Simulated Annealing Search Algorithm for the Determination of Activation Energies and Arrhenius Prefactors from Limited Experimental Kinetic Data

David B. Terry, Jessica L. Bader, and Michael Messina*

Department of Chemistry, University of North Carolina at Wilmington, Wilmington, North Carolina 28405

Received June 26, 1999

We present a kinetics search algorithm that can be used to determine the activation energies and Arrhenius prefactors for the elementary reaction steps in an overall reaction scheme from a small set of concentration versus time data. The algorithm assumes the availability of a small set of experimental kinetic data for species concentrations for at least two different temperatures. The search algorithm is based upon the minimization of a so-called error function that is the squared difference between the experimentally measured species concentrations and species concentrations that are predicted from a guessed set of activation energies and Arrhenius prefactors that govern each of the elementary steps in the reaction mechanism. The activation energy search algorithm presented in this work is based on the Simulated Annealing heuristic minimization algorithm. We test the search algorithm on three types of chemical reactions: a consecutive reaction with a reversible step, a parallel reaction, and a representative chemical fate reaction.

I. INTRODUCTION

Important microscopic information about chemical reactions can be obtained through macroscopic measurements in the laboratory. This important microscopic information includes the reaction mechanism, the rate constants that govern the elementary reaction steps, and the activation energies that describe the temperature dependence of these rate constants. Once this microscopic information is known, one can make chemical predictions about (1) the concentration versus time profile for all species for all times and (2) how the concentration versus time profiles are effected by the temperature. The macroscopic experimental information that is used to uncover the important microscopic information usually takes the form of a set of concentration versus time measurements for either, all, or a subset of species involved in the chemical reaction.

Chemical reactions of current interest, such as biochemical or environmental fate reactions, can involve many complex elementary steps, thus making extensive profiling of concentration versus time information prohibitive and expensive. In previous work, we described a chemical kinetics search algorithm that can find the set of rate constants that govern an overall chemical reaction from a limited amount of experimental concentration versus time data.[1,2] The determination of these rate constants allows for the prediction of species concentrations for all times at the single temperature for which the experimental data were available. Experimental data measured at a single temperature do not allow for the prediction of chemical kinetic behavior as a function of temperature.

Knowledge of the temperature behavior of chemical reactions is important for two major reasons: (1) uncovering further microscopic information about the reaction such as the activation energies and (2) temperature control of chemical reactions. If, as is usual, rate constants are modeled according to the Arrhenius equation, one can use the temperature behavior of chemical reactions to uncover activation energies and Arrhenius prefactors. The activation energies describe the energetic barriers reactants must overcome to become products for each of the elementary steps in a reaction. The Arrhenius prefactors can describe the importance of steric factors in a chemical reactions such as how reactants must be aligned in order for a reaction to occur. Temperature can also be used as a control parameter that can be tuned to effect the speed of the reaction and/or the relative ratio of species concentrations at desired times. In order to bring about a rational use of temperature as a control parameter, the experimentalist needs to know the activation energies that govern each of the elementary reaction steps in the overall reaction scheme. It would be desirable, then, to develop a computational algorithm that can determine the Arrhenius activation energies, and prefactors, from a limited set of experimental concentration versus time data and well-specified reaction mechanism. We describe such an algorithm in the present work.

We base this Arrhenius parameter search algorithm (APSA) on the minimization of a well-defined error function. This error function is the squared difference between experimentally measured concentration versus time data and computed concentration versus time data. The experimental data must contain concentration versus time data for a subset of the species involved in the reaction for at least two different temperatures. The computed concentration versus time data are obtained by solving the kinetic equations for the reaction computationally based on a set of guessed Arrhenius activation energies and prefactors. This guessed set of Arrhenius activation energies and prefactors is then improved iteratively by minimizing this error function.

In previous work, we have found that this error function contains many local minima and it is only the global minimum (or a minimum nearby) that leads to accurate

* Corresponding author e-mail: messinam@uncwil.edu; fax (910) 962-3013.

10.1021/ci990332r CCC: $18.00
Published on Web 09/21/1999

results. We will show that the error function based upon guessed Arrhenius activation energies and prefactors is also pock-marked by many local minima. In previous work,[1,2] we have shown that this error function can be effectively minimized by a heuristic search algorithm such as Simulated Annealing[3−6] (SA) or the Genetic Algorithm[7] (GA). In this previous work, we critically compared the efficiencies of the SA, GA, and conventional minimizer (Simplex)[8] and found that the SA was the most efficient at getting near a useful global minimum in the error function. The conventional minimizer was found to get stuck in the local minimum closest to the initial guess and lead to inaccurate results. In this work, we base the APSA on the SA since we have found that it is most efficient at finding a useful local minimum and leads to accurate results. As discussed in previous work,[2] the SA algorithm that is used is a continuous minimization method that is based on the Simplex algorithm. At an annealing temperature of zero, the simulated annealing algorithm becomes the downhill Simplex algorithm. Thus, for the initial set of iterations, the SA algorithm gets near a useful global minimum while the annealing temperature is high and smoothly switches to a downhill Simplex algorithm to polish up the values of the search parameters.

