{d\tau} \right|_{\tau=\tau_k} = \frac{P(\tau_{k+1}) - P(\tau_{k-1})}{2\Delta\tau} \quad (4.3)$$

where it is assumed that the  $\tau$ 's are evenly spaced. The derivatives in eq 4.3 are accurate to second order. For the end points, a different formula is necessary but because the cdf is flat at its beginning and end, the end points are left out here.

The bottom panel of Figure 6 shows a smoother version of the pdf. It is calculated using the derivatives of an interpolated version of the cdf in the top panel. The interpolation is done via the Gaussian process (GP)<sup>46,47</sup> model and is described in Appendix B. There are no parameters necessary to accomplish the interpolation, but there are “hyperparameters” that are adjusted via maximum likelihood to avoid overly complex models that lead to overfitting.<sup>46</sup> For the Gaussian process model, the derivatives can be evaluated analytically as described in Appendix C. Because the hyperparameters of the model are adjusted to avoid an overly complex model (see chapter 5 in ref 46), we view the generation of the ignition-time pdf using the GP model as an “optimal bandwidth” fit.<sup>48</sup> Many of the pdf estimates in this paper use the GP model.

Figure 6 shows an example where the Gaussian process model gives a more consistent representation of the density than an evaluation based on a finite-difference form of the derivative. We found that the case in Figure 6 was very sensitive to the spacing and positions of the points near the peak when a finite-difference



**Figure 7.** Comparison between the finite-difference form of the pdf (blue curves) and the form calculated from the Gaussian process model (black curves).

form of the derivative was used. With the GP model the pdf is much less sensitive to these features. As demonstrated in Appendix C (see ref 47, for example), the GP model forms an interpolated function as a linear combination of the points put into the model. The correlation of the points is embodied in the hyperparameter,  $\theta$ , which describes a “length scale” (eqs C.2a and C.2b). The values of  $\theta$ , which are calculated from maximum likelihood, are high for all cases studied in this paper. For the case in Figure 6 it has a value of  $3 \times 10^7 \text{ s}^{-2}$ , which means that, effectively, the interpolation uses only a few nearby points.

A more direct comparison of the finite-difference form of the pdf and the GP version is shown in the middle panel of Figure 7, where the finite-difference form is plotted as a blue line and the GP form as a black line. In this case there is a significant difference between the finite-difference and the GP versions of the ignition-time pdf, among the largest we have observed. The top and bottom panels for  $P = 1.07 \text{ bar}$  and  $P = 5.62 \text{ bar}$  show much closer agreement between the finite-difference and GP forms of the ignition-time pdf. Our experience so far is that the GP form performs better near the peak of the distribution than a finite-difference version of the pdf, but there is little difference between the two away from the peak. So many of the results presented in section VI use the finite-difference form, because fine details of the pdf are investigated there away from the peak.

## V. DECONSTRUCTING THE IGNITION-TIME PDF

**A. Motivation.** The HDMR expansion (eq 3.3) can be used to reproduce the ignition-time pdfs described in section IV.<sup>24,45</sup> We have found that comparisons of the true ignition-time pdf

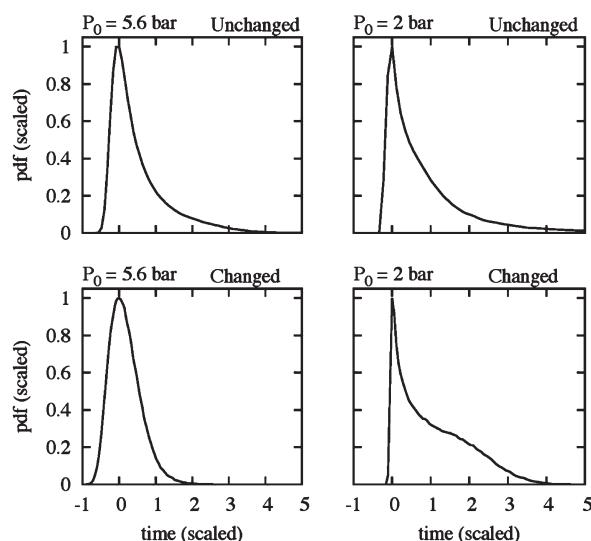
with reduced pdfs generated from various portions of the HDMR expansions provide a detailed picture of the processes that lead to the shape of the ignition-time pdf. These sorts of comparisons can provide insight into which reactions contribute to the shape of the ignition-time pdf and the time ranges that they contribute. We refer to the select investigation of portions of the HDMR expansion as “deconstruction”, because it provides a deeper understanding of a pdf.

The reproduction of the ignition-time pdf is also a better gauge of the convergence of the HDMR expansion<sup>24</sup> than the sum of the partial variances in the variance decomposition (eq 3.6). Because the terms in eq 3.6 must sum to 1, the variance decomposition is often used as a way to gauge convergence. Sums close to 1 are generally expected to result in a very accurate assessment of the degree of information contained in the expansion. However, our experience, as well as the calculations in ref 19, suggest that a better assessment of the degree of convergence is provided by the comparisons of the true pdf with the deconstructions. In addition, a deconstruction can indicate which time ranges of the pdf are converged.

One of the difficulties encountered in the convergence process is that the number of terms in the HDMR expansion gets very large at higher orders. For the H<sub>2</sub>/O<sub>2</sub> mechanism studied here, which is small, there are 25 first-order terms, 300 second-order terms, and 2300 third-order terms. The GUI\_HDMR software<sup>20</sup> provides a threshold method to exclude terms from the expansion that do not improve the overall fit. However, each term has to be tested using this approach. The deconstruction in section VD suggests an alternative efficient method to accurately implement the HDMR expansion with a limited number of terms.

Another reason for the deconstruction is the use of uncertainty in the validation of chemical-kinetic mechanisms. The theoretical validation procedure developed in refs 26 and 27 identifies reactions on the basis of their first-order indices described in section IIIA. Rate coefficients for the reactions with the highest first-order indices are then calculated to improve their values and to reduce the uncertainty in the observables.<sup>26,27</sup> It is assumed that a reduction in uncertainty of the reaction with the largest fractional variance (i.e., first-order sensitivity index) will lead to a reduction in the uncertainty of the observable that is targeted by the global sensitivity analysis if the rate coefficient for that reaction does not change significantly. Our experience so far suggests that this is generally true, as long as the largest fractional variance is a significant fraction of 1. For cases where the rate coefficient changes significantly along with the uncertainty, scaling of the observable may need to be introduced to observe the reduction of the overall uncertainty.<sup>27</sup> In addition, reduction of the uncertainty of a specific reaction targeted at a specific set of conditions may not reduce the uncertainty in all portions of the space of pressures, temperatures, and stoichiometries. The ignition-time pdf provides evidence as to why this occurs.

Most of the points in the previous three paragraphs are emphasized in Figure 8. The top row shows ignition-time pdfs for two pressures,  $P_0 = 5.6$  bar and  $P_0 = 2$  bar. As with all other results in this paper, these were generated for initially stoichiometric H<sub>2</sub>/O<sub>2</sub> mixtures with  $T_0 = 1000$  K. The system studied in the top row uses the same set of uncertainties as all other calculations in this paper (Appendix A). The ignition-time pdfs in the bottom row of Figure 8 were generated under all the same conditions as those in the top row, but the uncertainty of reaction 22,  $H + H_2O_2 \rightarrow H_2 + HO_2$ , has been reduced from 5 to 1.58, the same value as reaction 2, O + H<sub>2</sub> → OH + H.



**Figure 8.** Comparison of two sets of ignition-time pdfs for  $P_0 = 5.6$  bar and  $P_0 = 2$  bar. The top plots were generated with the system studied in this paper and the bottom is for a system with the uncertainty of reaction 22 reduced to 1.58.

To make a proper comparison of the ignition-time pdfs, we scale the coordinates in the following manner. The time is shifted so that the maximum value of the pdf is at 0.0. The value of the pdf is then scaled so that its maximum is 1.0. The times are further scaled to ensure that the integral of the pdf over the full time range is unity. The resulting functions are presented in Figure 8. Both of the pdfs in the top row can be compared to their unscaled counterparts. The one at  $P_0 = 5.6$  bar is presented in Figure 14 and the one at  $P_0 = 2$  bar in Figure 20.

In the top left plot of Figure 8, the first-order sensitivity index of reaction 22 is 0.87. The total variance of the ignition-time pdf is  $3.2 \times 10^{-6} s^2$  in the original time units. The theoretical validation procedure developed in refs 26 and 27 would identify this as the reaction that should be studied and it is assumed that it has been updated in the bottom panel of the first column of Figure 8. For convenience it is assumed that the updated rate coefficient is the same, but due to the accuracy of the theoretical calculation, the uncertainty is reduced to 1.58. A calculation of the pdf of the ignition delay times for this system is shown on the bottom left plot. Visual comparison of the two pdfs makes it clear that the reduction of the uncertainty of reaction 22 has led to a reduction in uncertainty in the ignition-time pdf. The value of the total variance also indicates that this is true. It is now  $1.1 \times 10^{-6} s^2$ , a reduction of nearly a factor of 3. This variance is calculated not merely from a subset of the original system, but as with the previous set of uncertainties (eq 3.1a), the uncertainty range is always scaled to be between 0.0 and 1.0. In addition to the reduction of total variance, the first-order sensitivity index of reaction 22 has been reduced to 0.40. This is still the highest partial variance, but other reactions have significant indices. For example, the following reactions have significant first-order sensitivity indices:  $S_1 = 0.19$ ,  $S_{18} = 0.13$ , and  $S_{20} = 0.23$ .

The reduction of the uncertainty evident in the left column is not evident in the right column. At the lower pressure (2 bar) reaction 22 does not have the highest sensitivity index in the original system (upper right pdf). Its value is 0.36, which is second to reaction 1, which has a value of 0.43. In addition, there

is a significant second-order component, not discussed here. The total variance of the ignition-time pdf in the original time units is  $1.2 \times 10^{-6} \text{ s}^2$ . The reduction in the uncertainty of reaction 22 does not appear to reduce the overall uncertainty, as a visual comparison of the top right and bottom right pdfs suggest. The overall variance actually goes up approximately 10%, to  $1.3 \times 10^{-6} \text{ s}^2$ . The first-order sensitivity index of reaction 22 drops from 0.36 to 0.052, with reaction 1 increasing from 0.43 to 0.87. As noted in the previous paragraph, the uncertainty range of reaction 22 is rescaled to be between 0.0 and 1.0. This rescaling does not contribute to the overall shape of the ignition-time pdf, but we have found that specific types of nonlinearity can lead to an increase of the variance in the rescaled units, as it does in this case. This same nonlinearity also leads to a significant change in the shape of the pdf, which is what the second column shows. Section VIC revisits Figure 8, where the rescaling will be discussed in more detail.

Visual inspection of the ignition-time pdfs in Figure 8 suggests that they encode additional information that is not evident in the variance, which, after all, is generated from averaging. As the top row indicates, there is a distinct difference in the shapes of the pdfs as pressure increases from 2 to 5.6 bar. Comparison of the two pdfs in the right column of Figure 8 indicates that the lack of reduction in variance is due to the development of a shoulder at the scaled time of about 1 (0.0018 s in unscaled time). The deconstructions discussed in this section, particularly the one in section VD, are useful for explaining the features of pdfs. The geometrical analysis in section VI is based on the deconstruction of section VD and it contributes additional information concerning the shape of the pdf. The combination of these two analyses often makes it possible to assign the features of a pdf to particular reactions or features of the HDMR expansions.

**B. Deconstruction 1: Order by Order.** All the deconstruction procedures used in this section start by generating approximations to the true ignition-time pdf from the HDMR expansion. The order-by-order deconstruction described here has been investigated in previous work,<sup>20,24,45</sup> where it was used to estimate the accuracy of HDMR expansions.

