

Central limit theorem for chemical kinetics in complex systems

Joel Z. Bandstra and Paul G. Tratnyek*

*Department of Environmental and Biomolecular Systems, Oregon Health & Science University,
20000 NW Walker Road, Beaverton, OR 97006-8921, USA Tratnyek@ebs.ogi.edu*

Received 13 August 2004; revised 24 September 2004

The prevalence of apparently first-order kinetics of reactant disappearance in complex systems with many possible reaction pathways is usually attributed to the dominance of a single rate limiting step. Here, we investigate another possible explanation: that apparently first-order kinetics might arise because the aggregate behavior of many processes, with varying order of reaction and rate constant, approaches a “central limit” that is indistinguishable from first-order behavior. This hypothesis was investigated by simulating systems of increasing complexity and deriving relationships between the apparent reaction order of such systems and various measures of their complexity. Transformation of a chemical species by parallel irreversible reactions that are zero-, first-, or second-order is found to converge to a central limit as the number of parallel reactions becomes large. When all three reaction orders are represented, on average, in equal proportions, this central limit is experimentally indistinguishable from first-order. A measure of apparent reaction order was used to investigate the nature of the convergence both stochastically and by deriving theoretical limits. The range of systems that exhibit a central limit that is approximately first-order is found to be broad. First-order like behavior is also found to be favored when the distribution of material among the parallel processes (due to differences in rate constants for the individual reactions) is more complex. Our results show that a first-order central limit exists for the kinetics of chemical systems and that the variable controlling the convergence is the physical complexity of reaction systems.

KEY WORDS: reaction order, complexity, first-order kinetics

1. Introduction

A first-order rate equation is often used to model the chemical kinetics of reactions that involve complex pathways and multiple mechanisms. The results of such modeling are often quite satisfactory even though theoretical considerations would suggest a more complicated rate law. Common explanations of this

* Corresponding author.

phenomenon include: (i) the first-order model is applied over a small extent of reaction, in which case, many rate laws will adequately describe the data, (ii) the data contain noise causing any fine structure to be obscured and deviation from the first-order model to be attributed to random measurement errors rather than model failure, and (iii) the observed kinetics reflect a rate limiting step that is first-order. It is the purpose of this work to investigate another possible explanation: that multiple processes (which are not necessarily first-order) may combine to produce kinetic behavior that is indistinguishable from first-order and that such combinations are more likely to exist when reactions occur in a complex environment.

This hypothesis may be viewed as a chemical kinetics analog to the central limit theorem of statistics. The central limit theorem states that, when scaled properly, a linear combination of random variables approaches a normal distribution as the number of random variables becomes large [1]. Some familiar implications of the central limit theorem include (i) the justification for assuming that most measurement errors are normally distributed and (ii) the result that sufficiently large stochastic systems obey a Fokker–Planck equation [2] and, in the case of chemical kinetics, a reaction rate equation [3]. Our hypothesis, then, is that there is also a central limit effect in the reaction order of chemical kinetics where the central limit is first-order behavior and the controlling variable is system complexity.

In a general sense, it seems that the notion of this central limit effect is widely held, but we have found no evidence that it has been systematically investigated. Our interest in the hypothesis arose from consideration of contaminant degradation kinetics in complex environmental media such as soils and sediments, but we expect that this investigation will have relevance in other contexts such as chemical engineering of bioreactors, catalysis on heterogeneous surfaces, radical reactions in plasmas and flames, atmospheric chemistry, etc.

2. Methods

We have investigated the nature and validity of our conjecture by examining a range of specific kinetic systems for the type of behavior hypothesized. This approach involves four elements: (i) generation of a variety of rate laws, (ii) solution of the rate laws, (iii) assessment of the degree of complexity of the corresponding reaction system, and (iv) assessment of the apparent order of the resulting disappearance curve.

2.1. *Generation of rate laws*

We selected parallel irreversible reactions as our test case because it is the simplest possible system that contains the features employed in formulating our conjecture. Considering the simplest case allows us to clearly explain the details

of the conjecture and allows us to easily develop the necessary analyses and interpret the results. We generated rate laws for an arbitrary number of parallel irreversible reactions in the form:

$$\frac{dA}{dt} = - \sum_{i=1}^{N_{\text{rxns}}} k_i \cdot A^{m_i} \tag{1}$$

with the initial condition:

$$A(0) = A_0, \tag{2}$$

where A denotes concentration (or activity) of the reacting chemical species, N_{rxns} denotes the number of parallel reactions, k_i denotes the rate constant for the i th reaction, and m_i denotes the reaction order for the i th reaction.

