Global Dynamic Optimization for Parameter Estimation in Chemical Kinetics

Adam B. Singer,[†] James W. Taylor, Paul I. Barton, and William H. Green*

Department of Chemical Engineering, MIT, 77 Massachusetts Ave., Cambridge, Massachusetts 02139

We present the first method guaranteed to find the best possible least-squares (χ^2) fit of experimental data by a nonlinear kinetic model. Several important advantages of knowing with certainty the best possible fit rather than a locally optimum fit are discussed and demonstrated using data from the recent literature. This is particularly important when the model and the data appear to be inconsistent. With the new method, one can rigorously demonstrate that a nonlinear kinetic model with several adjustable rate parameters is inconsistent with measured experimental data. The numerical method presented is a valuable tool in evaluating the validity of a complex kinetics model.

1. Introduction

Kineticists frequently desire to validate or disprove a proposed chemical reaction network by comparing it with experimental data. Models for chemical kinetics experiments are usually systems of nonlinear ordinary differential equations (ODEs) with several unknown adjustable parameters, p, posed as an initial value problem as shown in eq 1

$$\frac{dx}{dt} = f(t, x, q, p) \quad x(t_0) = x^0(q, p) \tag{1}$$

where the vector of state variables, x, usually includes time-dependent species concentrations or mass fractions and might also include other quantities such as temperature or density if they are time-varying during the experiment. The vectors q and p represent values that are constant throughout the process, such as rate constants for an isothermal system at atmospheric pressure. There are often dozens of numerical parameters q and p; here, we assume that most of these parameters, q, are well-established and can be safely held fixed, but some of them, p, are significantly uncertain and should be adjusted to improve the agreement between the model and the experiment.

Generally, the adjustable parameter vector, p, includes several unknown rate constants but could also include unknown molecular properties or initial conditions or both. Examples of uncertain initial conditions that could be adjusted to obtain a better fit are the temperature in a shock tube or the initial concentration of radicals formed in a flash photolysis experiment.

Once a kinetics model is formulated, the first question a kineticist asks is whether the model is consistent with experimental data. In other words, are there any physically reasonable values of p that would allow the kinetic model to match the data within its limits of uncertainty? The usual approach to try to answer this question is to vary p within a physically reasonable range to try to minimize the $\chi^2(p)$ defined in eq 2

$$\chi^2(p) = \sum_{i=1}^{N_{\text{data}}} \left(\frac{d_i - m_i(p)}{\sigma_i} \right)^2 \tag{2}$$

where d is the vector of data, m is the vector of model predictions, and σ is the vector of standard deviations. The value for d_i is the average over j replicate measurements, d_{ij}, such that $d_i = (1/N_{\text{replicates}})\sum_j d_{ij}$. The standard deviation used here is the unbiased indicator associated with the χ^2 distribution. The

$$\sigma_i = \left(\sqrt{\frac{1}{N_{\text{replicates}} - 1} \sum_{j=1}^{N_{\text{replicates}}} (d_{ij} - d_i)^2} \right) \tag{3}$$

model prediction, $m_i(p)$, is usually a simple linear function of the state variables at the time, t_i, where the corresponding data point, d_i, was measured, i.e., $m_i(x(t_i))$, where $x(t_i)$ is the solution of eq 1 for the specified choice of p.

An important value in the χ^2 distribution is the number of degrees of freedom in the system, ν, defined as the number of data points minus the number of adjustable parameters

$$\nu = N_{\text{data}} - N_{\text{parameters}} \tag{4}$$

For the large values of ν typical in kinetics, the χ^2 distribution is sharply peaked near its expectation value, $\langle \chi^2 \rangle = \nu$. If the model and the data are consistent, we expect that eq 2 will give a $\chi^2 \approx \nu$; if it gives a value significantly larger than ν, it is unlikely that the data and the model are consistent. If the model is correct, the probability that a data set would have a weighted sum of squares value greater than χ^2 is given by eq 5.

$$\Pr(\chi^2) = \int_{\chi^2}^{\infty} \frac{t^{\nu/2-1} e^{-t/2}}{2^{\nu/2} \Gamma\left(\frac{\nu}{2}\right)} dt = \frac{\Gamma\left(\frac{\nu}{2}, \frac{\chi^2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \tag{5}$$

The Pr value is the measure of likelihood that the data is consistent with the specified model, given the specified error, σ_i. The implicit assumptions associated with using this measure are that the model structure and all the numbers in the model are exactly correct and all the deviations between d_i and $m_i(p)$