The purpose of this work is to develop a computational algorithm based on the SA heuristic search algorithm that can predict the set of Arrhenius parameters that govern the elementary steps in an overall reaction mechanism from a set of limited concentration versus time data given the reaction mechanism. To facilitate this study, we test the APSA on three representative types of chemical reactions. These include a consecutive reaction with a reversible step, a parallel reaction, and a chemical fate reaction. For each of these reactions, we use a synthesized set of limited concentration versus time data for each of the reacting species and then use the APSA to find the complete set of Arrhenius parameters that govern the overall reaction from these data and the known reaction mechanism.

This paper is organized as follows. In section II, we sketch the general theory and equations that govern the APSA and the reaction schemes for the three types of reactions to be studied. In section III, we present numerical results that include the computational details of the algorithm and numerical results for each of these reactions. In section IV, we discuss the results and conclude.

II. THEORY

(A) General Theory. Given a general reaction scheme, we define two vectors: $\vec{\mathbf{c}}(t, T)$, the concentration of all the species at time t at temperature T, and $\vec{\mathbf{k}}(T)$, the set of rate constants governing the reaction which also depend on the temperature T. If, as usually assumed, the temperature dependence of the rate constants follows the Arrhenius equation, then the ith vector component of $\vec{\mathbf{k}}(T)$ can be written as

$$k_i(T) = A_i e^{-E_{A,i}/(RT)} \tag{1}$$

In eq 1, A_i, is the so-called Arrhenius prefactor that is usually assumed to be temperature independent, $E_{A,i}$ is the activation energy for rate constant i, and R is the gas constant expressed in J mol^{-1} K^{-1}. The ith vector component of $\vec{\mathbf{c}}(t, T)$, $c_i(t,T)$, represents the concentration of species i at time t and temperature T. The kinetic equation that governs the time evolution of the concentration of species i can then be written in terms of the unknown set of Arrhenius prefactors, which we denote as vector $\vec{\mathbf{A}}$, and the unknown set of activation energies, which we denote as $\vec{\mathbf{E}}_\mathbf{A}$

$$\dot{c}_i(t,T) = f_i[\vec{\mathbf{c}}(t,T),\vec{\mathbf{A}},\vec{\mathbf{E}}_A; t] \tag{2}$$

where f_i is a function dependent upon the instantaneous concentrations $\vec{\mathbf{c}}(t, T)$, the unknown Arrhenius prefactors, $\vec{\mathbf{A}}$, the unknown activation energies, $\vec{\mathbf{E}}_A$, and the time. The solution of the set, $\dot{\vec{\mathbf{c}}}(t,T)$, then yields the complete set of concentration versus time profiles, $\vec{\mathbf{c}}(t,T)$, for a given set of Arrhenius prefactors, $\vec{\mathbf{A}}$, and activation energies, $\vec{\mathbf{E}}_A$.

Say an incomplete set of experimental concentration versus time data is available at several different temperatures, which we term the target concentrations. For species i, there may exist target concentration data, $d_{i_j}(T_k)$, at several discrete times, t_{i_j}, and temperatures, T_k. The error function for species i is defined as

$$\epsilon_i(\vec{\mathbf{A}},\vec{\mathbf{E}}_A) = \sum_j \sum_k \left(\frac{c_i(t_{i_j}, T_k) - d_{i_j}(T_k)}{d_{i_j}(T_k)} \right)^2 \tag{3}$$

The error function for species i is the squared difference between the computed concentration, $c_i(t_j)$, for a guessed set of Arrhenius prefactors, $\vec{\mathbf{A}}$, and activation energies, $\vec{\mathbf{E}}_A$, and the target concentrations, d_{i_j} at the same times, t_{i_j}, and temperatures, T_k. As discussed in a previous work, the error function is normalized by the target concentrations. This prevents species that have higher concentrations during the course of the reaction from being weighted more heavily than species with lower concentrations. The total error function for the complete reaction then becomes

$$\epsilon(\vec{\mathbf{A}},\vec{\mathbf{E}}_A) = \sum_i \epsilon_i(\vec{\mathbf{A}},\vec{\mathbf{E}}_A) \tag{4}$$

This error function depends upon $\vec{\mathbf{A}}$ and $\vec{\mathbf{E}}_A$ since the time dependence of the concentrations, $\vec{\mathbf{c}}(t, T)$, is governed by the kinetic equations in eq 1.

The chemical kinetics search algorithm is comprised of the following steps: (1) make an initial guess of the set of $\vec{\mathbf{A}}$ and $\vec{\mathbf{E}}_\mathbf{A}$**;** (2) solve the set of differential equations for the set $\vec{\mathbf{c}}(t, T)$ based on these guesses according to eqs 2; (3) construct the error function from eqs 3 and 4; (4) minimize this error function to obtain updated values of $\vec{\mathbf{A}}$ and $\vec{\mathbf{E}}_A$; (5) use these values in step 1 and continue this iterative process until the error function is zero within some predefined tolerance.