A variety of rate laws may be obtained from equation (1) by employing various selection criteria for N_{rxns} , the values of k_i , and the values of m_i . We expect complexity of the system to increase with N_{rxns} , so we varied N_{rxns} from 1 to 50. We would like to interpret each reaction as an elementary reaction step, so we required that the value of each m_i be 0, 1, or 2 under the presumption that third-order and higher-order reaction steps are prohibitively unlikely. We selected the values of m_i stochastically when we wished to examine an ensemble of like reaction systems and combinatorially when we wished to examine all possible combinations of reaction orders for a given N_{rxns} . $P(m_i = x)$ denotes the probability that $m_i = x$ and n_x denotes the number of reactions with $m_i = x$. Note that $P(m_i = 0) + P(m_i = 1) + P(m_i = 2) = 1$ and $n_0 + n_1 + n_2 = N_{\text{rxns}}$.

We selected rate constants by first selecting a characteristic value, ξ_i , and then scaling ξ_i according to:

$$k_i = \begin{cases} \xi_i \cdot \frac{1}{N_{\text{rxns}} \cdot \langle \xi \rangle} \cdot A_0 \cdot \frac{7}{8 \cdot \ln(8)}, & m_i = 0, \\ \xi_i \cdot \frac{1}{N_{\text{rxns}} \cdot \langle \xi \rangle}, & m_i = 1, \\ \xi_i \cdot \frac{1}{N_{\text{rxns}} \cdot \langle \xi \rangle} \cdot \frac{1}{A_0} \cdot \frac{7}{\ln(8)}, & m_i = 2, \end{cases} \tag{3}$$

where $\langle \xi \rangle$ denotes the average value of the distribution from which each ξ_i is selected. Equation 3 contains three scaling factors: $1/(N_{\text{rxns}} \cdot \langle \xi \rangle)$, $A_0^{(1-m_i)}$, and a third reaction order dependent factor. The scaling factor, $1/(N_{\text{rxns}} \cdot \langle \xi \rangle)$, makes equation (1) independent of N_{rxns} and $\langle \xi \rangle$. This scaling is equivalent to scaling time by $N_{\text{rxns}} \cdot \langle \xi \rangle$. The scaling factor, $A_0^{(1-m_i)}$ (equals 1 for $m_i = 1$), causes A/A_0 to be independent of A_0 .

Rate constants for processes of different reaction order are dimensionally inconsistent and, therefore, cannot be compared directly. The third scaling factor is needed to force reactions to contribute to the overall system development in

proportion to their value of ξ_i , regardless of reaction order. We derived the values for this factor by requiring a zero-order only system and a second-order only system to proceed to an arbitrary concentration endpoint in the same amount of time as a first-order only system. The endpoint used was three half-lives, the selection of which is discussed later. In the case of stochastic selection of reaction orders, the ξ_i for each reaction was selected randomly from a normal distribution which was truncated at zero. In the case of combinatorial selection of reaction orders, all ξ_i 's were set equal to $\langle \xi \rangle$.

The manner of rate constant generation described above gives rise to groups of reaction systems with the same value of N_{rxns} , m_i selection criteria, and ξ_i selection criteria. In the discussion that follows, we will call these groups ensembles of like reaction systems. An ensemble member, then, is an individual reaction system and the ensemble size is the number of reaction systems belonging to a particular ensemble.