* Corresponding author: whgreen@mit.edu.
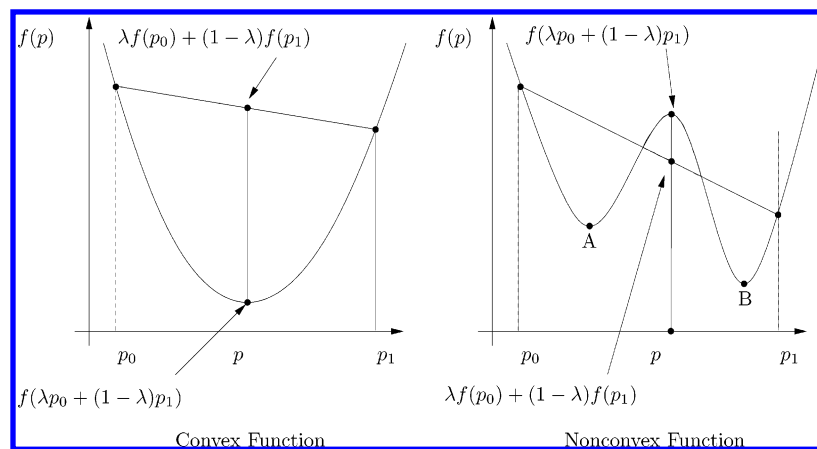† Present Address: ExxonMobil Upstream Research Company, Houston, TX.

Figure 1. A simple example of a convex and nonconvex function. All local minima of a convex function are also global minima. On the other hand, a nonconvex function may exhibit suboptimal local minima (A) in addition to a global minimum (B).[12]

are due to normally distributed random errors. If we measure data, d, that yield $\text{Pr}(\chi^2) > 0.75$, we have a lot of confidence that the data are consistent with the model, but if $\text{Pr}(\chi^2) < 0.25$, the data and the model are likely inconsistent. Intermediate values of Pr suggest that there would be value in repeating the experiment in order to reach an unambiguous conclusion.

If we could show that the Pr value is small, or correspondingly that $\chi^2(p)$ is large for *all* physically reasonable values of the adjustable parameters p, then we could say the model is inconsistent with the data, or equivalently that the data disproves the model. The most obvious way to do this would be to find the lowest possible $\chi^2(p)$ over the entire physically reasonable range of parameters. This can be accomplished by using an optimization algorithm to find the minimum of the $\chi^2(p)$.

Many algorithms exist for finding the minimum of objective functions such as χ^2. One of the best of these algorithms is the Levenberg–Marquardt (LM)[1,2] method. However, all of these methods have difficulty relating kinetic data to nonlinear chemical kinetic models for several reasons.

1. Usually several of the unknown parameters are correlated, i.e., the $\chi^2(p)$ surface is typically very flat in some regions of p space, and in these regions, its Hessian is near-singular, making it difficult for many methods to identify a minimum.

2. Experimental data often contain noise, which roughens the objective function surface. These rough patches can trap an optimization algorithm into a local minimum. That minimum, however, may not be a global minimum of $\chi^2(p)$ over the whole parameter space.

Until recently, it was impossible to determine if the minimum obtained from optimizing this class of problems was a local minimum or a global minimum. In many situations, despite the numerical problems, kineticists would nonetheless be able to find a set of parameters that made the model consistent with experimental data, i.e., $\text{Pr} > 0.75$, and immediately turn to determining the range of parameters that give good fits, i.e., the uncertainties in the parameter values, as discussed extensively in the literature.[3–6]

However, a kineticist that could not find a parameter set which made the model consistent with experimental data would be left in a quandary. Is the model an incorrect description of the chemistry or did I just not find the correct set of fitting parameters?

Here, we present the first numerical method which allows one to find a global minimum of $\chi^2(p)$ for nonlinear ODE IVP models that do not have analytical solutions, and so conclusively determine whether a proposed kinetic model is inconsistent with a set of experimental data. The method is demonstrated using

data and a kinetic model drawn from a recent study[7] of the reaction of cyclohexadienyl radical with oxygen.

Being able to reliably identify global optima would be useful in many problems that arise in physical chemistry. Because good numerical methods for global optimization of nonconvex problems were not available, the most common approach has been to restrict oneself to models which yield convex (often linear) optimization problems or to approximate the objective function with a convex response surface, so that all local minima are global minima.[5,6,8] This is probably one of the main reasons that "textbook" models which lead to linear least-squares optimizations are so popular. Unfortunately, many chemical systems are in reality nonlinear and nonconvex, so that approximations are required to make the resulting optimization problems convex, and it can be difficult to bound the errors introduced by these approximations. A popular approach which does not approximate the model or the objective function is multistart, a stochastic method where a large number of local optimizations are performed from various initial guesses, in the hope that at least one of them will hit the global optimum.[9] In practice, multistart and the convex approximation methods are effective ways to find good fits for cases where the model and the data are consistent. However, none of these methods are guaranteed to find the best fit and, therefore, cannot disprove a model with confidence.