To generate an approximation of the ignition-time pdf from the HDMR expansion, eq 3.3 is rewritten for convenience:

$$\tau(\{u_i\}) = \tau_0 + \sum \tau_i^{(1)}(u_i) + \sum_{i < j} \sum_j \tau_{ij}^{(2)}(u_i, u_j) + \dots \quad (4.4)$$

This approximation serves as a surrogate for the full calculation described in section II. A set of ignition delay times is generated from eq 4.4 at the same 250 000 points as the full calculations. The calculation then follows the procedure described in section IV. For example, the cumulative distribution function is calculated in the usual manner:

$$P^a(\tau) = \frac{N(\tau_{\text{ign}}^a < \tau)}{N} \quad (4.5a)$$

but the designation " $\tau_{\text{ign}}^a$ " refers to an approximation to the true ignition delay time based on the expansion in eq 4.4. The approximation in eq 4.5a might be first order

$$\tau_{\text{ign}}^a(\{u_i\}) = \tau_0 + \sum \tau_i^{(1)}(u_i) \quad (4.5b)$$

or second-order

$$\tau_{\text{ign}}^a(\{u_i\}) = \tau_0 + \sum \tau_i^{(1)}(u_i) + \sum_{i < j} \sum_j \tau_{ij}^{(2)}(u_i, u_j) \quad (4.5c)$$

or higher. Generally, the expansion is taken to second order, because the number of terms in third order and higher is very large. For the hydrogen system there are 300 second-order terms and 2300 third-order terms.

The ignition-time pdf is calculated from the approximate cdf in the usual manner:

$$p^a(\tau) = \frac{dP^a(\tau)}{d\tau} \quad (4.6)$$

It is the process of investigating the ignition-time pdf as the order of the HDMR is increased that we refer to here as "deconstruction".

Comparisons are made for the cumulative distribution function and the pdf at first-order in Figure 9 (eq 4.5b) and second-order in Figure 10 (eq 4.5c). Because of the convergence of the Monte Carlo procedure with sample size (Appendix B), the only second-order terms used are ones where both of the reaction pairs have first-order sensitivity coefficients that are 0.001 or higher. With this restriction the following set of reactions are used for the second-order fits (the numbers refer to the reaction indices in Appendix A): 1, 2, 3, 14, 15, and 22 at  $P_0 = 1.07 \text{ bar}$ ; 1, 15, 22 at  $P_0 = 1.59 \text{ bar}$ ; and 1, 14, 15, 18, 19, 20, 22 at  $P_0 = 5.62 \text{ bar}$ . In these plots the top row shows the comparisons of the true cumulative distribution in black and the expansions in blue. The same set of pressures is used as in most of the previous figures. The ignition-time pdf comparisons are shown in the bottom row of Figures 9 and 10.

Comparisons of the HDMR fits with the true distribution and density functions demonstrate that the first-order fits are only reasonable at  $P = 1.07 \text{ bar}$  and that the second-order fits are very good at  $P = 1.07 \text{ bar}$  and  $P = 5.62 \text{ bar}$ . The second-order fit is very good at  $P = 1.59 \text{ bar}$  except at the peak, which, as noted above, is difficult to accurately model.

These results demonstrate that comparisons of the pdfs generated from an HDMR expansion and the actual pdfs are more sensitive to the accuracy of the HDMR expansions than the sensitivity indices. For example, the sum of the first-order indices is 0.98 for  $P = 5.62 \text{ bar}$  in Figure 9, but the fit of the ignition-time pdf is significantly different from the true pdf. Although the visual inspection of the pdfs is useful, numerical calculations of fitting errors give even more information and some of these will be shown in the final deconstruction in section IV.

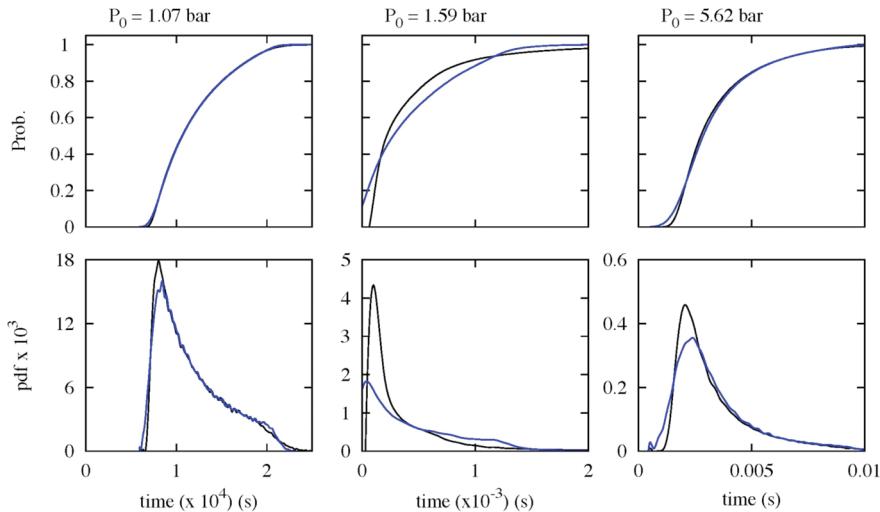
**C. Deconstruction 2: Term Ordering.** A deconstruction of the ignition-time pdf is provided by a sum of terms based on their sensitivity indices, whether the term is first-order, second-order, or higher. This is investigated for the case in the middle column of Figure 3. At this pressure (1.59 bar) the order of the sensitivity indices is

$$S_1 > S_{1,22} > S_{22}$$

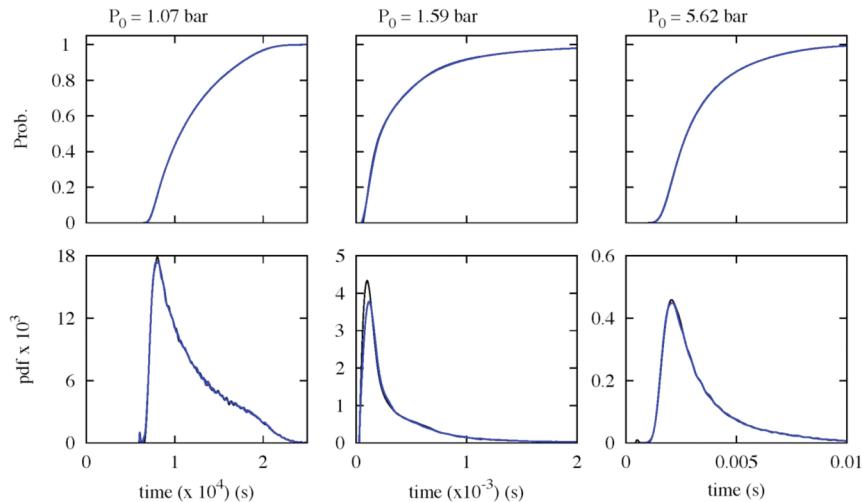
with the reaction numbers defined in Appendix A. The series of fits from these three sums (from top to bottom in Figure 11) are

$$\tau_{\text{ign}}^a(u_1) = \tau_0 + \tau_1^{(1)}(u_1) \quad (4.7a)$$

$$\tau_{\text{ign}}^a(u_1, u_{22}) = \tau_0 + \tau_1^{(1)}(u_1) + \tau_{1,22}^{(2)}(u_1, u_{22}) \quad (4.7b)$$



**Figure 9.** Comparisons of the full cumulative densities (top row) and ignition-time pdfs for the same cases as Figure 7. The black curves show the full pdf and the blue curves the first-order expansions described in the text. These plots are generated with all the first-order terms.



**Figure 10.** Comparisons of the full cumulative densities (top row) and ignition-time pdfs for the same cases as Figure 8. The black curves show the full pdf and the blue curves the second-order expansions described in the text (eq 4.5c).

$$\tau_{\text{ign}}^{\text{a}}(u_1, u_2) = \tau_0 + \tau_1^{(1)}(u_1) + \tau_{1,22}^{(2)}(u_1, u_{22}) + \tau_{22}^{(1)}(u_{22}) \quad (4.7\text{c})$$

Figure 11 demonstrates that the convergence of this deconstruction is uneven, but that the final fit at the bottom is better than the fit in the middle, bottom plot of Figure 10, which includes many extra terms. Because of the restrictions on the number of second-order terms, the fit in Figure 10 includes all first-order terms and the second-order terms associated with the following second-order indices: S<sub>1,15</sub>, S<sub>1,22</sub>, and S<sub>15,22</sub>.

The deconstructions studied here can also be used to reproduce plots like those in Figure 3. As in Figure 3, a subset of the points is plotted: 2500 out of 250 000. The three columns in Figure 12 are generated from the three expansions in eqs 4.7a–4.7c and are used to study the differences between the fits and the true data. The expansion in eq 4.7a is one-dimensional in reaction 1, and thus the top plot in the first column is a one-dimensional curve. The

expansion does not depend at all on the reaction H + H<sub>2</sub>O<sub>2</sub> (22), and thus the projection of the 25-dimensional space of the reactions onto this single reaction shows a scatter of points, with no correlation evident.

In column 2 of Figure 12, the expansion of eq 4.7b is used. Here there is now a correlation for both reactions and it is observed that the spread of points takes on different widths as the uncertainty parameters are varied, as was noted in section IIIA in the discussion of Figure 3. Because the expansion is two-dimensional and has coupling between the two reactions the scatter of points in the plot results simply from projecting a two-dimensional surface onto a plane. Finally, in column 3 of Figure 12, where an accurate fit has been obtained, the scatter of points indicates that the three-term expansion in eq 4.7c faithfully reproduces the two plots in the middle column of Figure 3.

**D. Deconstruction 3: Reaction Ordering.** The bottom plot of Figure 11 and the right-hand column of Figure 12 indicate that the ignition delay time distribution is two-dimensional in the space of the reactions at P<sub>0</sub> = 1.59 bar. These results suggest that

an expansion based on the reactions could generate an efficient representation of the ignition-time pdf. In this subsection we demonstrate that an expansion based on the reactions is most likely the best deconstruction, combining accuracy with increased understanding. Because the expansion is done on a reaction-by-reaction basis, the increase in the number of terms in the expansion as the order is increased is not a big

problem, because higher order terms are generated only as needed.

The deconstruction of the HDMR expansion as reactions are added is written as

$$\tau_{\text{ign}}^m(\{u_i\}_m) = \tau_0 + \sum_{j=1}^m \tau_{\{u_j\}}^{(j)}(\{u_j\}) \quad (4.8a)$$

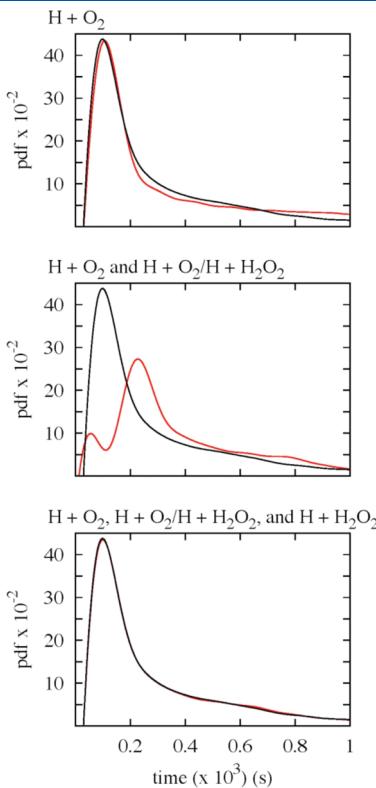
where the index “*m*” refers to an approximation based on using the first *m* reactions. The set  $\{u_i\}_m$  refers to the set of *m* uncertainty progress variables (eq 3.1b) at the *m*th level. For a three-reaction fit the expansion has the following form:

$$\begin{aligned} \tau_{\text{ign}}^3(\{u_{\text{im}}\}) = & \tau_0 + \tau_1^{(1)}(u_1) + [\tau_2^{(1)}(u_2) + \tau_{1,2}^{(2)}(u_1, u_2)] \\ & + [\tau_3^{(1)}(u_3) + \tau_{1,3}^{(2)}(u_1, u_3) + \tau_{2,3}^{(2)}(u_2, u_3)] \\ & + \tau_{1,3}^{(2)}(u_1, u_3) + \tau_{1,2,3}^{(3)}(u_1, u_2, u_3) \end{aligned} \quad (4.8b)$$

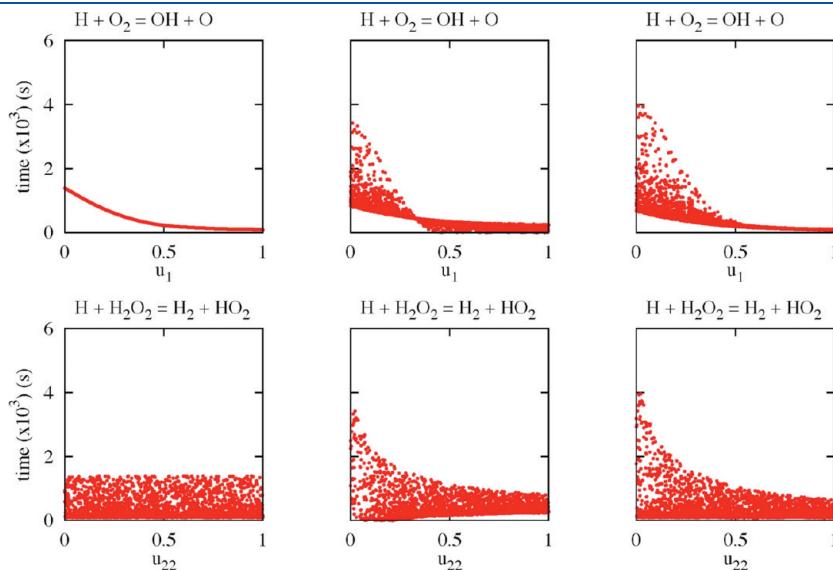
where the numbers “1”, “2”, and “3” refer to the ordering of the reactions and not the numbering in Appendix A. The reaction ordering is based on the first-order sensitivities of each reaction.