2.2. Solutions to the rate laws

Given the restrictions placed on m_i , an analytical solution to equation (1) may be obtained by direct integration. To aid integration, we re-write equation (1) as

$$\frac{dA}{dt} = -a \cdot A^2 - b \cdot A - c, \quad (4)$$

where

$$a = \sum_{\{i|m_i=2\}} k_i, \quad (5)$$

$$b = \sum_{\{i|m_i=1\}} k_i, \quad (6)$$

$$c = \sum_{\{i|m_i=0\}} k_i. \quad (7)$$

Integration and application of the initial condition (equation 2) yields:

$$A = \begin{cases} -\frac{b}{2a} + \frac{\sqrt{b^2-4ac}}{2a} \cdot \frac{(\sqrt{b^2-4ac+(b+2a \cdot A_0)})/(\sqrt{b^2-4ac-(b+2a \cdot A_0)}) \cdot e^{\sqrt{b^2-4ac}t-1}}{(\sqrt{b^2-4ac+(b+2a \cdot A_0)})/(\sqrt{b^2-4ac-(b+2a \cdot A_0)}) \cdot e^{\sqrt{b^2-4ac}t+1}}, & t < t_{A=0}, \\ 0, & t \geq t_{A=0}. \end{cases} \quad (8)$$

This solution, while general, contains a number of cases that result in singularities when calculating disappearance curves. In these cases the calculations were simplified by taking limits and/or employing Euler's formula. For example, if $a = c = 0$, then equation (8) takes on an indeterminate form and must be calculated

by taking the limit (using L'Hospital's rule) as a and c approach zero. The resulting equation, in this case, is the solution to a first-order rate law with rate constant equal to b .

2.3. Quantification of complexity

In selecting an approach to quantify complexity, we considered two generally accepted features of complex systems: (i) complex systems are comprised of many interrelated parts, and (ii) complex systems exhibit both ordered and random behavior [4–8]. A number of complexity metrics have been proposed [7,9–13], but application of these metrics to specific systems can be nontrivial and the interpretation of these metrics is still an area of research. Our preliminary efforts using the Shiner metric [13] produced results consistent with the findings reported below. However, we found that two lower level quantities, the number of reactions, N_{rxns} , and the pathway entropy (defined later), gave comparable results for the systems considered here. These surrogate complexity metrics proved to be preferable for present purposes due to ease of implementation and interpretation.

Complex systems are comprised of many interrelated parts. For the systems that we have considered, N_{rxns} is the relevant measure of the number of parts. Because we are not changing the way in which the reactions are related, it is intuitive to use N_{rxns} as a measure of complexity. This is consistent with the Shiner metric, which we found to be dominated by $\ln(N_{\text{rxns}})$ for the systems considered.

In addition to examining the effect of N_{rxns} , we also want to compare reaction systems with the same N_{rxns} . Early work in information theory by Shannon [14] led to widespread use of information or Shannon entropy as a measure of complexity

$$S = - \sum_i P_i \cdot \ln(P_i), \quad (9)$$

where S is a measure of the degree of randomness for the probability distribution, P . Because complex systems are neither completely ordered nor completely random, it has been recognized that the relationship between Shannon entropy and complexity includes a maximum at an intermediate entropy value [6,11,15].

For the purposes of this study, a Shannon entropy pertaining to chemical transformation may be calculated from the set of probabilities describing the likelihood that a randomly selected molecule will undergo a particular transformation, or follow a particular pathway, during the course of reaction. These pathway probabilities were calculated, as shown in equations (10)–(12), by dividing the limit of the concentration of the i th product as time approaches infinity by the total amount of material in the system (A_0).

$$P_i = \frac{k_i}{A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}} \cdot \ln \left(\frac{A_0 \cdot b + 2 \cdot c + A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}}{A_0 \cdot b + 2 \cdot c - A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}} \right), \quad m_i = 0, \quad (10)$$

$$P_i = \frac{k_i}{A_0 \cdot 2 \cdot a} \left(\ln \left(\frac{A_0 \cdot b + c + A_0^2 \cdot a}{c} \right) - \frac{b}{\sqrt{b^2 - 4 \cdot a \cdot c}} \cdot \ln \left(\frac{A_0 \cdot b + 2 \cdot c + A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}}{A_0 \cdot b + 2 \cdot c - A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}} \right) \right), \quad m_i = 1, \quad (11)$$

$$P_i = \frac{k_i}{A_0 \cdot 2 \cdot a^2} \left(A_0 \cdot 2 \cdot a - b \cdot \ln \left(\frac{A_0 \cdot b + c + A_0^2 \cdot a}{c} \right) + \frac{b^2 - 2 \cdot a \cdot c}{\sqrt{b^2 - 4 \cdot a \cdot c}} \cdot \ln \left(\frac{A_0 \cdot b + 2 \cdot c + A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}}{A_0 \cdot b + 2 \cdot c - A_0 \cdot \sqrt{b^2 - 4 \cdot a \cdot c}} \right) \right), \quad m_i = 2. \quad (12)$$