Several global optimization methods exist for solving chemical engineering and kinetics problems.[10,11] However these methods require an explicit nonconvex algebraic model[10,11] and cannot solve problems that can only be expressed as a system of differential equations. Here, we present the first method suitable for the common situation where the objective function is only known implicitly, through the numerical solution of a system of differential equations.

2. Theory and Implementation

A simple example of local and global minima is shown in Figure 1. For an objective function that is convex, all local minima are also global minima; thus, a local optimization procedure will always find a global minimum. Most kinetic models, however, lead to nonconvex least-squares problems, or, as Figure 1 shows, problems containing a concave and a convex portion. When an objective function contains both concave and convex regions, suboptimal local minima can exist; thus, a local optimization procedure may not always find a global minimum.

The first algorithm guaranteed to find a global minimum was created in the late 1960s by Falk and Soland[13] and is called
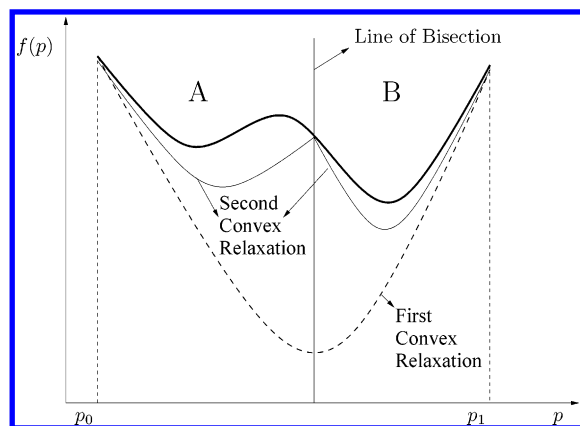
Figure 2. An example of a convex relaxation of a nonconvex function. The large parabola represents the first convex relaxation of the function. The region is then bisected at the convex minimum to form regions A and B. The relaxation is repeated to find new lower bounds for each region. By finding a point p in region B where $f(p)$ is less than the convex lower bound for region A, one can show that a global minimum cannot exist in region A. Region A can now be discarded from the search.

branch-and-bound. What branch-and-bound does is subdivide the region containing the feasible set into two separate regions. The algorithm then determines an upper and lower bound on the minima for each region. For example, in Figure 2, if the upper bound on the minimum of the objective function (bold curve) in region B is lower than the lower bound on the objective function in region A, then region A is rejected because it cannot contain a global minimum. Region B is then divided and the process repeated, narrowing in on the region in p space where a global minimum lies. Subdivision and evaluation of the upper and lower bounds continue until the bounds converge to within a specified tolerance on the global minimum.

For this algorithm to work, one must be able to determine upper and lower bounds on the minimum in any subregion, and the bounds must converge as the region is subdivided. Determining an upper bound on a global minimum is easy: choose any point p_{guess} in the range, then $\chi^2(p_{guess})$ is an upper bound on the minimum of $\chi^2(p)$. A tighter upper bound can found by performing a standard optimization starting from p_{guess} to find a local minimum. Unfortunately, it is much more difficult to determine provable lower bounds that tighten appropriately as the region is subdivided. Performing a local optimization does not guarantee a lower bound, since the region may contain suboptimal local minima.

The success of the new algorithm used here comes from its ability to construct convex relaxations of $\chi^2(p)$ in any specified subregion, Figure 2. A convex relaxation, $c(p)$, is a curve that is convex and underestimates the objective function at all points

$$c(p) \le \chi^2(p) \quad \forall\, p \in [p_0, p_1] \qquad (6)$$

Because $c(p)$ is convex, if you find a local minimum at p_{min}, it is guaranteed to be a global minimum of c

$$c(p_{min}) \le c(p) \quad \forall\, p \in [p_0, p_1] \qquad (7)$$

Combining the two produces the following expression

$$c(p_{min}) \le \chi^2(p) \quad \forall\, p \in [p_0, p_1] \qquad (8)$$

Thus, any local minimum of the convex relaxation is a rigorous lower bound, which can be used in the branch-and-bound algorithm.

The hard part is to construct $c(p)$ so that it is guaranteed to be a convex relaxation of $\chi^2(p)$, even though $\chi^2(p)$ is only evaluated via the numerical solution of a system of nonlinear ODEs. Here, we give a brief synopsis of how this is done. All the details, with proofs, are given in refs 12 and 14−16. Information on global optimization in general can be found in the excellent book by Horst and Tuy.[17]

The procedure for constructing convergent convex relaxations for problems with nonlinear ODEs embedded involves a sequence of bounding operations. First, it is necessary to estimate the image of the parameter set under the solution of the ODEs, i.e., find pointwise in time $x^L(t)$ and $x^U(t)$ such that

$$x^L(t) \le x(t, p) \le x^U(t), \quad \forall\, p \in [p_0, p_1] \qquad (9)$$

the so-called *implied state bounds*. These estimates can be obtained via the classical theory of differential inequalities,[18] but if the ODEs are non-quasimonotone,[18] these estimates can be very weak. In ref 14, we show how a priori knowledge concerning physical bounds and solution invariants (e.g., overall mass conservation) can be incorporated into the differential inequalities in order to compute much tighter implied state bounds.