Figure 13 shows a set of pdfs generated from a set of reaction-ordered expansions for the case studied in Figures 11 and 12 and a number of earlier figures. The top panel shows an expansion based on a single reaction, the second panel one that sums terms for two reactions and the bottom two panels show two three-reaction expansions. For most cases we have studied it has only been necessary to expand to second order, so that the last term in eq 4.8b is left out, and this is what is presented in the third panel of Figure 13. However, in this particular case it is necessary to use the third-order term in the expansion to accurately fit the ignition-time pdf and that is what is done in the bottom panel of Figure 13.

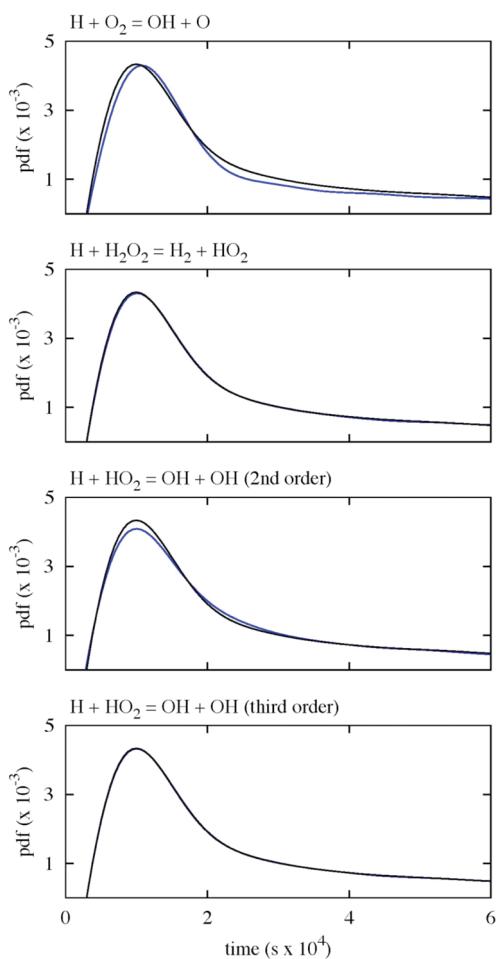
Some caution should be exercised when the procedure presented in this subsection is used. There could be cases where two reactions both have small first-order terms, but a large second-order term, and cause the reaction ordering to break down. While it is clear that one can devise models that behave in this manner,



**Figure 11.** Term-by term expansions at  $P_0 = 1.59$  bar are shown for the first term (top) (eq 4.7a), the sum of the first two (middle) (eq 4.7b), and the sum of the first three terms (bottom) (eq 4.7c).



**Figure 12.** Two projections of each of the expansions in Figure 11 (the columns). These plots can be compared to the middle column of Figure 4.



**Figure 13.** Set of pdfs generated from a reaction-ordered expansion, as described in this subsection. The black curves show the full pdf and the blue the expansions.  $P_0 = 1.59$  bar.

we have not observed such behavior in our studies. An effect like this would lead to a poor fit to the ignition-time pdf and this should be evident with a close examination of such a fit.

The accuracy demonstrated in Figure 13 is also evident at two other pressures studied throughout this paper. The top row of Figure 14 shows results for  $P_0 = 1.07$  bar and the bottom row for  $P_0 = 5.62$  bar. For completeness, the middle row repeats the results of Figure 13 at  $P_0 = 1.59$  bar.

All the results presented in Figure 14 are generated from the GP model described in section IV and Appendix C. The black curves show the full ignition-time pdf and the blue curves show the deconstructions as outlined in this subsection. No terms higher than second-order are used in Figure 14, except for the third-order term used at  $P_0 = 1.59$  bar, as it was in the bottom panel of Figure 13. One can observe that the first two blue curves in the top row have some additional structure at the lowest values of time, a result of the GP interpolation that is not present in finite-difference versions of the fit. This extra structure disappears for the last two fits in the top row, where convergence is much better. Despite this structure, we view the GP model as a better way to represent the density. The extra structure that appears for the expansions in the top row never appears for the fit of the full pdf and merely reflects a lack of convergence. It is possible to smooth this structure by using a "noise" term in the GP model

(see, for example, chapter 6 in ref 47), but we have not pursued that in this paper.

The reaction ordering advocated in this subsection is useful because it deconstructs the ignition-time pdf on a reaction-by-reaction basis. Our experience also indicates that the fitting error decreases in a monotonic manner with the addition of reactions based on their first-order sensitivity indices, something that does not always occur with the term ordering deconstruction in section VC. We define the average root-mean-square error as

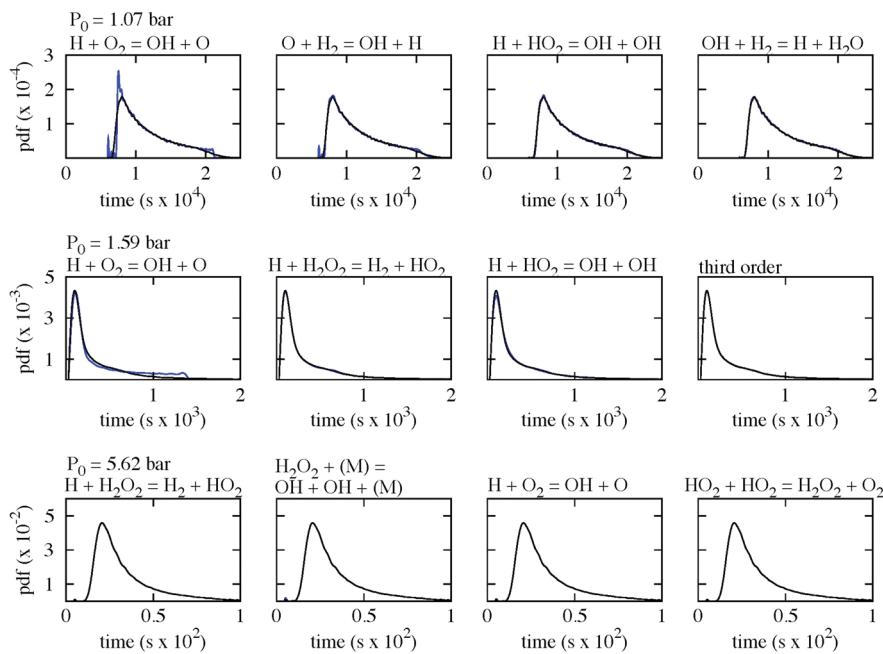
$$\bar{E} = \sqrt{\frac{\sum (\tau - \tau_f)^2}{n}} \quad (4.9)$$

where  $n$  is the sample size with  $\tau$  and  $\tau_f$  the ignition delay time from a simulation and one generated by the fit to it (eq 4.8a), respectively. Figure 15 shows the errors for the cases shown in Figure 14, with the results at  $P_0 = 1.07$  bar and  $P_0 = 5.62$  extended to a larger number of reactions than in Figure 14. The results labeled as level "0" are when only the average value of  $\tau$ ,  $\tau_0$  is used in eq 4.9. An interesting aspect of the plots in Figure 15 is the relative decline in the error when a specific reaction is added. For example, there is a large decline in both the top and bottom plots when just the first reaction is added, and a more gradual decline afterward. Such a decline is obviously connected to the significant fraction of variance described by that reaction (see Figure 14 for the specific reaction).

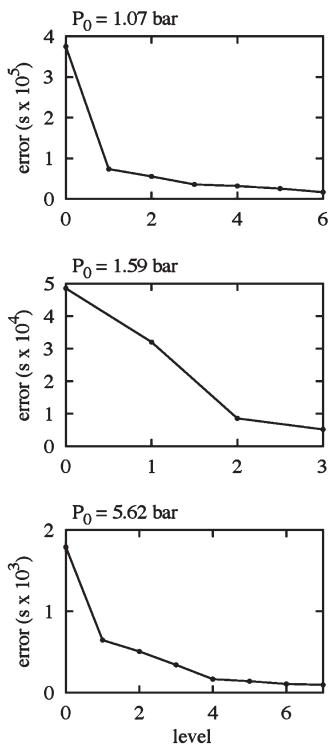
The results presented so far in this subsection do not give a complete picture of the utility of the present deconstruction and Figure 16 shows an interesting case that is similar to others we have encountered for other initial conditions. The plots in Figure 16 show results for an ignition simulation that has the same initial conditions as the others in this subsection, except  $P_0 = 0.5$  bar. We show pdfs generated from a one-reaction fit in the top panel and a two-reaction fit in the bottom panel. The pdf approximations are drawn in blue and the true pdf is drawn as a black line. The top panel demonstrates that the reaction  $H + O_2 = OH + O$  contributes to the ignition-time pdf over a well-defined time range and makes no contribution at long and short times. When a second reaction is added in the bottom panel, the full extent of the pdf is reproduced, including the distinct drop in the pdf after 0.0175 ms, suggesting that the second reaction contributes significantly to the longest time behavior when added to the sum in eq 4.8a.

## VI. GEOMETRICAL INTERPRETATION OF THE REACTION-ORDER DECONSTRUCTION

**A. Second-Order Functions.** A geometrical interpretation of the ignition-time pdf using the deconstruction of section VD is useful for understanding the shape of the ignition-time pdf in particular the sharp changes that occur as a function of  $\tau$ . We demonstrate this for the ignition-time pdf in Figure 16. For the conditions of this case,  $P_0 = 0.5$  bar and  $T_0 = 1000$  K, the reactions that have the highest first-order sensitivity indices are  $H + O_2 = OH + O$  (0.92) and  $O + H_2 = OH + H$  (0.057). There is a small second-order term for these two reactions,  $S_{12} = 0.000\ 13$ . Single reaction cumulative distribution functions (eq 4.7a) for these two reactions are shown in Figure 17. The ignition-time pdf of the top reaction is drawn on the top panel of Figure 16 and the combination of the two (eq 4.8a) with the second-order term gives the blue curve in the bottom panel of Figure 16. The dots on



**Figure 14.** Series of reaction-by-reaction deconstructions at  $P_0 = 1.07$  bar (top row),  $P_0 = 1.59$  bar (middle), and  $P_0 = 5.62$  (bottom). The full pdf is shown as a black curve and the deconstructions are shown as blue curves. No plot includes terms higher than second order, except the fourth plot in the second row.



**Figure 15.** Average error for the three pressures in Figure 14 as defined in eq 4.9.

each curve in Figure 17 are near the maximum of the pdf at  $\tau = 0.000\ 12$  s and the sharp change near  $\tau = 0.000\ 18$  s.