The minimization of the error function is the critical step in the algorithm, and a mechanism is needed to perform this minimization. In previous work, we demonstrated that a conventional functional minimizer such as the gradient descent or downhill Simplex algorithm will converge toward the local minimum nearest the initial parameter guesses, which can lead to inaccurate results. Instead, the error function is minimized via a heuristic search algorithm. In a previous work, we critically compared the relative efficiencies of two different heuristic search algorithms, the SA and GA. We found that the SA is the more efficient heuristic

search algorithm for the rate constant search algorithm.[2] Thus, we base the activation energy search algorithm on the SA. To implement the chemical kinetics search algorithm using the SA, one supplies the function to be optimized, in this case the error function defined in eq 4 to the SA, and the SA finds the set of parameters, in this case the set of $\vec{\mathbf{A}}$ and $\vec{\mathbf{E}}_A$, that minimize the value of $\epsilon(\vec{\mathbf{A}},\vec{\mathbf{E}}_A)$.

(B) Chemical Reactions under Study. (i) Consecutive Reaction with a Reversible Step. The first type of reaction we will consider is a consecutive reaction with a reversible step. These reactions are ubiquitous in chemical kinetics, examples of these reactions including the Lindemann−Hinshelwood mechanism for unimolecular decay and the Michaelis−Menten mechanism of enzyme catalysis.[9] The reaction mechanism for this reaction type is shown below:

$$\mathrm{A} \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} \nu\mathrm{B}$$

$$\nu\mathrm{B} \xrightarrow{k_2} \mathrm{C} \tag{5}$$

In eq 5, ν is a stoichiometric coefficient that we will vary in order to test the algorithm for different orders of the elementary reactions. Each of the rate constants $\{k_1, k_{-1}, k_2\}$ can be written in the Arrhenius form as

$$k_1(T) = A_1 \mathrm{e}^{-E_{\mathrm{A},1}/(RT)} \tag{6a}$$

$$k_{-1}(T) = A_{-1} \mathrm{e}^{-E_{\mathrm{A},-1}/(\mathrm{RT})} \tag{6b}$$

and

$$k_2(T) = A_2 \mathrm{e}^{-E_{\mathrm{A},2}/(\mathrm{RT})} \tag{6c}$$

We will consider the set of Arrhenius prefactors, $\vec{\mathbf{A}} = \{A_1, A_{-1}, A_2\}$, and the set activation energies, $\vec{\mathbf{E}}_A = \{E_{A_1}, E_{A_{-1}}, E_{A_2}\}$, for each of the elementary steps as unknowns to be determined from the search algorithm. The differential equations whose solutions yield the complete concentration versus time profile at a particular temperature T are given by

$$\frac{\mathrm{d[A]}}{\mathrm{d}t} = -k_1(T)[\mathrm{A}] + k_{-1}(T)[\mathrm{B}]^{\nu} \tag{7a}$$

$$\frac{\mathrm{d[B]}}{\mathrm{d}t} = -k_2(T)[\mathrm{B}]^{\nu} - k_{-1}(T)[\mathrm{B}]^{\nu} + k_1(T)[\mathrm{A}] \tag{7b}$$

$$\frac{\mathrm{d[C]}}{\mathrm{d}t} = k_2(T)[\mathrm{B}]^{\nu} \tag{7c}$$

where the rate constants $\{k_1(T), k_{-1}(T), k_2(T)\}$ are defined in eqs 6a−c. Note that ν becomes the reaction order in the differential equations that govern the time evolution of the species.

(ii) Parallel Reaction. The second type of reaction we will study is a parallel reaction. Parallel reactions occur widely in chemical kinetics and can represent three types of general processes: (1) an initial reactant decomposes into different products, (2) a single substance reacts with several different reactants, or (3) a set of different reactants decompose into the same product.[9] We consider the parallel reaction of type 1 shown below:

$$\nu\mathrm{A} \xrightarrow{k_1} \mathrm{B}$$

$$\nu\mathrm{A} \xrightarrow{k_2} \mathrm{C}$$

$$\nu\mathrm{A} \xrightarrow{k_3} \mathrm{D} \tag{8}$$

Again, each of the three rate constants, $\{k_1, k_2, k_3\}$, can be written in the Arrhenius form as in eqs 6a−c. We will consider the set of Arrhenius prefactors, $\vec{\mathbf{A}} = \{A_1, A_2, A_3\}$, and the set activation energies, $\vec{\mathbf{E}}_A = \{E_{A_1}, E_{A_2}, E_{A_3}\}$, for each of the elementary steps as unknowns to be determined from the search algorithm. As for the consecutive reaction type, ν in eq 8 is a stoichiometric coefficient that we will vary in order to test the algorithm for different orders of the elementary reactions. The differential equations whose solutions yield the complete concentration versus time profile for the species are

$$\frac{\mathrm{d[A]}}{\mathrm{d}t} = -k_1(T)[\mathrm{A}]^{\nu} - k_2(T)[\mathrm{A}]^{\nu} - k_3(T)[\mathrm{A}]^{\nu} \tag{9a}$$

$$\frac{\mathrm{d[B]}}{\mathrm{d}t} = k_1(T)[\mathrm{A}]^{\nu} \tag{9b}$$

$$\frac{\mathrm{d[C]}}{\mathrm{d}t} = k_2(T)[\mathrm{A}]^{\nu} \tag{9c}$$

and

$$\frac{\mathrm{d[D]}}{\mathrm{d}t} = k_3(T)[\mathrm{A}]^{\nu} \tag{9d}$$

(iii) Chemical Fate Reaction. The last type of chemical reaction we will consider is a chemical fate process. The chemical fate reaction distinguishes itself from the above consecutive and parallel reactions in that there exist product channels to which the reactants can escape and the concentration in these particular product channels is unknown. Mathematically this means that the sum of the right-hand side of the differential equations that govern the evolution of the species concentrations does not sum to zero. This means that the values of the measured species concentrations are not conserved. To make this more clear, we consider the chemical fate mechanism shown below:

$$\mathrm{A} \xrightarrow{k_1} \mathrm{I} \tag{10a}$$

$$\mathrm{B} \xrightarrow{k_2} \mathrm{J} \tag{10b}$$

$$\mathrm{A} \xrightarrow{k_3} \mathrm{B} \tag{10c}$$

Here A and B are chemical reactants, and I and J are possible partitions for the fate of the reactants A and B. We assume that once A and B escape into product channels I and J, their concentrations cannot be measured. This means that the sum of the concentrations of species A and B is not conserved. We assume that the concentrations of species A and B can be measured, and the set of differential equations that govern their time dependence is

$$\frac{dA}{dt} = -k_1(T)[A] - k_3(T)[A] \quad (11a)$$

$$\frac{dB}{dt} = -k_2(T)[B] + k_3(T)[A] \quad (11b)$$

Careful inspection of eqs 11a and 11b indicates that the sum of the right-hand side is non-zero, meaning that the sum of the concentration of A and B is not conserved, which is not surprising since both reactants can escape into the product channels I and J. Once again, each of the three rate constants, $\{k_1, k_2, k_3\}$, can be written in the Arrhenius form as in eqs 6a−c. We will consider the set of Arrhenius prefactors, $\vec{\mathbf{A}} = \{A_1, A_2, A_3\}$, and the set of activation energies, $\vec{\mathbf{E}}_A = \{E_{A_1}, E_{A_2}, E_{A_3}\}$, for each of the elementary steps as unknowns to be determined from the search algorithm.

III. COMPUTATIONAL RESULTS

(A) Computational Details. In the actual application of the APSA, the target data would be measured experimentally. In order to test the ability of the above-described APSA on the three reaction types described in the previous section, we synthesize these target data computationally. We do this by choosing the sets of Arrhenius prefactors and activation energies for each reaction. We then solve the differential equations described in eqs 7, 9, and 11 computationally to generate a set of target data for each reaction. These differential equations were solved numerically via a fourth-order Runge−Kutta adaptive step-size integrator.[8]

The APSA algorithm is based on a SA minimization subroutine. The SA was run with an annealing schedule given by the following equation:

$$T_{k+1} = \lambda T_k \quad (12)$$

where T_{k+1} is the annealing temperature (not to be confused with the thermodynamic temperature) after N function evaluations. The initial temperature and temperature reduction factor, λ, were chosen as 10.0 and 0.67 for all examples, respectively. The temperature is reduced after $N = 100N$ moves, where N is the number of unknown parameters in the particular problem. The lower bounds on all parameters were set to zero, and the upper bounds on all parameters were initially set to 100 for the examples presented in this paper. Each time the upper bound was exceeded, we increased the upper bound and restarted the algorithm. For all reactions studied in this paper, we have found that 2×10^5 iterations took about 5 min on an SGI Indigo2 workstation.

(B) Numerical Results. (i) Consecutive Reaction. For the consecutive chemical reaction scheme that is outlined in Eq 5, we synthesized a limited set of known concentration data for the reaction by choosing the set of Arrhenius prefactors to have the values of $\vec{\mathbf{A}} = \{100\ s^{-1}, 500\ s^{-1}, 1000\ s^{-1}\}$ and the set of activation energies to have the values $\vec{\boldsymbol{E}}_A = \{5.74\ kJ\ mol^{-1}, 15.50\ kJ\ mol^{-1}, 28.72\ kJ\ mol^{-1}\}$. We choose this set of Arrhenius prefactors and Activation energies for all reaction types presented in this paper. [Note that the units on the Arrhenius prefactors are only consistent for the case when the stoichiometric coefficient $\nu = 1$. We keep the magnitudes of these prefactors the same when $\nu = 2$, but the reader should understand that the units on the prefactors will change]. Thus, the rate constants for this reaction span over 4 orders of magnitude. We studied this reaction for two values of the stoichiometric coefficient, $\nu = 1$ and $\nu = 2$, in eq 5. For each value of ν, we synthesized the set of known target data by solving the differential equations in eqs 7a−c with the initial concentration values of $\{[A]_0 = 1.0\ M, [B]_0 = 0, [C]_0 = 0\}$. We synthesized target data for three different temperatures, $T = 300$, 325, and 350 K. The concentration versus time profile for this reaction for $\nu = 1$ and $T = 300$ K is shown in Figure 1a.