Once bounds for x and p are known, it is possible to construct, pointwise in time, convex and concave relaxations of the right-hand sides (RHSs) of eq 1 on the sets defined by the time varying bounds. Because the RHSs are typically elementary functions, established methods for constructing convex relaxations, such as that of ref 19, can be employed for this purpose. From the properties of convex and concave functions, any linearizations to these convex and concave functions will under- and overestimate the convex and concave relaxations, respectively. Hence, affine functions, functions of the form $c(t, p) = M(t)p + n(t)$, that under- and overestimate the RHSs of eq 1 on the bounding set can be constructed. These affine functions, in conjunction with the theory of differential inequalities,[18] are used to construct the RHSs of linear ODEs whose solutions under- and overestimate, respectively, the solution of eq 1 for all values of p in the parameter set shown in eq 9. As the upper and lower bounds on the range of p narrow, they converge to $x(t, p)$. With bounds on $x(t, p)$, it is now possible to construct bounds on $\chi^2(p)$, because $\chi^2(p)$ is a functional of x. Because the solutions of these linear ODEs are pointwise in time affine in p,[16] these solutions of the linear ODEs can be used in conjunction with the results of refs 19 and 16, to construct convex relaxations of point and integral form functionals defined in terms of functions of the state variables.[14] $\chi^2(p)$ is an example of a point form function (multiple points).

In addition, as the parameter set employed in this construction shrinks to a point, both the state bounds and the convex relaxations will converge to the solution of eq 1 and the original nonconvex functional, respectively.

The global optimizer (GDOC) for $\chi^2(p)$ consists of many distinct parts: a branch-and-bound algorithm, a local optimizer, a numerical integrator, software for constructing the convex relaxations, and a residual evaluator, which is used to evaluate the right-hand sides (RHSs) of the ODEs. The branch-and-bound performed on the parameter set is done using an in-house branch-and-bound library, libBandB version 3.2.[20] Local optimization is performed using NPSOL version 5.0.[21] NPSOL is a proprietary package and does not come with the GDOC distribution; instead, SLSQP, a free optimization code is included. Numerical integration is performed using an extended discontinuity-locked version of CVODES,[22] which is discussed in ref 12. Residual evaluation is performed via code generated
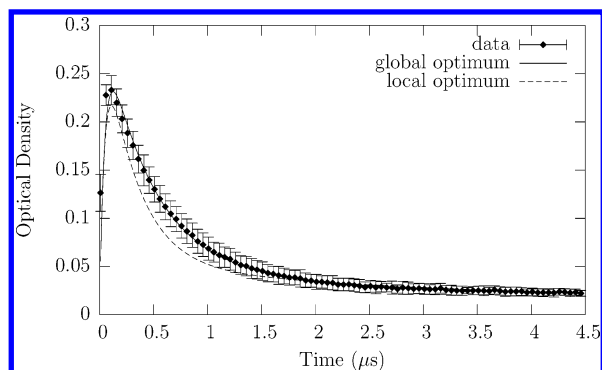
Figure 3. Globally and locally optimal fits to c-C_6H_7 absorption at 298 K with three adjustable parameters. Error bars shown are $\pm\sigma_i$, and for clarity, only every fifth data point is shown. The local optimum ($\chi^2 = 500$) is not significant at the 25% confidence level, while a global optimum ($\chi^2 = 128$) is consistent with the data (>99% confidence). In a previous work, Taylor et al.[7] used local optimization to fit these data and were unable to find a parameter set that provided a significant fit. The GDOC program, however, found a global solution for the three parameter fit which shows their model to be statistically significant.

by the compiler component of the GDOC package.[12] Together, these components form the global dynamic optimization code (GDOC) used in this work. To solve parameter estimation problems, the end user needs only to define a problem in the GDOC input language, which the GDOC package then parses.

The GDOC global optimization package and the example cases are available for download over the Internet at http://yoric.mit.edu/gdoc/. All of the details, including information on obtaining key numerical packages such as NPSOL and CVODES, which are neither developed nor distributed by the authors, are explained at http://yoric.mit.edu/gdoc/#WhatIsGDOC. The algorithm appears to scale as a low power of the number of state variables but scales exponentially with the number of adjustable parameters. Hence, even models involving large reaction mechanisms with many reacting species can be handled, but only a relatively small number of parameters can be optimized simultaneously.