Application of equation (9) to the pathway probabilities (equations (10)–(12)) yields a pathway entropy that we denote as S_{pthwy} . For a fixed N_{rxns} , S_{pthwy} is a measure of the uniformity in the pathway probability distribution. A perfectly uniform pathway probability distribution yields the maximum S_{pthwy} , $\ln(N_{\text{rxns}})$. A distribution in which a pathway probability, P_i , equals one (and all others equal zero) yields $S_{\text{pthwy}} = 0$. In all cases considered here, S_{pthwy} is close to the maximum value, $\ln(N_{\text{rxns}})$, causing the relationship between complexity and S_{pthwy} to be approximately linear with negative slope. Since we are concerned only with relative complexity, we may use S_{pthwy} as our measure, noting that reaction systems with larger S_{pthwy} values are less complex.

2.4. Metric for apparent order of reaction

Many procedures have been described for determination of the apparent order of a reaction from experimental data, and their relative strengths and weaknesses are frequently discussed in textbooks of chemical kinetics. The most intuitive and widely used approach involves plotting transformed time series data and examining the plot for linearity. In this study, we start by using this approach to provide a qualitative but direct indication of reaction order. Since we are interested in examining for behavior that appears to be first-order, we used logarithmically transformed concentration versus time plots.

It also proved useful to have a quantitative measure of apparent order. To this end, we fit the solution to equation (13) to simulated data generated by evaluating equation (8) at 1000 time points, evenly spaced between $t = 0$ and the time required to reach the three half-lives concentration.

$$\frac{dA}{dt} = -k \cdot A^M. \quad (13)$$

Fitting was performed by Chi-square minimization using the Levenberg–Marquardt method [16] with both k and M as fitting parameters. M , then, is a quantitative measure of apparent reaction order.

We used three half-lives as the endpoint for these calculations, as well as for the normalization of rate constants (see above). This value was chosen to provide enough extent of reaction to reveal deviations from first order behavior on a log concentration versus time plot [17] while preserving enough resolution of the initial rate behavior to be consistent with the way many experimental protocols are performed.

3. Results

3.1. Convergence with increasing N_{rxns}

Ensembles of 50 similar rate laws were generated for $N_{\text{rxns}} = 1, 5, 20,$ and 50 with a uniform distribution on the reaction order ($P(m_i = 0) = P(m_i = 1) = P(m_i = 2) = 1/3$) and a normal distribution ($\mu = 1, \sigma = 0.3$) on the raw rate parameter, ξ . The resulting disappearance behavior is depicted in figure 1 on a semi-log concentration versus time plot. Recall that (in this type of plot) purely second-order behavior is a curve with positive concavity, purely first-order behavior is a straight line, and purely zero-order behavior is a curve with negative concavity.

The $N_{\text{rxns}} = 1$ ensemble (figure 1(a)) exhibits all three integer-order possibilities in approximately equal proportions. With increasing N_{rxns} (figures 1(b)–(d)), the likelihood of integer order behavior becomes smaller and the ensemble members tend toward a limiting case (indicated by the dashed line in figure 1) that is close to first-order. Behavior that appears to be first-order is possible for any N_{rxns} while behavior that appears zero-order or second-order becomes increasingly improbable as N_{rxns} increases.

The limiting case, shown as a dashed line in figure 1, is described by a rate law of the form of equation (4). As N_{rxns} becomes large, equations (5–7) become infinite series which converge to the following:

$$a_{\text{lim}} = \lim_{N_{\text{rxns}} \rightarrow \infty} a = P(m_i = 2) \cdot \frac{1}{A_0} \cdot \frac{7}{\ln(8)}, \tag{14}$$

$$b_{\text{lim}} = \lim_{N_{\text{rxns}} \rightarrow \infty} b = P(m_i = 1), \tag{15}$$

$$c_{\text{lim}} = \lim_{N_{\text{rxns}} \rightarrow \infty} c = P(m_i = 0) \cdot A_0 \cdot \frac{7}{8 \cdot \ln(8)}. \tag{16}$$

On a semi-log plot, the limit exhibits a slight “S” shape; concave up at high concentrations and concave down at low concentrations. The reason for this curvature and its magnitude will be examined later.