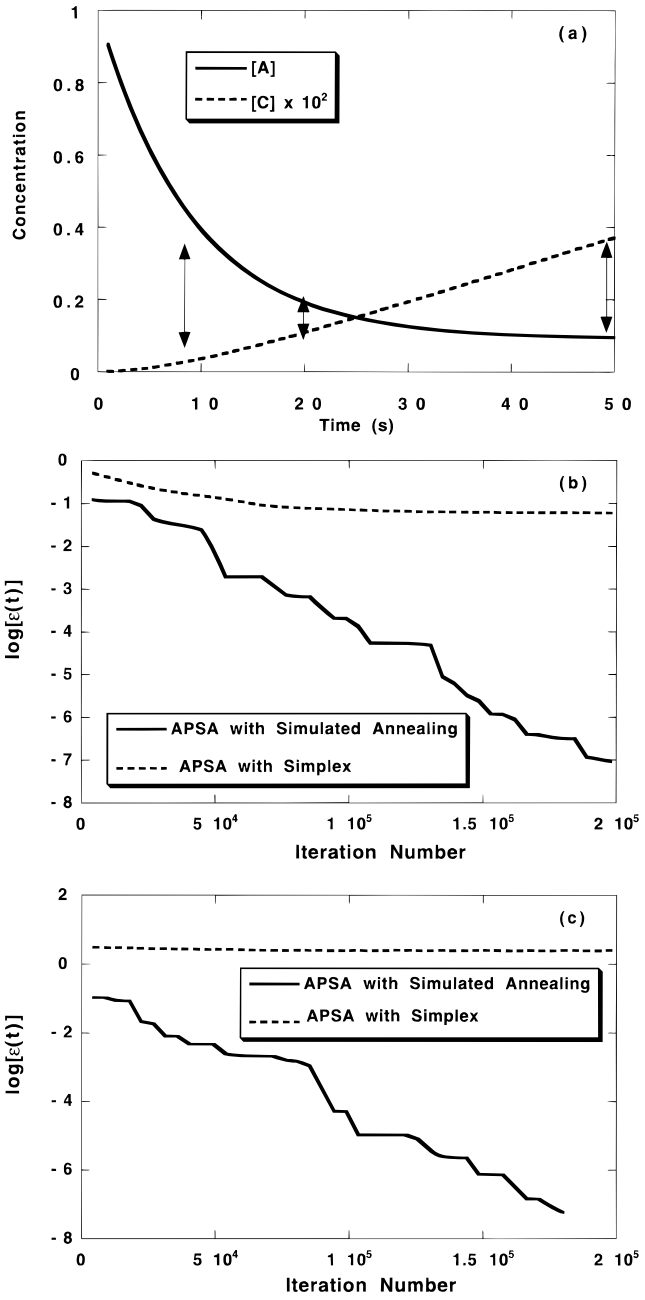

Figure 1. . (a) Concentration versus time profile for the consecutive reaction when $\nu = 1$ and $T = 300$ K. The solid and dashed lines show species A and C, respectively. The arrows indicate the target data. (b) Comparison of the convergence for the SA and Simplex-based APSA for the consecutive reaction when $\nu = 1$. The solid and dashed lines show the convergence of the SA and Simplex-based algorithms, respectively. (c) Comparison of the convergence for the SA and Simplex-based APSA for the consecutive reaction when $\nu = 2$. The solid and dashed lines show the convergence of the SA and Simplex-based algorithms, respectively.

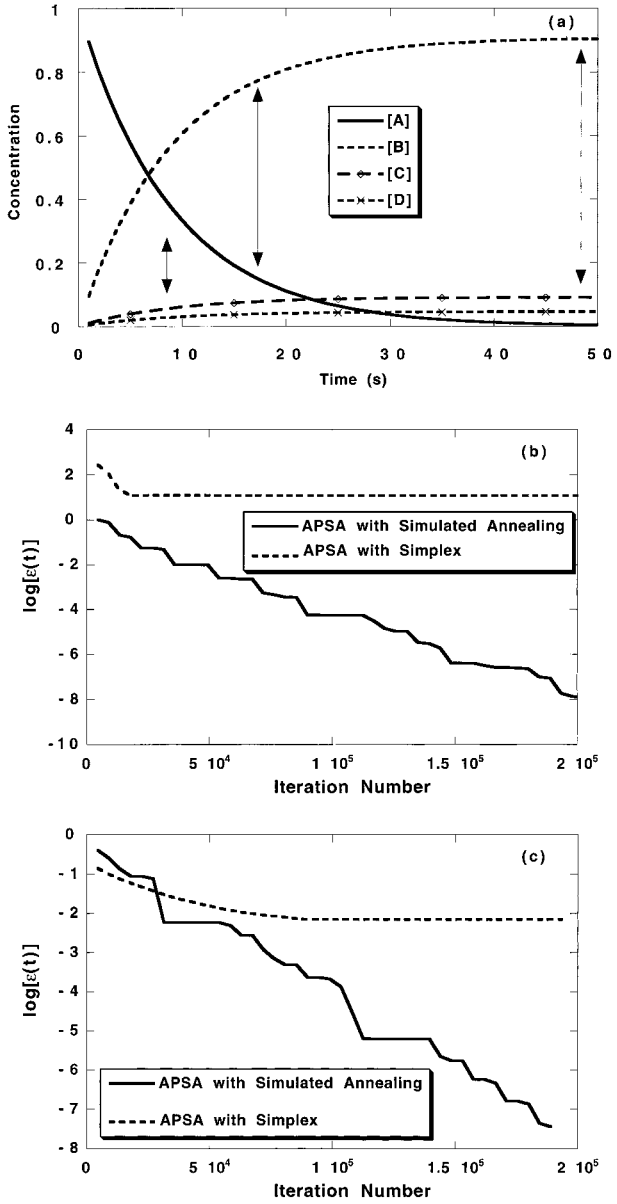

Figure 2. (a) Concentration versus time profile for the parallel reaction when $\nu = 1$ and $T = 300$ K. The solid, short-dashed, long-dashed (connecting diamonds), and short-dashed (connecting cross marks) lines show species A, B, C, and D, respectively. The arrows indicate-the target data. (b) Comparison of the convergence for the SA and Simplex-based APSA for the parallel reaction when $\nu = 1$. The solid and dashed lines show the convergence of the SA and Simplex-based algorithms, respectively. (c) Comparison of the convergence for the SA and Simplex-based APSA for the parallel reaction when $\nu = 2$. The solid and dashed lines show the convergence of the SA and Simplex-based algorithms, respectively.