3. Case Study: Transient Absorption Experiments and Model

To illustrate the utility of the global optimization technique discussed in this paper, GDOC was applied to a parameter estimation problem studied extensively by Taylor et al.[7] They measured the reaction rate of resonantly stabilized cyclohexadienyl radicals with molecular oxygen in cyclohexane solvent using laser flash photolysis. The reaction appears to be first order in oxygen concentration, but there is a large discrepancy between the rate constants for decay of the cyclohexadienyl radical measured in the gas phase[23,24] and in solvents,[25,26] and several different kinetic models have been proposed to interpret the experimental findings. Many data traces were measured and compared with various kinetic models in the course of the wide-ranging study of Taylor et al.; those interested in the chemistry should see the original paper and the doctoral thesis.[7,27] Here, we focus on two particular data traces measured in the course of that work, to see whether they can disprove one of the proposed kinetic models.

The first data trace, Figure 3, was measured by photolyzing a 298 K solution containing 0.4 M 1,4-cyclohexadiene and 0.1 M di-*tert*-butyl peroxide in equilibrium with 0.2 bar O_2 and 0.8 bar Ar. The second data trace, Figure 4, was measured under identical conditions but at 323 K. The spectrometer averages 30 individual laser shots, then sends the data set to the computer.
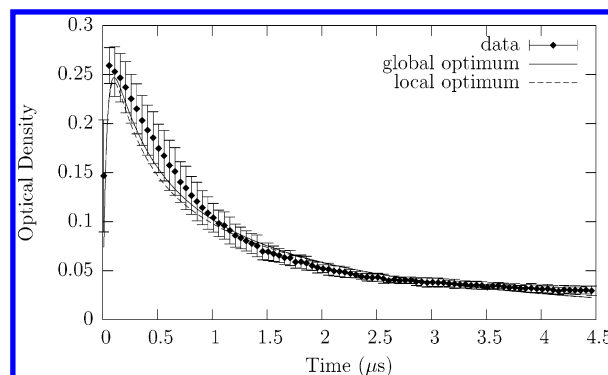

Figure 4. Globally and locally optimal solutions to a fit of c-C_6H_7 absorption at 323 K with three adjustable parameters. Error bars shown are $\pm\sigma_i$, and for clarity, only every fifth data point is shown. The local optimum ($\chi^2 = 500$) corresponds to a 5% level of confidence in the model fit, while the global optimum ($\chi^2 = 476$) corresponds to only a 16% level of confidence. Since the GDOC program was unable to find a global optimum with a level of confidence greater than the cutoff of 25%, we can reject the model fit of the data. There are no sets of parameters which will improve the fit.

TABLE 1: Proposed Kinetic Model[a]

no.	reaction	k_{298} [$M^{-1}\mu s^{-1}$ or μs^{-1}]
1	$(CH_3)_3CO + 1,4\text{-}C_6H_8 \rightarrow c\text{-}C_6H_7 + (CH_3)_3COH$	53
2	$c\text{-}C_6H_7 + O_2 \rightleftharpoons p\text{-}C_6H_7OO$	[1, 1200]
3	$c\text{-}C_6H_7 + O_2 \rightleftharpoons o\text{-}C_6H_7OO$	[1, 1200]
4	$o\text{-}C_6H_7OO \rightarrow C_6H_6 + HO_2$	[0.001, 100]
5	$2\ c\text{-}C_6H_7 \rightarrow$ products	1200

[a] k_1 and k_5 were fixed. k_2, k_3, and k_4 were adjusted to fit the data within the stated bounds.

Equation 2 treats each of these data points as if they are single shots and not the result of 30 replicates. The means of the data are, thus, well-determined, and the standard deviations are slightly overestimated by this procedure. The laser flash photolyzes the peroxide to form *tert*-butoxy radicals, which rapidly abstract an H atom from the cyclohexadiene to form cyclohexadienyl radicals (c-C_6H_7). c-C_6H_7 concentration is then probed using the absorbance of 316 nm light. *Ortho* and *para*-cyclohexadienylperoxy radicals are formed in the reaction and also absorb weakly at 316 nm. The measured absorbance, $\boldsymbol{d}$, is modeled using eq 10.

$$m_i = (\epsilon_{c-C_6H_7}[c\text{-}C_6H_7](t_i) + \epsilon_{ortho}[o\text{-}C_6H_7OO](t_i) + \epsilon_{para}[p\text{-}C_6H_7OO](t_i))l \quad (10)$$

The proposed mechanism which Taylor et al.[7,27] were trying to disprove with the data is shown in Table 1. In this proposed mechanism, reactions 1, 4, and 5 were assumed to be irreversible, while reactions 2 and 3 were treated as reversible. In the fits, the rate constants were constrained to lie within a physically reasonable range.