The blue curves in Figure 16 were generated using eq 4.8a. As described in the previous section, these pdfs were generated from the expansions. The expansions were evaluated at the Monte

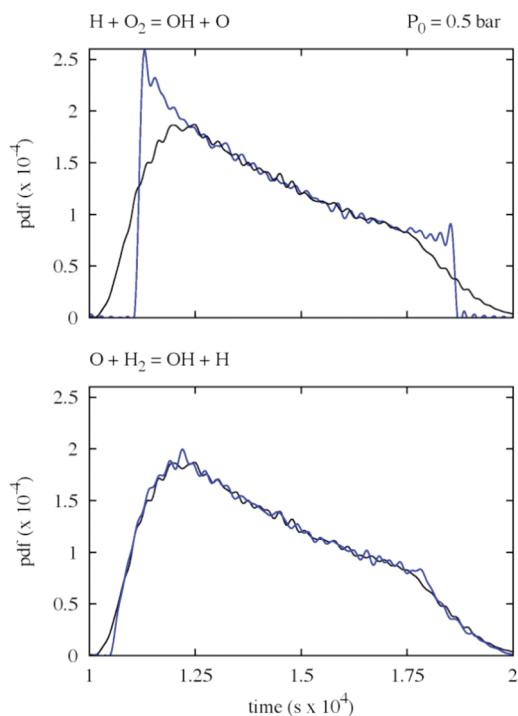
Carlo points that were used to generate the full ignition-time pdfs and many other results in this paper. A more accurate method to generate the first-order and second-order expansions is through the geometry of the surfaces generated from the expansions. Because the second-order expansions describe most of the behavior of the ignition-time pdfs for the  $H_2/O_2$  system this can be very useful.

Figure 17 indicates that the reaction in the top panel does not provide the correct density at low and high values of  $\tau$ , something previously observed in the ignition-time pdf generated from this cdf, shown in the top panel of Figure 16. On the other hand, the single reaction cdf of the second most important reaction does have nonzero values at the shortest and longest times, as indicated in the bottom panel, and it is not surprising that the combination of the two gives a very reasonable representation of the ignition-time pdf, as indicated on the bottom panel of Figure 16. However, neither of these cdfs taken separately have sufficient information to describe the changes that occur near  $\tau = 0.000\ 12$  s and  $\tau = 0.000\ 18$  s.

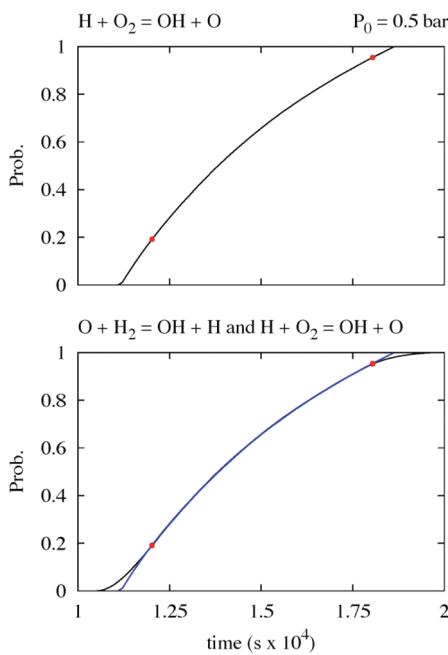
Although the single-reaction cdfs do not adequately describe the full cdfs, it is useful to introduce the geometrical picture with one of these cases. The single-reaction first-order function for reaction 1 is shown in the top panel of Figure 18. Once again, as with all first-order functions in this paper it is approximated by an eighth-order polynomial. To generate a cdf from this function at a specific value of  $\tau_\beta$ , it is necessary to find the total length along the  $u_1$ -axis for all  $u_1$ 's whose  $\tau$ -values lie below  $\tau_\beta$ . For a monotonically decreasing function, this involves finding the intersection of the first-order function and the line  $\tau = \tau_\beta$ . If we designate the intersection point as  $u_{1\beta}$ , then the probability is

$$P(\tau_\beta) = 1 - u_{1\beta} \quad (6.1)$$

The two horizontal lines drawn on the top panel of Figure 18 are  $\tau_\beta = 0.000\ 12$  s and  $\tau_\beta = 0.000\ 18$  s. The cumulative



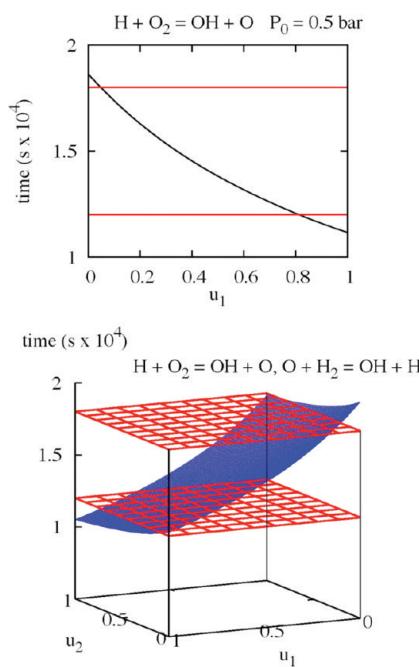
**Figure 16.** Comparison of one reaction and two reaction fits (blue curves) to the full ignition-time pdf (black) for the case described in the text.



**Figure 17.** One-reaction cumulative density functions as described in the text. The top shows the CDF for the  $H + O_2$  reaction and the bottom repeats this reaction in blue, along with the CDF for the  $O + H_2$  reaction. The locations of the dots are described in the text.

distribution function values can be estimated from inspection of the plot as  $P(0.000\ 12) = 0.2$  ( $1 - 0.8$ ) and  $P(0.000\ 18) = 0.95$ .

The geometric picture is extended to a second-order function in the bottom panel of Figure 18. As with all second-order functions, this is an expansion to eighth order in the two relevant



**Figure 18.** Cdfs from one-reaction functions (top) and two-reaction functions (bottom) are the fraction of the unit interval and unit square, respectively. The lines and planes are  $\tau_\beta = 0.000\ 12$  s and  $\tau_\beta = 0.000\ 18$  s. See text for further details.

reactions. The cumulative distribution function is calculated as the fraction of the unit square whose  $\tau$ -values along the second-order surface lie below the plane  $\tau = \tau_\beta$ . The two planes drawn on the bottom panel of Figure 18 are once again  $\tau_\beta = 0.000\ 12$  s and  $\tau_\beta = 0.000\ 18$  s. The area defined in this manner changes with  $\tau_\beta$  and the ignition-time pdf reflects this change. The cdf and pdf are written as

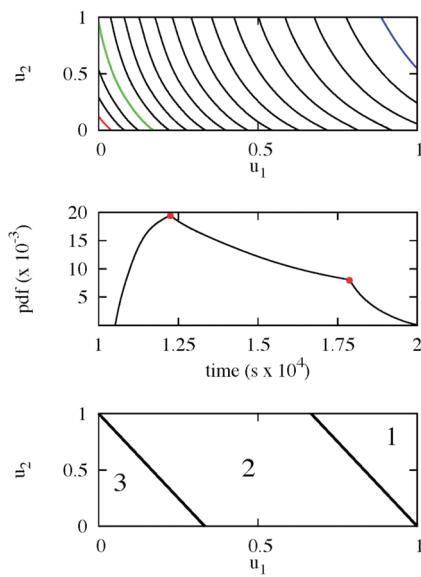
$$P(\tau) = A(\tau) \quad (6.2a)$$

$$p(\tau) = \frac{dA(\tau)}{d\tau} \quad (6.2b)$$

Extra complications arise in two dimensions compared to one dimension. The domain is a unit square and significant changes occur in the ignition-time pdf when the second-order surface touches a corner as it does for the two planes illustrated in the bottom panel of Figure 18. Under such a circumstance, there is a change in the area function in eq 6.2a. There are more complications for second-order surfaces that are not monotonic. Very few of the surfaces we have investigated in the  $H_2/O_2$  system have such behavior to any noticeable extent, and we do not discuss any cases like this in the present paper.

To demonstrate how these changes occur, it is useful to investigate contour plots of the second-order functions. The second-order function in the bottom panel of Figure 18 is drawn as a set of contours in the top panel of Figure 19. There are 17 contours plotted in the panel at intervals of  $5 \times 10^{-6}$  s, with the contour heights ranging from 0.000 11 s in the upper right (blue curve) to 0.000 195 s in the lower left (red curve). The contour at  $\tau = 0.000\ 18$  s is drawn as a green curve. The rest of the contours are drawn as black lines.

The areas of the regions enclosed by the contours and the appropriate portions of the boundaries define the cumulative



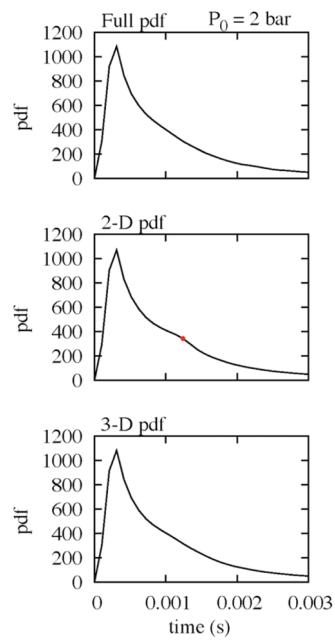
**Figure 19.** Top: contour plot of the two-reaction function of Figure 18. Middle: pdf derived from the areas enclosed by the contours. Features of the pdf are delineated by the two dots drawn on it and correspond to three regions of the unit square (bottom plot).

distribution function (eq 6.2a). Several hundreds of these areas are used to estimate the ignition-time pdf (eq 6.2b), and it is plotted in the second panel of Figure 19. The sharp changes in the ignition-time pdf are correlated with three regions. Region 1 in the bottom panel of Figure 19 describes the initial rise of the ignition-time pdf. At the boundary between region 1 and 2 there is a change in the form of the pdf as indicated by the dot on the ignition-time pdf in the middle panel. In region 2 the behavior is different than in region 1 and the change from region 2 to 3 leads to another change in the behavior of the ignition-time pdf (the second dot on the curve in the middle panel).

**B. Three-Dimensional Geometrical Analysis.** The 2-D (two reaction) ignition-time pdfs often reproduce the distinct features of the full ignition-time pdf for the  $H_2/O_2$  system, but some of the fine features of the 2-D pdfs may not be present in the full pdf. Although the sharp changes that occur in the 2-D pdf in the second panel of Figure 19 are present in the full pdf of Figure 16, this is not always true and we show an example here.

Figure 20 shows the full pdf for the same initial temperature (1000 K) and stoichiometry used throughout the paper. The initial pressure is now 2 bar. All the pdfs shown in Figure 20 were generated with the finite difference version of the derivative used to convert the cdfs to pdfs (eq 4.3). The 2-D pdf in the middle panel of Figure 20 has a distinct change near  $\tau = 0.00125$  s as indicated by the dot. This feature is not present in the full pdf (top panel). The two reactions in the 2-D pdf are  $H + O_2 = OH + O$  (reaction 1) and  $H + H_2O_2 = H_2 + HO_2$  (reaction 22). To understand the change between the 2-D and the full pdf, we first investigate the three-reaction pdf, which is presented in the bottom panel in Figure 20. The 3-D expansion has the additional reaction,  $HO_2 + H = OH + OH$ , the 15th in Appendix A. The bottom panel of Figure 20 demonstrates that the three-reaction pdf, like the full pdf, does not have the distinct change that is present in the 2-D pdf.

We now demonstrate geometrically how the sharp feature at  $\tau = 0.00125$  s disappears. The top left panel of Figure 21, like the top panel of Figure 19, shows a series of contours for the two-



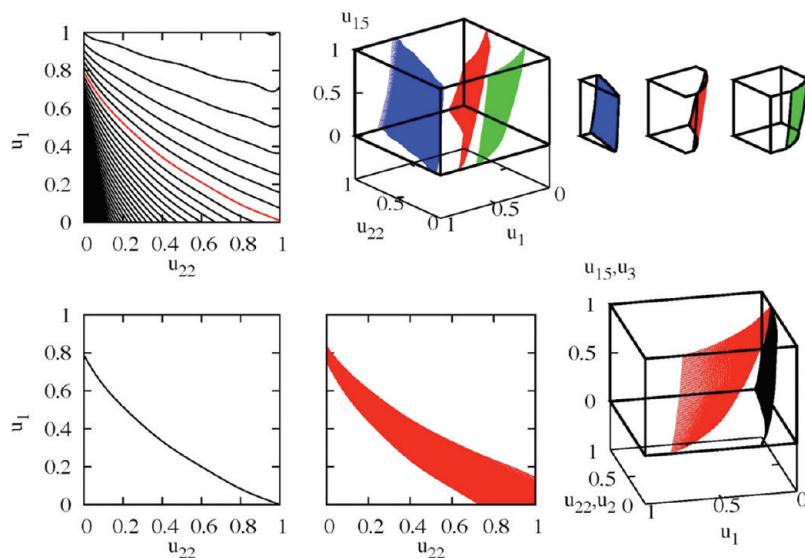
**Figure 20.** Full pdf at 2 bar (top) and pdfs calculated from 2-D (middle) and 3-D expansions (bottom). See text for further details.

reaction expansion (eq 4.8a) using reactions 1 and 22. The contour at  $\tau = 0.00125$  s is drawn as a red line in this panel. This contour, like the one in Figure 19, ends at a corner ( $u_{22}, u_1 = (1.0, 0.0)$ ) and marks a change in shape of the region enclosed by the contour from triangular to trapezoidal (Appendix D). The analysis presented in eqs D.1a–D.7 of Appendix D demonstrates that this can lead to changes in the ignition-time pdf, which is what is observed in the middle panel of Figure 20 at  $\tau = 0.00125$  s.