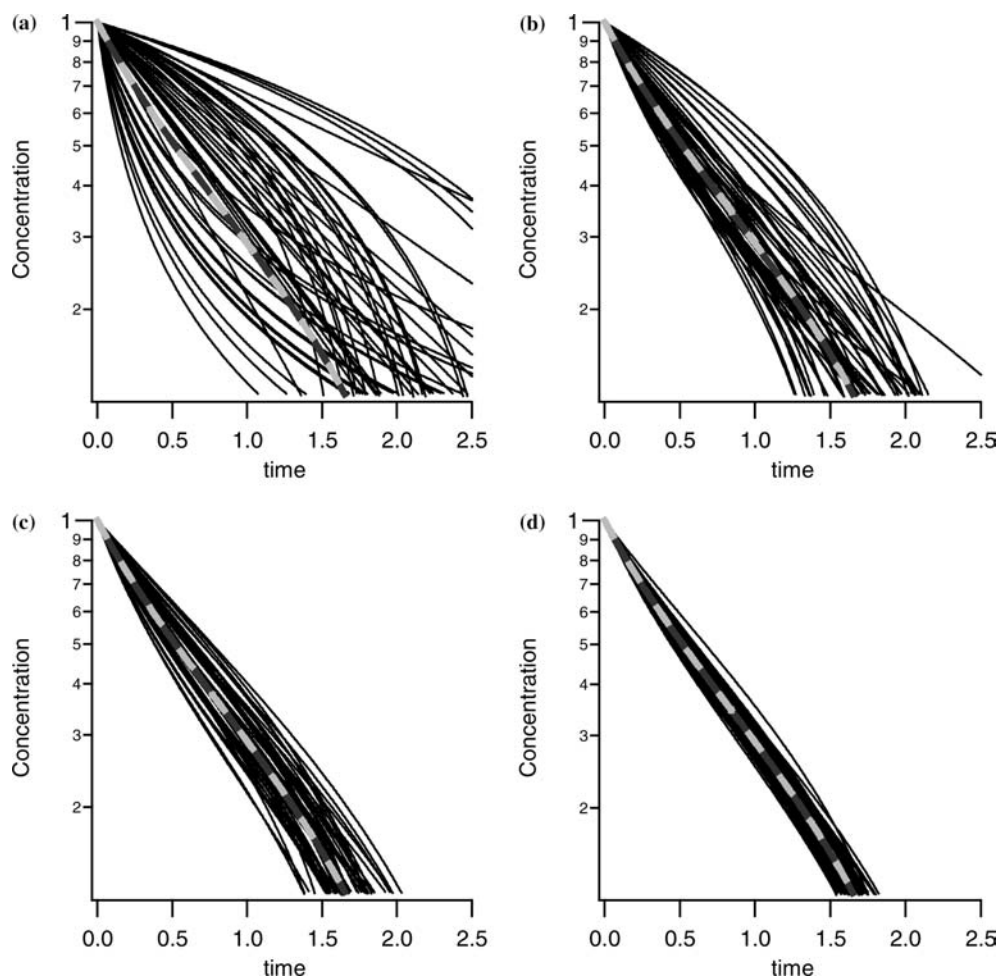


Figure 1. Disappearance curves for ensembles (ensemble size = 50) with $N_{\text{rxns}} =$ (a) 1, (b) 5, (c) 20, and (d) 50. Rate constants and reaction orders selected stochastically. The limiting disappearance curve for $N_{\text{rxns}} \rightarrow \infty$ is marked by the dashed line in each plot.

The behavior exhibited in figure 1 may be further examined by plotting the apparent reaction order, M , (as defined by equation (13)) against N_{rxns} . This is shown in figure 2 for ensembles of reactions generated in the same manner as those shown in figure 1 but with the ensemble size increased to 100. Each point represents a particular M for a member of an ensemble with the number of reactions indicated on the abscissa. As is the case for $N_{\text{rxns}} = 1$, some points may lie exactly on top of each other even if the corresponding disappearance curves do not. The average order and the average \pm one standard deviation for each N_{rxns} are plotted as solid lines. As N_{rxns} increases, the average order approaches a limiting value and the standard deviation approaches zero. Although the limiting