There are too many parameters in the model to determine them all from one or even several transient absorption data traces. Taylor et al.[7,27] fixed several parameters on the basis of prior literature information or their own measurements or calculations, as shown in Tables 1 and 2. For example, the rate constants for reactions 1 and 5 were taken from the literature,[28,29] and the reverse rate constants for reactions 2 and 3 were set using equilibrium constants they computed (on the basis in part on photoacoustic calorimetry experiments in the literature[30]). In various ways, they measured ϵ, l, and the initial concentration

TABLE 2: Model Parameters Which Were Not Adjusted in the Fit from Refs 7 and 27

model parameter	298 K	323 K
$\epsilon_{c-C_6H_7}$	2100 M^{-1} cm^{-1}	2100 M^{-1} cm^{-1}
$\epsilon_{o-C_6H_7OO}$, $\epsilon_{p-C_6H_7OO}$	200 M^{-1} cm^{-1}	200 M^{-1} cm^{-1}
l	0.7 cm	0.7 cm
$[1,4-C_6H_8]_0$	0.400 M	0.400 M
$[(CH_3)_3CO]_0$	1.53×10^{-4} M	1.53×10^{-4} M
$[O_2]_0$	0.0019 M	0.0014 M
$[c-C_6H_7]_0$, $[o-C_6H_7OO]_0$, $[p-C_6H_7OO]_0$	0 M	0 M
$K_2 = k_2/k_{-2}$	2081 M^{-1}	385 M^{-1}
$K_3 = k_3/k_{-3}$	4162 M^{-1}	770 M^{-1}
k_1	53 M^{-1} μs^{-1}	73 M^{-1} μs^{-1}
k_5	1200 M^{-1} μs^{-1}	1750 M^{-1} μs^{-1}

of 1,4-C$_6$H$_8$, and they used the published[31] Henry's law coefficient for O$_2$ along with their flow measurements to determine its concentration.[7,27] In this model, it is assumed that the only radical formed by the laser flash is (CH$_3$)$_3$CO, i.e., $[c-C_6H_7]_0 = [C_6H_7OO]_0 = [HO_2]_0 = 0$. The concentration of (CH$_3$)$_3$CO was determined using the known absorption cross-section of DTBP[32] and a measurement of the UV light power density exiting the reaction cell. This left three adjustable parameters to be determined in the least-squares fitting procedure: k_2, k_3, and k_4. The dynamic model has six important state variables: $[(CH_3)_3CO]$, $[1,4-C_6H_8]$, $[c-C_6H_7]$, $[O_2]$, $[o-C_6H_7OO]$, and $[p-C_6H_7OO]$; these are the elements of $x(t)$.

The estimates for the variation in the data were taken by comparing three different experimental runs taken under identical conditions for each temperature. The average and standard deviations (eq 3) were used to create the data points with error bars in Figures 3 and 4. The error bars drawn consist of one standard deviation on either side of the data point on the y-axis.

The tolerances used by GDOC to search for a global optimum were RTOL = 10^{-4} and ATOL = 10^{-5}. For this particular problem, the absolute tolerance was the active stopping condition. The tolerance for NPSOL was set to 10^{-5}, and the ATOL for CVODES was set to 10^{-10}. A multistart analysis[33] showed that these two cases have hundreds of local minima. However, many of these local minima have similar p and objective (χ^2) values. Presumably, they arise because the noise in the data, d, roughens the $\chi^2(p)$ surface. From an analysis of the roughness,[12] we set the objective function tolerance for the branch-and-bound procedure equal to 10 (the smallest χ^2 values are O(100)).

4. Results

In refs 7 and 27, experimental data were compared to several different proposed kinetic models using conventional optimization methods to find local minima. The goal was to identify kinetic model(s) consistent with a large and varied set of experimental and theoretical information on the reactions of cyclohexadienyl radical. For clarity, we focus on an apparently simple subproblem that needs to be solved in order to achieve that goal: fitting a kinetic data trace to a single model.

For the data set shown in Figure 3, the best local fit obtained previously[7,27] using the kinetic model detailed in Table 1 had a significant discrepancy from the experimental data. The authors assumed that this poor fit indicated that something was wrong with the structure of the model and published several speculations about possible failings in the model which could rationalize the discrepancy.[7,27] However, a globally optimized fit obtained in the present work fits the data set very well with a χ^2 value of 128, as shown in Figure 3. The model actually is consistent with the data, and there is no discrepancy that needs to be explained! The authors of refs 7 and 27 were misled, because
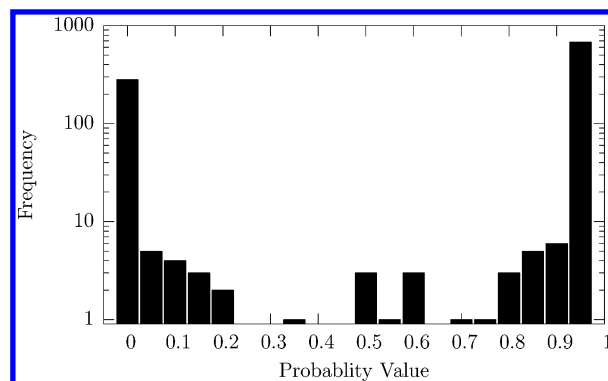