A comparison of the contours of the two-reaction and three-reaction cases at  $P = 2$  bar demonstrates how the difference between the 2-D and 3-D pdfs arises, and thus describes how the difference between the 2-D and the full pdf arises. For three-reaction expansions the contours of the HDMR expansion are 2-D surfaces. It is straightforward to generate these surfaces from the expansion in eq 4.8b and three of these contours are shown in the top middle panel of Figure 21. The blue surface is at  $\tau = 0.0003$  s, the red is at  $\tau = 0.00125$  s, and the green is at  $\tau = 0.0025$  s. The red-surface contour for the three-reaction case corresponds to the corner line contour of the two-reaction case, drawn as a red curve on the top left plot of Figure 21.

The contour surfaces and the side of the cube drawn on the middle panel of Figure 21 enclose volumes that are presented on the top right of Figure 21. One edge of each volume matches one of the contour surfaces in the middle panel. The plots in the top right indicate that the volumes needed to define the cumulative distribution function change from a shape similar to a triangular prism on the left to a pentagonal prism in the third plot on the top right. The middle volume in Figure 21 is a hybrid, with the bottom part of the shape a trapezoidal prism and the top a pentagonal prism.

The change in the pdfs between the middle and bottom panel of Figure 20 results for two reasons. The first reason is that all 2-D contours of the 3-D function near the contour at  $\tau = 0.00125$  s show the same hybrid shape shown in the top row of Figure 21, and this shape causes a smoothing of the features observed in Figure 20. A second reason that is more important is that there is



**Figure 21.** Top left: contour plot of a two-reaction expansion at 2 bar. Middle top: three contour surfaces of a three-reaction expansion at this pressure. Top right: set of three shapes whose volumes describe three points along the cdf. Bottom row, left to right: corner contour of the two-reaction surface at 2 bar, projection of the contour surface of the three-reaction expansion for the same contour value, and comparison of the corner contour surfaces at  $P_0 = 2$  bar (red) and  $P_0 = 0.5$  bar (black).

a sufficiently strong dependence on the  $u_{15}$ -coordinate to cause a smoothing of the sharp feature. This is demonstrated with a comparison between the 1-D corner contour of the two-reaction surface on the lower left of Figure 21 with a projection of the 2-D contour surface of the three-reaction function onto the  $(u_{22}, u_1)$  plane in the middle bottom panel of Figure 21.

The behavior outlined here is different than the situation studied in Figures 17–19 at  $P_0 = 0.5$  bar. For that case the sharp feature for the 2-D ignition-time pdf persists for the full pdf. In the earlier case there is a much weaker dependence on the third reaction for the three-reaction case than there is at  $P_0 = 2$  bar. This is demonstrated on the bottom right of Figure 21. The black surface shows the corner contour for  $P_0 = 0.5$  bar, which has a very weak dependence on the third reaction (reaction 3) than does the red contour at  $P_0 = 2$  bar, where the third reaction is reaction 15. Because of these features the sharp feature at 0.00178 s for the ignition-time pdf at  $P_0 = 0.5$  bar persists in going from two to three reactions and persists for the full ignition-time pdf in Figure 16.

**C. Figure 8 Revisited.** The procedures presented in the previous two subsections allow a better understanding of the changes that occur in the pdf at  $P_0 = 2$  bar when the uncertainty in reaction 22 (Appendix A) is reduced as presented in Figure 8. We carry out the analysis at this pressure, because most of the difference between the two pdfs can be observed in the two-reaction case. There are changes that occur at this pressure in going from two to three reactions as discussed in the previous subsection. However, these do not contribute significantly to the changes that occur when the uncertainty is reduced.

Finite-difference versions of the two-reaction pdfs are shown in the top panel of Figure 22. These were previously presented in Figure 8 in scaled time units. The red curve is for  $f = 1.58$  and the black for  $f = 5.0$ . The feature that causes the variance to increase when the uncertainty in reaction 22 is decreased is the broad shoulder that begins near  $\tau = 1.6$  ms and ends near  $\tau = 4$  ms. A contour plot of the two-reaction case for  $f = 5$  is shown in the middle panel of Figure 22 as a series of 20 evenly spaced contours

that start in the upper right at 0.15 ms and end in the lower left at 7.7 ms.

This contour plot in the middle panel can be compared to the one for  $f = 1.58$  shown in the bottom panel. The contours drawn on the bottom panel were generated at the same set of values as the ones in the middle panel, but because the uncertainty in reaction 22 has been reduced, the range of  $\tau$  has been significantly reduced, and there are only 12 contours now. The contour values range from 0.54 ms for the topmost contour to 4.9 ms on the very bottom left.

The changes observed in the top panel in going from  $f = 5$  to  $f = 1.58$  result from a two-reaction function that is much less steep in the bottom panel than its counterpart in the middle panel. The lowering of the steepness leads to a reduced value of the area derivative (eq 6.2b). In addition, the low value of the slope leads to a very sharp feature near  $\tau = 3.2$  ms, consistent with the model of eq D.7, where the change in the steepness of the pdf is inversely proportional to the slope of the contour.

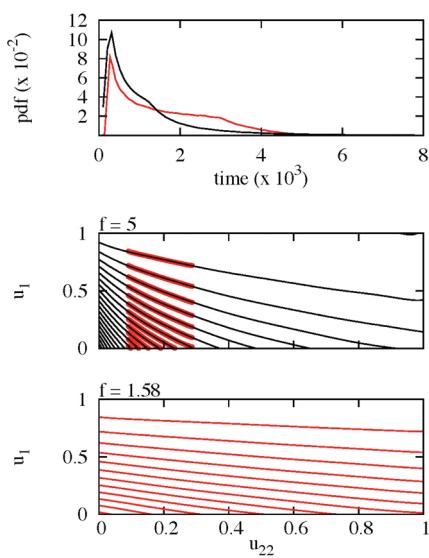
The results presented in Figures 8 and 22 at the reduced  $f$ -value of 1.58 were generated from a full calculation of the ignition properties as described in section II. However, because no reaction rates were changed with system update, it is not necessary to rerun the calculation. The response surface for the reduced  $f$ -value is merely a subset of the response surface for the larger  $f$ -value and we demonstrate this for the two-reaction contour surface in the middle panel.

The series of dots on the middle plot lie along the contours. These were generated from the contours of the surface in the bottom panel in the following manner. Define the two  $f$ -factors as  $f_1$  and  $f_2$ , with  $f_1 > f_2$ . The contours in the bottom panel were then scaled:

$$u_{22}(f=f_1) = a_1 + a_2 u_{22}(f=f_2) \quad (6.3)$$

where  $a_1$  and  $a_2$  are defined as

$$a_1 = \left( \frac{1}{f_2} - \frac{1}{f_1} \right) \frac{1}{q_1} \quad (6.4a)$$



**Figure 22.** Comparison of two two-reaction pdfs for the case of Figure 8 (top): (red curve)  $f = 1.58$ ; (black curve)  $f = 5$ . Contour plots of the two-reaction pdfs for  $f = 5$  are shown in the middle panel and for  $f = 1.58$  in the bottom panel. The contour values are discussed in the text. The dots on the middle panel are also discussed in the text.

$$a_2 = \frac{q_2}{q_1} \quad (6.4b)$$

where  $r_1$ ,  $q_1$ , and  $q_2$  are defined as

$$q_k = f_k - \frac{1}{f_k} \quad k = 1, 2 \quad (6.5)$$

These equations merely state that a response surface for a system where a single  $f$  is lowered and no reaction rates have been changed can be generated directly from the higher uncertainty surface with no additional calculations like those outlined in section II. The transformation of the higher-uncertainty surface to the lower-uncertainty surface is accomplished by trimming the first surface at the new uncertainty limits and then rescaling back to the range 0.0–1.0.

## VII. DISCUSSION AND CONCLUSION

Two previous papers have used global sensitivity analysis<sup>26,27</sup> to study the sensitivity of ignition delay times to individual reactions. The goal of those papers was to develop a prescription for theoretical validation of chemical-kinetic mechanisms. The screening based on global sensitivity analysis led to the re-evaluation of two important rate coefficients, which led to improved estimates of ignition delay times and the reduction of the uncertainty of the ignition delay times. The success of those two papers has led us to examine in more detail several features used there in this paper.

The H<sub>2</sub>/O<sub>2</sub> mechanism was chosen for study here because it is important, being a submechanism of any hydrocarbon combustion, and because it is small and allows for detailed examination. The global sensitivity analysis demonstrated that there was a distinct transition in the ignition sensitivity. For a fixed temperature we showed that there was a transition from H + O<sub>2</sub> = OH + O to H + H<sub>2</sub>O<sub>2</sub> = H<sub>2</sub> + HO<sub>2</sub>, as the initial pressure was varied.

We demonstrated that this transition led to a transition in how the global sensitivity analysis performed. At low pressure, when the first reaction dominated, the first-order sensitivity coefficients described most of the variance in the ignition delay times. This situation also occurs at high pressure when the second reaction dominates. However, at intermediate pressures the two reactions compete and, to account for all the variance in the data, it is necessary to include second-order terms in the HDMR expansion used to generate the global sensitivity analysis.

Many of the interesting results reported here result from reaction pathways that typically occur for ignition processes in H<sub>2</sub>/O<sub>2</sub> mixtures (see for example, ref 49). As conditions change, the pathways can change, as noted in the previous paragraph. However, unlike a typical modeling calculation for hydrogen combustion where transitions are abrupt, the ignition process is studied here under uncertain rate parameters, and the shape of the ignition-time distribution reflects the degree to which given pathways contribute. Therefore, the ignition-time pdf becomes a useful object to generate and study.

We demonstrated in section IV that a good way to reproduce the ignition-time pdf was to differentiate the cumulative distribution function. The cumulative distribution function was converted to the probability density function using two approximations to the derivative necessary for the transformation. A finite-difference form for this derivative generally performed well, but for sharp peaks there could be a breakdown. For such cases we used the Gaussian process model to interpolate the cdf. Analytical derivatives of the Gaussian process model were then used to generate the pdf. Appendix C has additional information about this procedure.

To further examine the transitions that occur as pressure is changed, we examined in more detail the way the ignition-time pdf (the true measure of the uncertainty) was expanded in terms of the HDMR components in section V. The expansion of the ignition-time pdf is a much more sensitive measure of the accuracy of the HDMR expansion than the amount of variance captured at a given order. It was demonstrated that even in cases where the first-order sensitivity indices account for 98% or 99% of the variance, there could be significant errors in the expansion of the pdf in terms of first-order terms. This led us to further study ways to “deconstruct” the ignition-time pdf into various components for accuracy and understanding. The results we have generated so far suggest that an expansion in a reaction-by-reaction basis (section VD) gives the most accurate, efficient, and insightful expansion of the ignition-time pdf. In addition, this expansion limits the amount of higher-order terms that need to be calculated, as described in section VA.

The deconstruction in section VD was used to create a geometrical interpretation in section VI of the cdfs and pdfs. This allowed a detailed understanding of the important features of the pdfs. It was shown that sharp changes in the pdf could be understood in terms of the geometry of the two-reaction and three-reaction deconstructions of the HDMR expansion, and the way these features change with pressure. Section VIC demonstrated that a higher uncertainty response surface could be transformed to a lower uncertainty response surface without any additional calculations, provided no reaction rates are changed and the new uncertainties fall within the range of the old uncertainties. The resulting pdf then can be generated from the new response surface. This procedure can be employed to separate the effects of reaction-rate change and uncertainty change when mechanisms are updated with a change in both a reaction rate and its uncertainty.