The arrows indicate the target data points which we use as the known concentration versus time data. It should be pointed out that we only choose three time data points each for reactant A and final product C. Further, we assume we have no knowledge about the concentration of the intermediate species B. For many reactions, experimental concentration versus time profiles of intermediates are hard to measure, so we mimic this experimental situation by assuming we have no knowledge of the concentration versus time behavior of the intermediate B.

We now treat the Arrhenius prefactors, $\vec{\mathbf{A}} = \{A_1, A_{-1}, A_2\}$, and the activation energies, for each of the elementary steps, $\vec{\mathbf{E}}_A = \{E_{A_1}, E_{A_{-1}}, E_{A_2}\}$, as unknowns and use the APSA to find them from the limited concentration data. We assume no knowledge of the Arrhenius prefactors or activation energies and thus choose their initial values as zero. We stop the APSA after 2×10^5 iterations, which we have found to be quite sufficient in finding accurate values for the Arrhenius parameters. We show the results for the SA-based APSA in the third column in Table 1 for the consecutive reaction with the stoichiometric coefficient $\nu = 1$. Table 1 shows that all rate constants are reproduced by the search algorithm. The fourth column in Table 1 shows the results for an APSA that is based on a conventional minimization scheme, this being the Simplex algorithm. As shown by the results in column four in Table 1, a conventional minimizer is not able to find accurate values for the activation energies or Arrhenius prefactors. To further demonstrate the need for coupling the APSA to a heuristic minimization algorithm, we compare the convergence properties of the APSA when it is based on the SA and the Simplex minimization algorithm. In Figure 1b, we show the plot of the log of the error function versus the iteration number for the consecutive reaction when the stoichiometric coefficient $\nu = 1$. The solid line shows the error function for the SA-based APSA, while the dashed line shows the error function for the Simplex-based APSA. As shown by Figure lb, the SA-based algorithm converges very nicely toward an error of $\epsilon = 1 \times 10^{-7}$ after 2×10^5 iterations. The Simplex-based algorithm, on the other hand, does not converge toward smaller error and, in fact, gets stuck in a local minimum where the lowest value of the error is $\epsilon = 1 \times 10^{-1}$ after 2×10^5 iterations.

Table 1. Consecutive Reaction

parameters	exact	APSA $\nu = 1$	Simplex $\nu = 1$	APSA $\nu = 2$	Simplex $\nu = 2$
A_1, s^{-1}	100.00	100.08	40.43	100.19	72.50
A_{-1}, s^{-1}	500.00	500.14	18.62	501.99	146.85
A_2, s^{-1}	1000.00	998.42	0.04	999.18	317.95
$E_{A,1}$, kJ mol^{-1}	5.74	5.74	3.43	5.75	4.91
$E_{A,-1}$, kJ mol^{-1}	15.50	15.5	6.93	15.51	12.25
$E_{A,2}$, kJ mol^{-1}	28.72	28.71	2.67	28.71	25.66

We show the results for the SA-based APSA in the fifth column in Table 1 for the consecutive reaction with the stoichiometric coefficient $\nu = 2$. Again the SA-based APSA is able to reproduce all the Arrhenius parameters to great accuracy. The last column in Table 1 shows the results for the Simplex-based APSA for the stoichiometric coefficient $\nu = 2$. Again, these results indicate that the APSA that is based on a conventional minimizer is not able to find accurate values for the activation energies or Arrhenius prefactors.

In Figure 1c, we show the plot of the log of the error function versus the iteration number for the consecutive reaction when the stoichiometric coefficient $\nu = 2$. The solid line shows the error function for the SA-based APSA, while the dashed line shows the error function for the Simplex-based APSA. Again, Figure 1c shows that the SA-based algorithm converges very nicely toward an error of $\epsilon = 1 \times 10^{-7}$ after 2×10^5 iterations. Again, the Simplex-based algorithm does not converge toward smaller error and, in fact, gets stuck in a local minimum where the lowest value of the error is $\epsilon = 3$ after 2×10^5 iteratioris.