Figure 5. Histogram of the local minima found from nonlinear χ^2 fits to the data in Figure 3. Each of the 1000 different local optimizations performed began from a random initial guess spanning the physically reasonable range of the adjustable parameters. The local optima were binned on the basis of the probability the model is consistent with the data, i.e., by the level of confidence. At 298 K, many different parameter sets exist which provide significant fits (Pr > 0.75); however, there is a reasonable likelihood of finding fits that are not consistent with the data. Only 4 of the 1000 initial guesses converged to a globally optimal best fit.

their fitting program converged to a local minimum on the χ^2 surface significantly different from the true global minimum.

The previous study did not identify a global optimum of the objective function, since a good initial guess is needed to find the best overall fitted parameters. With the multistart algorithm, over 1000 different nonlinear local optimizations were run from random initial guesses spanning the physical boundaries of the parameters. Only 4 of the initial guesses found a global optimum, and $^1/_3$ of the guesses led to poor fits (Pr ≤ 0.25). A histogram of the 1000 multistart optima found is shown in Figure 5.

A second data set from the same transient absorption study,[7,27] measured at a higher temperature than the data in Figure 3, is displayed in Figure 4. Again, the fit obtained using the conventional local least-squares minimization approach does not fall within the experimental error bars. By using global optimization, the fit improves and has a χ^2 value of 476, which corresponds to a probability of 0.16. Even with this best possible choice for the three parameters, the deviations between the model and the data are significantly larger than expected from the error bars on the data. We can conclude with a high degree of confidence that the model is not consistent with the data for any choice of the parameters.

Again, by using the multistart algorithm, 1000 local optimizations of the model parameters were performed against the data shown in Figure 4. Figure 6 shows that finding the most significant optima is much more difficult in comparison to the 298 K data. The chances of randomly finding the best possible fit to the data are, literally, 1 in 1000.

There are of course many possible reasons why a model could be inconsistent with data, and this particular model is based on quite a large number of assumptions. In this particular case, it seems most likely that the heats of reaction for steps 2 and 3 computed in refs 7 and 27 are slightly off, causing the computed equilibrium constants K_2 and K_3 to have the wrong temperature dependence. Taylor et al.[7,27] noted that their computed heats of reaction could be in error by a few kcal/mol, more than enough to resolve the discrepancy in Figure 4. As discussed extensively in refs 7 and 27, there are many other possible problems with the model in Table 1. Even if it happened to perfectly match all the available data, that would not *prove* it was correct. What
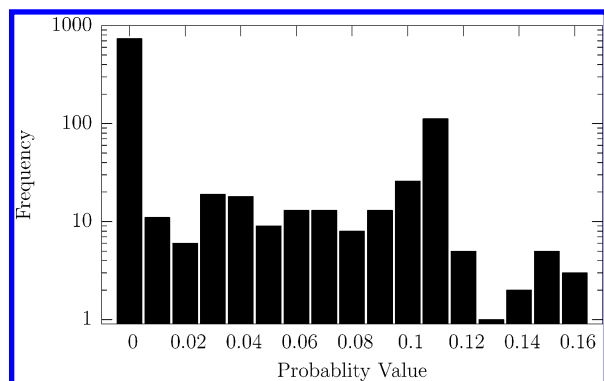
Figure 6. Histogram of the local minima found from nonlinear χ^2 fits to the data in Figure 4. Each of the 1000 different local optimizations performed began from a random initial guess spanning the physically reasonable range of the adjustable parameters. The local optima were binned on the basis of the probability the model is consistent with the data, i.e., by the level of confidence. Only 1 of the 1000 initial guesses converged to a globally optimal best fit. The best fit in this example has a $Pr(\chi^2) = 0.163$, which indicates that the model and the data are inconsistent.

is new is that there is a method for rigorously *disproving* a chemical kinetic model.

5. Conclusions

A computer program has been developed that is suitable for finding a global solution of the most common type of least-squares problem that arises in chemical kinetics, based on the recent breakthroughs in global dynamic optimization by Singer and Barton.[14-16] Using this new mathematical method, one can for the first time rigorously determine whether a model with several adjustable parameters can be made to fit kinetic data, or whether it is fundamentally inconsistent with the data. The model has been applied to two transient absorption traces measured in a recent study of cyclohexadienyl radical chemistry. In the 298 K example, the new method showed that the model was consistent with the data. However, it was proven that a globally optimum set of parameters was not consistent with the data at 323 K. This new addition to the kineticists' numerical tool kit is expected to prove useful for properly comparing models and data. The software described here and the associated documentation has been made available via the Internet at http://yoric.mit.edu/gdoc.