This paper and the previous two papers<sup>26,27</sup> have only varied the rate coefficient through the  $A$ -factor, ignoring more complicated variations that might occur. In particular, it would be interesting to extend the global sensitivity analysis employed here to the pressure dependence of those reactions whose rate coefficients are pressure-dependent. It would be interesting to include a more complex variation, because of the added uncertainty of the pressure dependence and we intend on studying this in the future.

We end this section with a brief discussion of the present approach compared to two earlier approaches. The first comparison is to the RS-HDMR approach of ref 18. In the present work and in ref 18, the various nonlinear functions needed to accomplish the expansion are written as a polynomial expansion (e.g., eqs 3.2a and 3.8). In the present approach, based on ref 28, the expansion coefficients are calculated by linear regression. In the RS-HDMR approach, the expansion is written in terms of orthogonal polynomials and the expansion coefficients are calculated by projection onto the polynomials and require the calculation of a multidimensional integral. It would be interesting to compare the relative efficiencies of these two approaches, but we have not done that yet.

A second approach that has been used that is different than the present approach is an expansion using polynomial chaos,<sup>50</sup> reviewed in ref 51 and discussed in further detail in ref 52. Reference 52 states that the polynomial chaos expansion is a form of HDMR expansion and can be more efficient than a typical HDMR expansion as long as the dimension of the underlying system is not too large. However, the expansion becomes increasingly computationally intensive as the dimension of a system gets larger, and even the size of the system used here (25) would be computationally intensive. Apparently, one of the reasons for this is that a polynomial chaos expansion incorporates higher-order terms in the expansion more readily and a typical HDMR expansion cuts off the higher order terms. This suggests that, when higher order terms are important, a polynomial chaos expansion may be preferred over other HDMR approaches, despite its computational cost.

## APPENDIX A: THE SYSTEM AND ITS UNCERTAINTIES

Table A1 lists all the reactions and their uncertainties. The reaction rates can be found in ref 35. The uncertainties (eq 3.1a)

were chosen on the basis of the references cited in the table, but all uncertainties were capped at  $f = 5.0$ , even though some of the literature values were higher. Because of the complexity of their reaction rates, two of the reactions ( $\text{HO}_2 + \text{HO}_2$  and  $\text{OH} + \text{H}_2\text{O}_2$ ) have rate expressions that are the sum of two Arrhenius forms. It was decided that these duplicate reactions (18/19 and 24/25) would be treated separately when the uncertainty was varied, and in fact, reactions 24 and 25 have different uncertainty factors.

In this paper and refs 26 and 27 the probability density function of the uncertainties has been taken to be uniform, a commonly used density for global sensitivity analysis.<sup>9</sup> One alternative to the uniform pdf for rate coefficients, which must be positive, is the log-normal pdf, used in ref 50 and discussed in detail in ref 58. Due to the relative uncertainty in the uncertainties of the rate coefficients (see the discussion in section 2.6 of ref 53), a uniform distribution seems to be a very reasonable choice, however, and that is what we have implemented in this paper.

## APPENDIX B: CONVERGENCE STUDIES

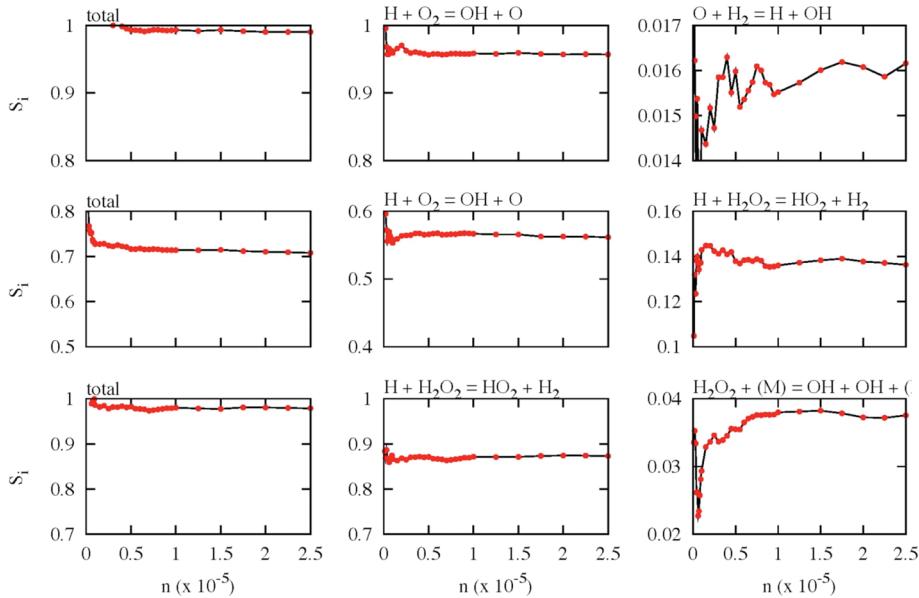
As noted in the text, the ignition-time pdf is a very sensitive measure of the convergence of the HDMR expansion. Therefore, we have done extensive studies of the convergence properties of the procedures used here, demonstrating the degree of convergence on the first-order sensitivity indices. These include convergence with the size of the sample used in the Monte Carlo simulation and the convergence of the polynomial fitting procedure. We have also investigated the possibility of overfitting as the polynomial order is increased, as was done in ref 28. We tried several standard techniques<sup>59</sup> for assessing whether overfitting occurred and have found no evidence of significant overfitting when the maximum polynomial order is 8, as it is in this paper.

**1. Sample Size.** In section IV it was stated that the error in first-order indices was approximately 0.001. This estimate is based on results like those shown in Figure 23. This figure shows the estimates of the sum of the first-order indices (column one) and the first-order indices for the top two reactions (columns 2 and 3) at three pressures (each row) that have been studied in this paper. These plots show the first-order sensitivity indices as calculated with the standard eighth-order polynomial fit used

Table A1. Reaction Uncertainties

reaction	$f^a$	reaction	$f^a$
1. $\text{H} + \text{O}_2 = \text{O} + \text{OH}$	1.26 <sup>53</sup>	14. $\text{HO}_2 + \text{H} = \text{H}_2 + \text{O}_2$	2.00 <sup>54</sup>
2. $\text{O} + \text{H}_2 = \text{H} + \text{OH}$	1.58 <sup>53</sup>	15. $\text{HO}_2 + \text{H} = \text{OH} + \text{OH}$	2.00 <sup>54</sup>
3. $\text{H}_2 + \text{OH} = \text{H}_2\text{O} + \text{H}$	2.00 <sup>54</sup>	16. $\text{HO}_2 + \text{O} = \text{O}_2 + \text{OH}$	3.16 <sup>57</sup>
4. $\text{O} + \text{H}_2\text{O} = \text{OH} + \text{OH}$	2.50 <sup>54</sup>	17. $\text{HO}_2 + \text{OH} = \text{H}_2\text{O} + \text{O}_2$	3.16 <sup>57</sup>
5. $\text{H}_2 + \text{M} = \text{H} + \text{H} + \text{M}$	3.00 <sup>54</sup>	18. $\text{HO}_2 + \text{HO}_2 = \text{H}_2\text{O}_2 + \text{O}_2$	5.00 <sup>55</sup>
6. $\text{H}_2 + \text{Ar} = \text{H} + \text{H} + \text{Ar}$	3.00 <sup>54</sup>	19. $\text{HO}_2 + \text{HO}_2 = \text{H}_2\text{O}_2 + \text{O}_2$	5.00 <sup>55</sup>
7. $\text{H}_2 + \text{He} = \text{H} + \text{H} + \text{He}$	3.00 <sup>54</sup>	20. $\text{H}_2\text{O}_2(+\text{M}) = \text{OH} + \text{OH}(+\text{M})$	3.16 <sup>57</sup>
8. $\text{O} + \text{O} + \text{M} = \text{O}_2 + \text{M}$	2.00 <sup>54</sup>	21. $\text{H}_2\text{O}_2 + \text{H} = \text{H}_2\text{O} + \text{OH}$	5.00 <sup>54</sup>
9. $\text{O} + \text{O} + \text{Ar} = \text{O}_2 + \text{Ar}$	2.00 <sup>54</sup>	22. $\text{H}_2\text{O}_2 + \text{H} = \text{HO}_2 + \text{H}_2$	5.00 <sup>54</sup>
10. $\text{O} + \text{O} + \text{He} = \text{O}_2 + \text{He}$	2.00 <sup>54</sup>	23. $\text{H}_2\text{O}_2 + \text{O} = \text{OH} + \text{HO}_2$	3.00 <sup>54</sup>
11. $\text{O} + \text{H} + \text{M} = \text{OH} + \text{M}$	5.00 <sup>54</sup>	24. $\text{H}_2\text{O}_2 + \text{OH} = \text{HO}_2 + \text{H}_2\text{O}$	1.26 <sup>56</sup>
12. $\text{H} + \text{OH} + \text{M} = \text{H}_2\text{O} + \text{M}$	2.00 <sup>54</sup>	25. $\text{H}_2\text{O}_2 + \text{OH} = \text{HO}_2 + \text{H}_2\text{O}$	5.00 <sup>57</sup>
13. $\text{H} + \text{O}_2(+\text{M}) = \text{HO}_2(+\text{M})$	3.16 <sup>57</sup>		

<sup>a</sup> This is used as defined in eq 3.1a.



**Figure 23.** Convergence as a function of sample size is shown here. The first column shows how the sums of the first-order sensitivity indices converge and the other two columns how select individual reactions change with sample size. Pressures are  $P_0 = 1.07$  bar,  $P_0 = 1.59$  bar, and  $P_0 = 5.62$  bar, in the top, middle, and bottom rows, respectively.

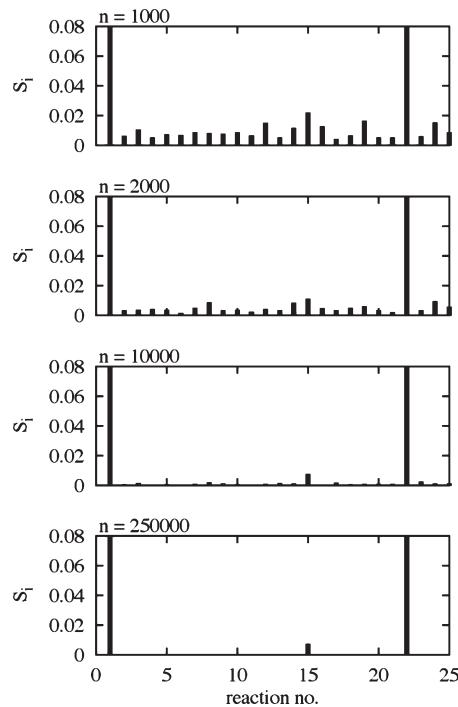
throughout the paper for sample sizes up to 250 000, the size used throughout the paper. It can be observed that the convergence in these plots is 0.001 or lower, and thus 0.001 is the estimated error used throughout the paper for the first-order indices.

The convergence of the sensitivity indices can be observed in the sensitivity spectrum (see Figure 4), which is shown on a much-reduced  $y$ -scale in Figure 24. These plots show results at  $P_0 = 1.59$  bar, which is the middle spectrum in Figure 4 and the middle row of Figure 23. As in the more converged result at the bottom and in Figure 4, the two largest sensitivity indices are for reactions 1 and 22.

These plots demonstrate that for the most part the screening process is complete at a sample size of 1000. Our general experience is that a sample size of approximately a factor of 40 more than the number of reactions is sufficient for reasonable screening provided that the largest sensitivity indices are at least 0.3.