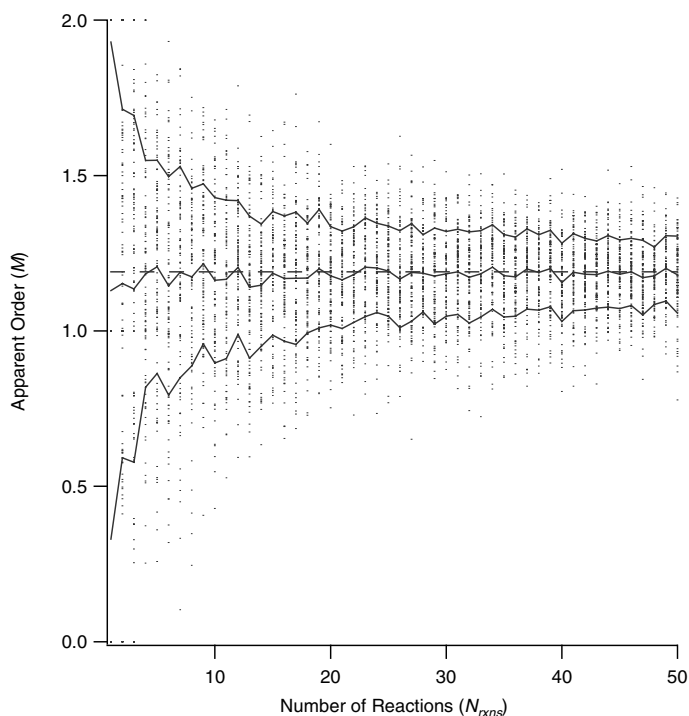


Figure 2. Apparent order calculated for statistically generated ensembles (ensemble size = 100) with N_{rxns} varying from 1 to 50. Solid lines denote the average order and standard deviation about the average for each ensemble. The limiting apparent order is shown by the dashed line.

apparent order (= 1.19) is somewhat greater than one, a first-order model may be fit to such a limiting disappearance curve. We calculated the average value of the absolute residuals for such a fit to be $0.01A_0$ concentration units. This is smaller than the random errors in most experimental studies.

To verify the interpretation given to figure 2, we can remove the distribution on ξ and examine every possible combination of reaction orders for a given N_{rxns} . A combination of reaction orders is defined by the sequence, (n_0, n_1, n_2) , where n_x is the number of parallel reactions of order x . All possible combinations corresponding to a single N_{rxns} may be found by taking all permutations of (n_0, n_1, n_2) under the condition $n_0 + n_1 + n_2 = N_{\text{rxns}}$. The number of permutations is given by the combinatorial expression.

$$C^R(3, N_{\text{RxnS}}) = \frac{(1 + N_{\text{RxnS}}) \cdot (2 + N_{\text{RxnS}})}{2}. \quad (17)$$

Here, all permutations are not equally probable. The weighting of (n_0, n_1, n_2) is given by

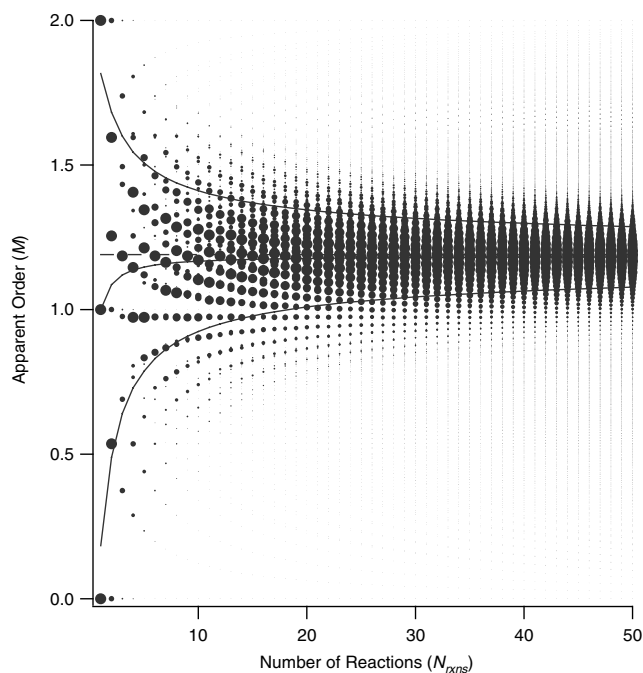


Figure 3. Apparent order of all possible combinations of zero, first, and second-order reactions for N_{rxns} from 1 to 50. The point size indicates the weight of the corresponding permutation relative to the other permutations for the same N_{rxns} . Average and standard deviation curves are shown (as in Figure 2) and represent the theoretical limit of such lines for infinitely large, statistically generated ensembles.