References and Notes

(1) Levenberg, K. *Q. Appl. Math.* **1944**, *2*, 164−168.

(2) Marquardt, D. *SIAM J. Appl. Math.* **1963**, *11*, 431−441.

(3) Hessler, J. P. *Int. J. Chem. Kinet.* **1997**, *29*, 803−817.

(4) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: Cambridge (U.K.) and New York, 1992.

(5) Frenklach, M.; Wang, H.; Rabinowitz, M. J. *Prog. Energy Combust. Sci.* **1992**, *18*, 47−73.

(6) Frenklach, M.; Packard, A.; Seiler, P.; Feeley, R. *Int. J. Chem. Kinet.* **2004**, *36*, 57−66.

(7) Taylor, J. W.; Ehlker, G.; Carstensen, H.-H.; Ruslen, L.; Field, R. W.; Green, W. H. *J. Phys. Chem. A* **2004**, *108*, 7193−7203.

(8) Bhattacharjee, B.; Schwer, D. A.; Barton, P. I.; Green, W. H. *Combust. Flame* **2003**, *135*, 191−208.

(9) Venkatesh, P. K.; Cohen, M. H.; Carr, R. W.; Dean, A. M. *Phys. Rev. E* **1997**, *55*, 6219−6232.

(10) Floudas, C. A.; Visweswaran, V. *Comput. Chem. Eng.* **1990**, *14*, 1397−1417.

(11) Adjiman, C. S.; Floudas, C. A. *J. Global Optim.* **1996**, *9*, 23−40.

(12) Singer, A. B. Global Dynamic Optimization. Thesis, Massachusetts Institute of Technology, 2004.

(13) Falk, J. E.; Soland, R. M. *Manage. Sci.* **1969**, *15*, 550−569.

(14) Singer, A. B.; Barton, P. I. *SIAM J. Sci. Comput.* In press.

(15) Singer, A. B.; Barton, P. I. *J. Global Optim.* In press.

(16) Singer, A. B.; Barton, P. I. *J. Optim. Theory Appl.* **2004**, *121*, 613−646.

(17) Horst, R.; Tuy, H. *Global Optimization;* Springer-Verlag: Berlin, 1993.

(18) Walter, W. *Differential and integral inequalities*; Ergebnisse der Mathematik und ihrer Grenzgebiete Springer-Verlag: Berlin, New York, 1970 (translated by Lisa Rosenblatt and Lawrence Shampine).

(19) McCormick, G. P. *Math. Programming* **1976**, *10*, 147−175.

(20) Singer, A. B. *LibBandB.a Version 3.2 Manual*; Technical Report; Massachusetts Institute of Technology, 2004.

(21) Gill, P. E.; Murray, W.; Saunders, M. A.; Wright, M. H. *User's Guide for NPSOL 5.0: A Fortran Package for Nonlinear Programming*; Technical Report; Stanford University, 1998.

(22) Hindmarsh, A. C.; Serban, R. *User documentation for CVODES, An ODE solver with sensitivity analysis capabilities*; Technical Report; Lawrence Livermore National Laboratory, Livermore, CA, 2002.

(23) Berho, F.; Lesclaux, R. *Phys. Chem. Chem. Phys.* **2001**, *3*, 970−979.

(24) Estupiñán, E.; Villenave, E.; Raoult, S.; Rayez, J. C.; Rayez, M. T.; Lesclaux, R. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4840−4845.

(25) Pan, X.-M.; Schuchmann, M. N.; von Sonntag, C. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 1021−1028.

(26) Maillard, B.; Ingold, K. U.; Scaiano, J. C. *J. Am. Chem. Soc.* **1983**, *105*, 5095−5099.

(27) Taylor, J. W. Direct Measurement and Analysis of Cyclohexadienyl Oxidation. Thesis, Massachusetts Institute of Technology, 2005.

(28) Effio, A.; Griller, D.; Ingold, K.; Scaiano, J.; Sheng, S. *J. Am. Chem. Soc.* **1980**, *102*, 6063−6068.

(29) Arends, I. W. C. E.; Mulder, P.; Clark, K. B.; Wayner, D. D. M. *J. Phys. Chem.* **1995**, *99*, 8182−8189.

(30) Kranenburg, M.; Ciriano, M. V.; Cherkasov, A.; Mulder, P. *J. Phys. Chem. A* **2000**, *104*, 915−921.

(31) Suresh, A. K.; Sridhar, T.; Potter, O. E. *AIChE J.* **1988**, *34*, 55−68.

(32) Bell, E. R.; Rust, F. F.; Vaughan, W. E. *J. Am. Chem. Soc.* **1950**, *72*, 337−38.

(33) Kan, A. H. G. R.; Timmer, G. T. *Math. Programming* **1987**, *39*, 57−78.