(ii) Parallel Reaction. For the parallel chemical reaction scheme that is outlined in eq 8, we synthesized a limited set of known concentration data for the reaction by choosing

Table 2. Parallel Reaction

parameters	exact	APSA $\nu = 1$	Simplex $\nu = 1$	APSA $\nu = 2$	Simplex $\nu = 2$
A_1, s^{-1}	100.00	99.92	59.08	100.01	72.50
A_2, s^{-1}	500.00	499.60	334.08	500.02	146.85
A_3, s^{-1}	1000.00	999.95	609.37	999.96	317.95
$E_{A,1}$, kJ mol^{-1}	5.74	5.74	4.33	5.74	4.91
E_{A2}, kJ mol^{-1}	15.50	15.49	14.42	15.50	12.25
E_{A3}, kJ mol^{-1}	28.72	28.72	27.38	28.71	25.66

the set of Arrhenius prefactors and the set of activation energies to have the same values as that for the consecutive reaction presented above. The initial concentrations of the species were chosen as $\{[A]_0 = 1.0$ M, $[B]_0 = 0$, $[C]_0 = 0$, $[D]_0 = 0\}$. As for the consecutive reaction, we studied this reaction for two values of the stoichiometric coefficient, $\nu = 1$ and 2, in eq 8. The concentration versus time profile for this reaction when $\nu = 1$ and $T = 300$ K is shown in Figure 2a. The arrows indicate the data points that we use as the known concentration versus time data.

Again, we treat the Arrhenius prefactors, $\vec{\mathbf{A}} = \{A_1, A_2, A_3\}$, and the activation energies, for each of the elementary steps, $\vec{\mathbf{E}}_A = \{E_{A_1}, E_{A_2}, E_{A_3}\}$, as unknowns and use the APSA to find them from the limited concentration data. We assume no knowledge of the Arrhenius prefactors or activation energies and thus choose their initial values as zero. We show the results for the SA-based APSA in the third column in Table 2 for the parallel reaction with the stoichiometric coefficient $\nu = 1$. Table 2 shows that all rate constants are reproduced by the search algorithm for the parallel reaction with $\nu = 1$. The fourth column in Table 2 shows the results for the Simplex-based APSA. Again, the conventional minimizer is not able to find accurate values for the activation energies or Arrhenius prefactors. In Figure 2b, we show the plot of the log of the error function versus the iteration number for the parallel reaction when the stoichiometric coefficient $\nu = 1$. The solid line shows the error function for the SA-based APSA, while the dashed line shows the error function for the Simplex-based APSA. As shown by Figure 2b, the SA-based algorithm converges very nicely toward an error of $\epsilon = 1 \times 10^{-8}$ after 2×10^5 iterations. The Simplex-based algorithm, on the other hand, does not converge toward smaller error and, in fact, gets stuck in a local minimum where the lowest value of the error is $\epsilon = 10$ after 2×10^5 iterations.

We show the results for the SA-based APSA in the fifth column in Table 2 for the parallel reaction with the stoichiometric coefficient $\nu = 2$. Again, the SA-based APSA is able to reproduce all the Arrhenius parameters to great accuracy. The last column in Table 2 shows the results for the Simplex-based APSA for the stoichiometric coefficient $\nu = 2$. Again, these results indicate that the APSA that is based on a conventional minimizer is not able to find accurate values for the activation energies or Arrhenius prefactors.

In Figure 2c, we show the plot of the log of the error function versus the iteration number for the consecutive reaction when the stoichiometric coefficient $\nu = 2$. The solid line shows the error function for the SA-based APSA, while the dashed line shows the error function for the Simplex-based APSA. Again, Figure 2c shows that the SA-based algorithm converges very nicely toward an error of $\epsilon = 1 \times 10^{-7}$ after 2×10^5 iterations. Again, the Simplex-based

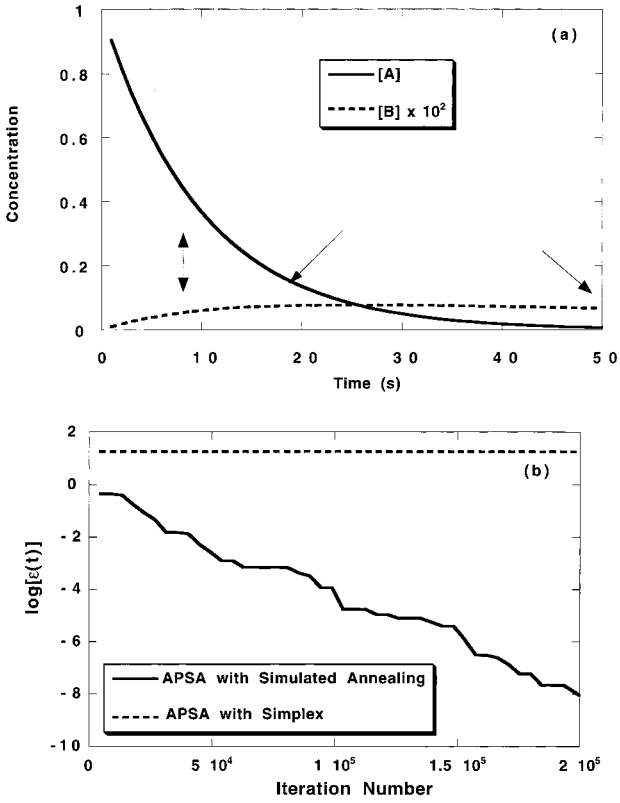

Figure 3. (a) Concentration versus time profile for the chemical fate reaction at $T = 300$ K. The solid and dashed lines show species A and B, respectively. The arrows indicate the target data. (b) Comparison of the convergence for the SA and Simplex-based APSA for the chemical fate reaction. The solid and dashed lines show the convergence of the SA and Simplex-based algorithms, respectively.

algorithm does not converge toward smaller error and, in fact, gets stuck in a local minimum where the lowest value of the error is $\epsilon = 1 \times 10^{-2}$ after 2×10^5 iterations.