$$p(n_0, n_1, n_2) = [p(m_i = 0)]^{n_0} \cdot [p(m_i = 1)]^{n_1} \cdot [p(m_i = 2)]^{n_2} \cdot \frac{N_{\text{rxns}}!}{n_0! \cdot n_1! \cdot n_2!}. \quad (18)$$

Figure 3 shows the results of this analysis, in the same format as figure 2 (M versus N_{rxns}). Larger points reflect greater relative weight (as defined by equation (18)). As N_{rxns} increases, the number of permutations increases and the probability density becomes increasingly localized in the region of the limiting order. As N_{rxns} approaches infinity, the number of permutations approaches infinity and the sum of the weights for permutations with apparent order equal to 1.19 (the limiting order) approaches one. The average apparent order and the average \pm one standard deviation curves plotted in figure 3 represent the theoretical limit (as the ensemble size approaches infinity) of the behavior that is approximated in figure 2.

3.2. Behavior of the limiting case

Equations (14)–(16) and (18) indicate that the limiting apparent order (dashed line in figures 2 and 3) is a function of the reaction order probability

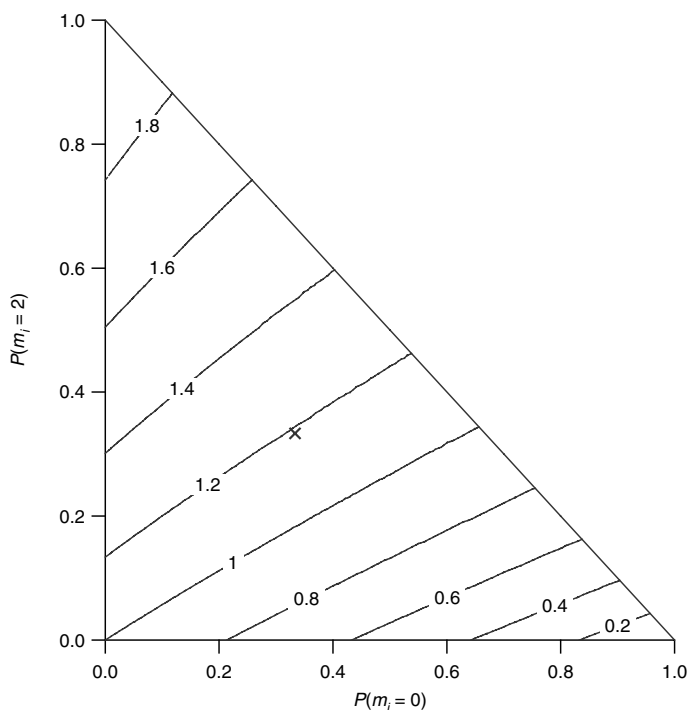


Figure 4. Contours of the limiting apparent order in the space of the probabilities used in selecting reaction orders. Purely second-order behavior is represented by point (0,1), purely first-order behavior by point (0,0), and purely zeroth-order behavior by point (1,0). The case considered in Figures 1–3 is marked with an “x”.

distribution. We have represented this functional dependence in figure 4 by plotting contours of constant limiting apparent order in the space of zero-order and second-order probabilities. In this figure, the first-order probability is implied by the values of the zero-order and second-order probabilities. Each corner of the triangle represents a case where all reactions have the same order. The center of the triangle (indicated by an “x”) is the case that we considered in figures 1–3. We note that the area between the contours for $M=0.8$ and $M=1.2$ is much larger than either of the areas above the $M=1.8$ contour or below the $M=0.2$ contour. This shows that many reaction order distributions will ultimately yield apparent behavior that is near to first-order.