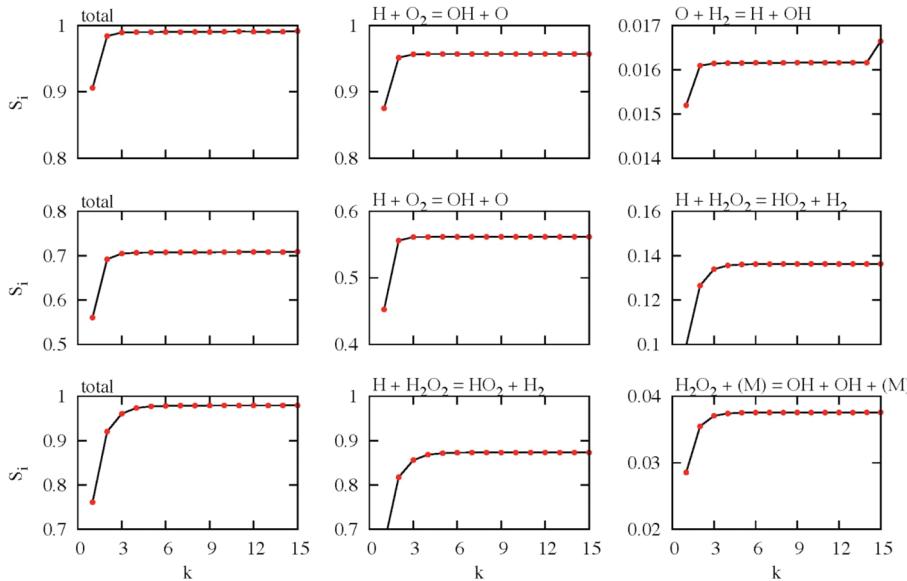
Figure 24 demonstrates that there is little change in the spectrum for sample sizes greater than 10 000. However, the accuracy of other measures in this paper, for example the deconstructions in section V and the study of second-order effects, needs large sample sizes. In most of our applications, where screening is the important task<sup>26,27,29</sup> much smaller sample sizes relative to the mechanism are used. For the methanol ignition studies in ref 26 and 27 a sample size of 20 000 points was used for a mechanism with 93 reactions and in an upcoming study of *n*-butanol ignition sample sizes of 50 000 were typically used for 1446 reactions.

**2. Polynomial Order.** The convergence with the degree of the polynomial is shown in Figure 25 for the same three sets as Figure 23, for the full set of 250 000 points. These plots demonstrate that convergence is generally very good at fourth order with accuracies comparable to what was expected for the largest sample sizes in Figure 23. They demonstrate that convergence is generally complete at sixth order. We have used



**Figure 24.** Convergence of the first-order sensitivity spectrum as a function of sample size is shown at  $P_0 = 1.59$  bar. The sample sizes are listed on the upper left of each panel.

eighth order throughout the paper to err on the conservative side. The top right panel demonstrates that overfitting errors can occur for the highest polynomial orders (here 15), as evidenced by the sudden increase in the first-order index there. We have generated many plots like those in Figure 25 and have never observed any effects like that at orders less than 12 for the 250 000-point samples.



**Figure 25.** Convergence with polynomial order of the first-order sensitivity indices are shown here. These are the same cases as Figure 23, for the full sample size of 250 000.

### APPENDIX C: OPTIMAL DENSITY ESTIMATION USING THE GAUSSIAN PROCESS MODEL

In the previous work of Ziehn and Tomlin<sup>19,20</sup> a kernel density estimator was used. In refs 26 and 27 we used a density estimation based on histograms. Section IV of this paper proposed using the derivative of the cumulative distribution function to generate the ignition-time pdf, a common practice. This approach permits an “optimal bandwidth” density estimation starting from an accurate interpolator for the cumulative distribution whose parameters can be chosen to avoid bias. Procedures for generating an optimal bandwidth using kernel density estimation (see the application in ref 45) are time-consuming for large data sets.

The procedure we use to make an optimal bandwidth calculation of the ignition-time pdf employs a Gaussian process (GP) model<sup>46,47</sup> of the cumulative distribution function. This estimator of the function is also referred to as “ordinary Kriging”.<sup>47,60</sup> Our development of the GP model follows most closely that of refs 47 and 61. The cumulative distribution functions in Figure 9 are generated at  $m$  points ( $\tau$  in eq 4.1) for a large sample of size 250 000. The  $m$  points are used as input to generate the GP model. Because the GP model used here is an interpolator, it goes through all  $m$  points. The estimator for a new point along the cumulative distribution function ( $\tau'$ ) using the GP model is

$$\tau' = \tau_0 + \kappa^T \mathbf{K}^{-1} (\boldsymbol{\tau} - \mathbf{1}\tau_0) \quad (C.1)$$

where  $\tau_0$  is a constant,  $\boldsymbol{\tau}$  is a vector of  $m$  ignition delay times along the cumulative distribution function,  $\mathbf{K}$  is an  $m \times 1$  vector,  $\mathbf{K}$  is an  $m \times m$  matrix, and  $\mathbf{1}$  refers to a vector of  $m$  1's.  $\mathbf{K}$  and  $\kappa$  are a correlation matrix and vector, respectively. They are calculated assuming a specific correlation among the points in the system. Here we use the most common form of this correlation, which is a squared exponential correlation:<sup>46</sup>

$$K_{ij} = e^{-\theta(\tau_i - \tau_j)^2} \quad (C.2a)$$

$$\kappa_n = e^{-\theta(\tau' - \tau_n)^2} \quad (C.2b)$$

The estimator in eq C.1 is a mean function of a distribution of functions.<sup>46</sup> In addition to the estimator, the Gaussian process model also includes a mean squared error. This error estimate includes an additional parameter.<sup>46</sup>

The successful interpolation of the cumulative distribution function in eq C.1 depends only on the initial set of points and does not depend on  $\tau_0$  or  $\theta$ , and thus they are referred to as “hyperparameters”. However, poorly chosen values of  $\tau_0$  and  $\theta$  may lead to overfitting, meaning that the values of the estimator away from the initial set behaves poorly. To avoid this problem, we follow the common practice of choosing the hyperparameters via maximization of a log likelihood function (for example, ref 47). We use an “augmented” likelihood function, which allows the maximum of the likelihood to be calculated from the single parameter “ $\theta$ ”.<sup>47</sup> The other hyperparameter relevant to the interpolation,  $\tau_0$ , is a function of  $\theta$  and the set of  $m$  points used for the interpolation.

Equations C.1 and C.2a indicate that all the  $\tau'$ -dependence in the estimator is in the transpose of  $\mathbf{K}$ . Defining the following vector

$$\mathbf{f} = \mathbf{K}^{-1} (\boldsymbol{\tau} - \mathbf{1}\tau_0) \quad (C.3)$$

leads to the simple form of the estimator:

$$\tau' = \tau_0 + \kappa^T \mathbf{f} \quad (C.4)$$

The interpolation of the cumulative distribution function then leads to the calculation of the pdf through the derivative. The derivative of the estimator of eq C.4 is

$$\frac{dP(\tau')}{d\tau'} = \left( \frac{d}{d\tau'} \kappa^T \right) \mathbf{f} \quad (C.5a)$$

This can be written explicitly in summation form as

$$\frac{dP(\tau')}{d\tau} = \sum_{j=1}^n f_j \frac{d}{d\tau} [e^{-\theta(\tau' - \tau_j)^2}] \quad (C.5b)$$

where  $\tau'$  has been renamed as  $\tau$ . It is straightforward to calculate the derivative in eq C.5b:

$$\frac{dP(\tau')}{d\tau'} = -2\theta \sum_{j=1}^n (\tau' - \tau_j) f_j e^{-\theta(\tau' - \tau_j)^2} \quad (\text{C.5c})$$

Equation C.5c is the estimator used in most of the plots in section IV (for example, Figure 14).

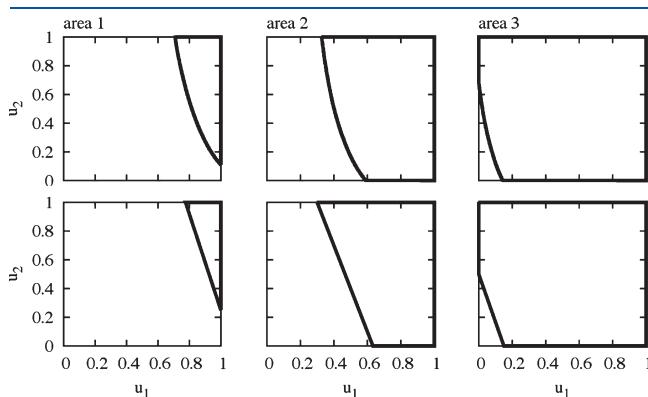
The density estimator in eq C.5c is a smooth and continuous function. The integration of this function over a range of ignition delay times leads to the following:

$$\int_{\tau_i}^{\tau_f} \frac{dP(\tau)}{d\tau} d\tau = P(\tau_f) - P(\tau_i) \quad (\text{C.6})$$

Equation C.6 generally guarantees that the integration of the pdf yields a number close to 1, because the limits we choose are such that  $P(\tau_i) = 0$  and  $P(\tau_f) \approx 1.0$ . This latter near equality can lead to some inaccuracy if there is a long tail for  $P(\tau)$ . If too many points are used in the tail to make the integral closer to 1, this may cause other portions of the  $P(\tau)$  to be inaccurate. In the cases studied here there is no problem in generating a function whose largest values are very close to 1, for example, 0.999 and higher. For cases where the representation of a long tail in the cdf is necessary, nonstationary versions of the Gaussian process model<sup>62,63</sup> might be necessary to retain accuracy over the whole domain.

## APPENDIX D: GEOMETRICAL MODEL FOR THE IGNITION-TIME PDF

A more quantitative application of the geometrical picture describes in more detail the sharp changes observed in the second panel of Figure 19 and the full pdf in Figure 16. Figure 26 isolates three of the regions from the top panel of Figure 19 in the top row and shows a set of polygons that are designed to mimic their shapes in the bottom row. This is a somewhat crude approximation of the contours in the top panel of Figure 19 but leads to a simple description of the discontinuity of the first derivatives observed in the ignition-time pdfs presented in section VIA that is qualitatively accurate. In fact, there are pressures and temperatures where this picture is quantitatively accurate, with the contours being more linear than they are in the top panel of Figure 26.



**Figure 26.** Top: three types of areas used to calculate the pdf in Figure 19. Bottom: linear versions of the areas, which are from left to right triangles, trapezoids, and pentagons.

The areas displayed in Figure 26 are formed from either the exact contours of  $\tau$  (top row) or the linear form (bottom row). These form one of the sides of a polygon that can have three shapes: triangular (left column), trapezoidal (center), or pentagonal (right). The changes in shape may lead to discontinuities in the derivative in eq 6.2b. These changes occur when a contour emanates from one of the corners of the unit cube. The contour that includes the corner at  $(u_1, u_2) = (1.0, 0.0)$  describes the switch from a triagonal shape to a trapezoidal shape and the contour that goes through  $(0.0, 1.0)$  describes the switch from trapezoidal to pentagonal. These contours are shown as dots on the ignition-time pdf in the second panel of Figure 19. We now explicitly demonstrate how the discontinuity at the trapezoidal/pentagonal switch occurs with a simple linear model.

Assume the contours have this form

$$u_2 = g(\tau) - \alpha u_1 \quad \alpha > 0 \quad (\text{D.1a})$$

or this form

$$u_1 = g(\tau) - \alpha u_2 \quad \alpha > 0 \quad (\text{D.1b})$$

where the contour is assumed to be linear and the slope is assumed to be the same for all contours and is negative. The intercept function,  $g(\tau)$ , changes with  $\tau$ . The area of the trapezoid formed by this line is

$$A^t(\tau) = P(\tau) = 1 - \frac{g(\tau)}{a} + \frac{1}{2a} \quad (\text{D.2})$$

The derivative of this expression gives the pdf

$$p^t(\tau) = -\frac{1}{a} \frac{dg}{d\tau} \quad (\text{D.3})$$

The area of the pentagonal shape is derived by subtracting the excluded triangle from the unit square and has the following value:

$$A^p(\tau) = P(\tau) = 1 - \frac{1}{2} \frac{[g(\tau)]^2}{a} \quad (\text{D.4})$$

The derivative of this expression gives the pdf,  $p^p(\tau)$ :

$$p^p(\tau) = -\frac{g(\tau)}{a} \frac{dg}{d\tau} \quad (\text{D.5})$$

The change in behavior occurs at  $(u_1, u_2) = (0.0, 1.0)$ , which requires that  $g(\tau) = 1.0$  there, guaranteeing that the pdf is continuous. However, the derivative of the pdfs are not the same at this point:

$$\frac{dp^t(\tau)}{d\tau} = -\frac{1}{a} \frac{d^2g}{d\tau^2} \quad (\text{D.6a})$$

$$\frac{dp^p(\tau)}{d\tau} = -\left[ \frac{g(\tau)}{a} \frac{d^2g}{d\tau^2} + \frac{1}{a} \left( \frac{dg}{d\tau} \right)^2 \right] \quad (\text{D.6b})$$

At the change point  $g(\tau) = 1$  and the difference between the derivatives is

$$\frac{dp^p(\tau)}{dt} - \frac{dp^t(\tau)}{dt} = -\frac{1}{a} \left( \frac{dg}{d\tau} \right)^2 \quad (\text{D.7})$$

This expression is consistent with what is observed. The slope along the pdf is negative at the change point and becomes more negative after the change point. This expression is within a factor of 2 from what is observed for the ignition-time pdf.