(iii) Chemical Fate Reaction. For the chemical fate reaction scheme that is outlined in eqs 10a−c, we synthesized a limited set of known concentration data for this reaction scheme by choosing the set of Arrhenius prefactors and the set of activation energies to have the same values as that for the consecutive reaction presented above. The initial concentrations were $\{[A]_0 = 1.0$ M, $[B]_0 = 0\}$. The concentration versus time profile for this reaction at $T = 300$ K is shown in Figure 3a. The arrows indicate the data points that we use as the known concentration versus time data.

Again, we treat the Arrhenius prefactors, $\vec{\mathbf{A}} = \{A_1, A_2, A_3\}$, and the activation energies, for each of the elementary steps, $\vec{\mathbf{E}}_A = \{E_{A_1}, E_{A_2}, E_{A_3}\}$, as unknowns and use the APSA to find them from the limited concentration data. Again, we assume no knowledge of the Arrhenius prefactors or activation energies and choose their initial values as zero. We show results for the SA-based APSA in the third column in Table 3 for the chemical fate reaction. Table 3 shows that all rate constants are reproduced by the search algorithm for the chemical fate reaction. The fourth column in Table 3 shows the results for the Simplex-based APSA. Again, the conventional minimizer is not able to find accurate values for the activation energies or Arrhenius prefactors. In Figure 3b, we show the plot of the log of the error function versus iteration number for the chemical fate reaction. The solid line shows the error function for the SA-based APSA, while

Table 3. Chemical Fate Reaction

parameters	exact	APSA $\nu = 1$	Simplex $\nu = 1$
A_1, s^{-1}	100.00	99.98	488.85
A_2, s^{-1}	500.00	500.00	973.48
A_3, s^{-1}	1000.00	999.93	625.09
E_{A1}, kJ mol^{-1}	5.74	5.74	5.04
E_{A2}, kJ mol^{-1}	15.50	15.50	2.1×10^{-3}
E_{A3}, kJ mol^{-1}	28.72	28.71	1.0×10^{-3}

the dashed line shows the error function for the Simplex-based APSA. As shown by Figure 3b, the SA-based algorithm converges very nicely toward an error of $\epsilon = 1 \times 10^{-8}$ after 2×10^5 iterations. The Simplex-based algorithm, on the other hand, does not converge toward smaller error and, in fact, gets stuck in a local minimum where the lowest value of the error is $\epsilon = 17$ after 2×10^5 iterations.

IV. DISCUSSION AND CONCLUSION

The results presented above for the APSA lead to two major conclusions: (1) an APSA that is based on a Simulated Annealing minimization algorithm can be used to find accurate values of Arrhenius prefactors and activation energies from limited concentration data; (2) an APSA that is based on a conventional minimizer may **not** find accurate values of Arrhenius prefactors and activation energies from limited concentration data.

The chemical kinetics search algorithm that is based on a heuristic functional optimizer is not only efficient and accurate but also robust. In a previous work, we tested the robustness of a similar SA-based search algorithm by performing three diagnostic tests.[1] In the first test, we restarted the search algorithm with different initial values for the unknown parameters. The second test involved the stability of the search algorithm when there is an artificially induced error in the target concentration data. The third test involved using less concentration versus time data, in this case using only two concentration versus time data points, as opposed to three concentration versus time data points. These three diagnostic tests were run for the SA-based APSA on the three reactions presented in this paper. For each test, the search algorithm was shown to be stable and yield roughly the same values [although there is some degradation in accuracy, i.e., ~20%, when only two concentration versus time data points are used] for the Arrhenius prefactors and activation energies that were presented in Tables 1−3.

A chemical kinetics search algorithm that is based on a conventional functional optimizer, such as the Simplex algorithm, was shown to become stuck in a local minimum that was closest to the initial guesses. The values of the rate constants found from the Simplex-based APSA were found to be inaccurate as compared to the values obtained from the SA-based APSA.

REFERENCES AND NOTES

(1) Von Arx, K. B.; Huffman, S. W.; Manock, J. J.; Messina, M. Using limited concentration data for the determination of rate constants with the genetic algorithm. *Environ. Sci. Technol.* **1998**, *32*, 3207−3212.
(2) Terry, D. B.; Messina, M. Heuristic search algorithms for the determination of rate constants and reaction mechanisms from limited concentration data. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1232−1238.
(3) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. Optimization by simulated annealing. *Science* **1983**, *220*, 671−680.
(4) Kirkpatrick, S. Optimization by simulated annealing: Quantitative studies. *J. Stat. Phys.* **1984**, *34*, 975−986.
(5) Otten, R. H. J. M.; van Ginneken, L. P. P. P. *The Annealing Algorithm*, Kluwer: Boston, 1989.
(6) Vanderbilt, D.; Louie, S. G. A monte-carlo simulated annealing approach to optimization over continuous variables. *J. Comp. Phys.* **1984**, *56*, 259−271.
(7) Goldberg, D. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison-Wesley: New York, 1989.
(8) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran 77*; Cambridge University Press: Cambridge, MA, 1992.
(9) Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. *Chemical Kinetics and Dynamics*; Prentice Hall: Englewood Cliffs, NJ, 1989.

CI990332R