A similar analysis can be done for the transition from triangular to trapezoidal, which is shown in the first two columns of Figure 26. The area of the triangular region within the linear approximation is

$$A^{\text{tr}}(\tau) = P(\tau) = \frac{1}{2a} [1 + a - g(\tau)]^2 \quad (\text{D.8})$$

The ignition-time pdf and its derivatives are

$$p^{\text{tr}}(\tau) = -\frac{1}{a} [1 + a - g(\tau)] \frac{dg}{d\tau} \quad (\text{D.9a})$$

$$\frac{dp^{\text{tr}}(\tau)}{d\tau} = \frac{1}{a} \left( \frac{dg}{d\tau} \right)^2 - \frac{1}{a} [1 + a - g(\tau)] \frac{d^2g}{d\tau^2} \quad (\text{D.9b})$$

At the change point,  $(u_1, u_2) = (1.0, 0.0)$ ,  $g(\tau) = a$ , so this derivative is

$$\frac{dp^{\text{tr}}(\tau)}{d\tau} = \frac{1}{a} \left( \frac{dg}{d\tau} \right)^2 - \frac{1}{a} \frac{d^2g}{d\tau^2} \quad (\text{D.10})$$

The difference in the derivatives for the trapezoidal region and the triangular region at  $(u_1, u_2) = (1.0, 0.0)$  is

$$\frac{dp^t(\tau)}{dt} - \frac{dp^{\text{tr}}(\tau)}{dt} = -\frac{1}{a} \left( \frac{dg}{d\tau} \right)^2 \quad (\text{D.11})$$

This demonstrates once again that there can be a significant change in the derivative. In this case the change is consistent with the peak observed in the middle panel of Figure 19.

## ACKNOWLEDGMENT

A.S.T. thanks Tilo Ziehn for helpful advice. This work was supported by the Division of Chemical Sciences, Geosciences, and Biosciences, the Office of Basic Energy Sciences, the U.S. Department of Energy, under contract number DE-AC02-06CH11357.

## REFERENCES

- (1) Westbrook, C. K.; Mizobuchi, Y.; Poinsot, T. J.; Smith, P. J.; Warnatz, J. *Proc. Combust. Inst.* **2005**, *30*, 125.
- (2) Zador, J.; Zseley, I. G.; Turanyi, T. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 1232.
- (3) Warnatz, J.; Maas, U.; Dibble, R. W. *Combustion*; Springer: New York, 2001. Kee, R. J.; Coltrin, M. E.; Glarborg, P. *Chemical Reacting Flow: Theory and Practice*; Wiley-Interscience: New York, 2003.
- (4) For an overview on sensitivity analysis, see : Saltelli, A.; Chan, K.; Scott, E. M. *Sensitivity Analysis*; John Wiley & Sons: Chichester, West Sussex, England, 2000.
- (5) Sobol, I. M. *MMCE* **1993**, *1* (4), 407–414.
- (6) Sobol, I. M. *Math Comput. Simul.* **2001**, *55*, 271–280.
- (7) Sobol, I. M. *Reliab. Eng. Syst. Saf.* **2003**, *79*, 187–193.
- (8) Sobol, I. M.; Kucherenko, S. *Wilmott Mag.* **2005**, *1*, 56–61.
- (9) For an overview of global sensitivity analysis, see, for example : Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Sasana, M.; Tarantola, S. *Global Sensitivity Analysis. The Primer*; John Wiley & Sons: Hoboken, NJ, 2008.
- (10) Cukier, R. I.; Fortuin, C. M.; Schuler, K. E.; Petschelt, A. G.; Schaiby, J. H. *J. Chem. Phys.* **1973**, *59*, 3873.
- (11) Cukier, R. I.; Fortuin, C. M.; Schuler, K. E.; Petschelt, A. G.; Schaiby, J. H. *J. Chem. Phys.* **1973**, *59*, 3879.
- (12) Cukier, R. I.; Schaiby, J. H.; Shuler, K. E. *J. Chem. Phys.* **1975**, *63*, 1140.
- (13) Cukier, R. I.; Levine, H. B.; Shuler, K. E. *J. Comput. Phys.* **1978**, *26*, 1.
- (14) McRae, G. J.; Tilden, J. W.; Seinfeld, John H. *Comput. Chem. Eng.* **1982**, *6*, 15.
- (15) Rabitz, H.; Alis, O. F.; Shorter, J.; Shim, K. *Comput. Phys. Commun.* **1999**, *117*, 11.
- (16) Rabitz, H. *J. Math. Chem.* **1999**, *25*, 197.
- (17) Li, G.; Rosenthal, C.; Rabitz, H. *J. Phys. Chem.* **2001**, *A105*, 7765–7777.
- (18) Li, G.; Wang, S.-W.; Rabitz, H. *J. Phys. Chem.* **2002**, *A106*, 8721–8733.
- (19) Ziehn, T. *Ph.D. Thesis*, The University of Leeds, Leeds, U.K., 2008.
- (20) Ziehn, T.; Tomlin, A. S. *Int. J. Chem. Kinet.* **2008**, *40*, 742.
- (21) Ziehn, T.; Tomlin, A. S. *Environ. Model. Soft.* **2009**, *24*, 775.
- (22) Ziehn, T.; Hughes, K. J.; Griffiths, J. F.; Porter, R.; Tomlin, A. S. *Combust. Theory Model.* **2009**, *13*, 589.
- (23) Li, G.; Rabitz, H.; Yelvington, P.; Oluwole, O.; Bacon, F.; Kolb, C.; Schoendorf, J. *J. Phys. Chem. A* **2010**, *114*, 6022.
- (24) Tomlin, A. S.; Ziehn, T. In *Coping with Complexity: Model Reduction and Data Analysis*; Gorban, A. N., Roose, D., Eds.; Springer: Berlin, 2010; p 9.
- (25) Scire, J. J., Jr.; Dryer, F. L.; Yetter, R. A. *Int. J. Chem. Kinet.* **2001**, *33*, 784.
- (26) Klippenstein, S. J.; Harding, L. B.; Davis, M. J.; Tomlin, A. S.; Skodje, R. T. *Proc. Combust. Inst.* **2011**, *33*, 351.
- (27) Skodje, R. T.; Tomlin, A. S.; Klippenstein, S. J.; Harding, L. B.; Davis, M. J. *J. Phys. Chem. A* **2010**, *114*, 8286.
- (28) Lewandowski, D.; Cooke, R. M.; Duintjer Tebbens, R. J. *ACM Trans. Model. Comput. Simul.* **2007**, *18* (1), Article 3.
- (29) Zhou, D. Y.; Skodje, R. T.; Davis, M. J. Manuscript to be submitted.
- (30) Dixon-Lewis, G.; Williams, D. J. *Comput. Chem. Kinet.* **1977**, *17*, 1.
- (31) For an overview see, for example: Konnov, A. A. *Combust. Flame* **2008**, *152*, 507.
- (32) Zseley, I. G.; Zador, J.; Turanyi, T. *Proc. Combust. Inst.* **2005**, *30*, 1273.
- (33) Zador, J.; Zseley, I. G.; Turanyi, T.; Ratto, M.; Tarantola, S.; Saltelli, A. *J. Phys. Chem. A* **2005**, *109*, 9795.
- (34) Zseley, I. G.; Zador, J.; Turanyi, T. *Int. J. Chem. Kinet.* **2008**, *40*, 754.
- (35) Li, J.; Zhao, Z.; Kazakov, A.; Dryer, F. L. *Int. J. Chem. Kinet.* **2004**, *36*, 566.
- (36) Li, J.; Zhao, Z.; Kazakov, A.; Chaos, M.; Dryer, F. L.; Scire, J. J., Jr. *Int. J. Chem. Kinet.* **2007**, *39*, 109.
- (37) Burke, M. P.; Chaos, M.; Dryer, F. L.; Ju, Y. *Combust. Flame* **2010**, *157*, 618.
- (38) Grcar, J. F.; Bell, J. B.; Day, M. S. *Proc. Combust. Inst.* **2009**, *32*, 1173.
- (39) Lutz, A. E.; Kee, R. J.; Miller, J. A. *Senkin: A Fortran program for predicting homogeneous gas phase chemical kinetics with sensitivity analysis*; SAND87-8248; Sandia National Laboratory: Livermore, CA, 1987.
- (40) For example: Chaos, M.; Dryer, F. L. *Int. J. Chem. Kinet.* **2010**, *42*, 143.
- (41) Kee, R. J.; Rupley, F.; Miller, J. A. *Chemkin-II: A Fortran chemical kinetics package for the analysis of gas-phase chemical kinetics*; SAND 89-8009; Sandia National Laboratory: Livermore, CA, 1990.

- (42) Efron, B.; Stein, C. *Ann. Stat.* **1981**, *9*, 596.
- (43) Cox, D. C. *IEEE Trans. Reliab.* **1982**, *R-31*, 265.
- (44) For a recent discussion with references, see, for example: Yu, W.; Harris, T. J. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 596.
- (45) Balakrishnan, S.; Roy, A.; Ierapetritou, M. G.; Flach, G. P.; Georgopoulos, P. G. *J. Hydrol.* **2005**, *307*, 204.
- (46) Rasmussen, C. E.; Williams, C. K. I. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, 2006.
- (47) Forrester, A. I. J.; Sobester, A.; Keane, A. *Engineering design via surrogate modeling. A practical guide*; Wiley: New York, 2008.
- (48) For a discussion of optimal bandwidths used in direct estimations of the density, see for example: Jones, M. C.; Marron, J. S.; Sheather, S. J. *J. Am. Stat. Assoc.* **1996**, *91*, 401.
- (49) Lee, D.; Hochgreb, S. *Int. J. Chem. Kinet.* **1998**, *30*, 385.
- (50) Reagan, M. T.; Najm, H. N.; Ghanem, R. G.; Knio, O. M. *Combust. Flame* **2003**, *132*, 545.
- (51) Najm, H. N. *Annu. Rev. Fluid Mech.* **2009**, *41*, 35.
- (52) Crestaux, T.; Le Maître, O.; Martinez, J.-M. *Rel. Eng. Syst. Saf.* **2009**, *94*, 1162.
- (53) Baulch, D. L.; Bowman, C. T.; Cobos, C. J.; Cox, R. A.; Just, Th.; Kerr, J. A.; Pilling, M. J.; Stocker, D.; Troe, J.; Tsang, W.; Walker, R. W.; Warnatz, J. *J. Phys. Chem. Ref. Data* **2005**, *34*, 757.
- (54) Tsang, W.; Hampson, R. F. *J. Phys. Chem. Ref. Data* **1986**, *15*, 1087.
- (55) Our estimate.
- (56) Atkinson, R.; Baulch, D. L.; Cox, D. L.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J. *Atmos. Chem. Phys.* **2004**, *4*, 1461.
- (57) Baulch, D. L.; Cobos, C. J.; Cox, R. A.; Esser, C.; Frank, P.; Just, T.; Kerr, J. A.; Pilling, M. J.; Troe, J.; Walker, R. W.; Warnatz, J. *J. Phys. Chem. Ref. Data* **1992**, *21*, 411.
- (58) Limpert, E.; Stahel, W.; Abbt, M. *BioScience* **2001**, *51*, 341.
- (59) Bishop, C. M. *Pattern Recognition and Machine Learning*; Springer: Berlin, 2006.
- (60) Cressie, N. A. C. *Statistics for Spatial Data*; Wiley: New York, 1993.
- (61) Jones, D. R. *J. Global Optimization* **2001**, *21*, 345–383.
- (62) Paciorek, C. J.; Schervish, M. J. *Environmetrics* **2006**, *17*, 483.
- (63) Gramacy, R. B. *J. Stat. Soft.* **2007**, *19*, 1.