The fitting routine used to generate figure 4 produces fits with chi-square equal to zero only for the reaction systems represented by the corners of figure 4. This is due to the tendency of the log-transformed disappearance curves to exhibit an “S” shape as noted in the discussion of figure 1. The contribution to the concavity of the log-transformed disappearance curve due to second-order

processes is positive and increases with concentration. First-order processes do not contribute to the concavity of the log-transformed disappearance curve and zero-order contributions are negative and increases with concentration (decreases in magnitude). As a result, the concavity of the log-transformed limiting case is initially dominated by second-order reactions and, therefore, positive, but finally dominated by zero-order reactions and, therefore, negative.

The log-transformed concavity is most pronounced when few of the processes are first-order (i.e. the $P(m_i = 0) + P(m_i = 2) = 1$ line in figure 4). For the worst case along the $M = 1$ contour, the average of the absolute residuals between the limiting disappearance curve and the first-order fit is equal to $0.002A_0$ concentration units. This value is significantly smaller than the error in a typical chemical kinetics experiment. Analysis of other cases along the $M = 1$ contour (not shown) indicates that this value decreases to zero at $P(m_i = 1) = 1$ slightly slower than linearly.

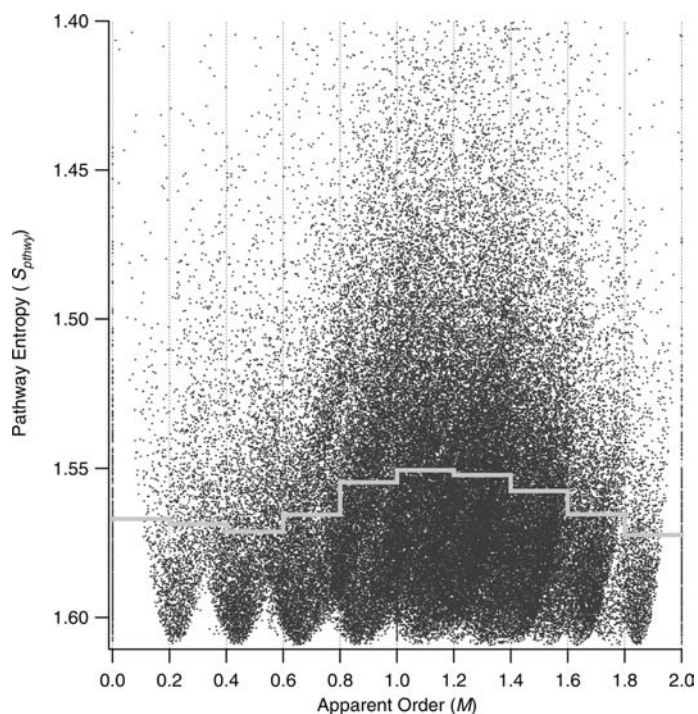


Figure 5. Scatter plot of pathway entropy (S_{pthwy}) versus apparent order (M) for an ensemble of 10,000 statistically generated parallel reactions with $N_{\text{rxns}} = 5$. Smaller S_{pthwy} (smaller values towards the top) indicates larger complexity. The solid line denotes the average S_{pthwy} for all ensemble members in the noted order range.

3.3. Convergence with decreasing S_{pthwy}

In the preceding discussion we have treated N_{rxns} as the variable controlling the tendency toward a central limit. It may also be instructive to look for centralizing tendencies within an ensemble of fixed N_{rxns} . To this end, we have calculated S_{pthwy} according to equations (9)–(12) for each member of an $N_{\text{rxns}} = 5$ ensemble generated in the same manner as those depicted in figures 1 and 2, but with the ensemble size increased to 10,000. Figure 5 shows the results of this calculation plotted against the apparent order, M . It is visually apparent from the distribution of the clusters of points that the average pathway entropy of ensemble members at or near the limiting order (for these conditions, 1.19) is larger than S_{pthwy} for the ensemble members on either side. This observation is reinforced by the maximum in the arithmetic mean of S_{pthwy} (represented in figure 5 by the white line), which was calculated for the ranges of order indicated with vertical dashed lines in the figure.

The ensemble members with an approximately equal number of each reaction type tend to yield apparent order close to one. Such ensemble members also have, on average, smaller pathway entropy. This is because, when multiple reaction types are present, the pathways with lower reaction order tend to convert the most material over the full course of the reaction, yielding a less ordered distribution of material amongst the individual reactions. Somewhat less ordered pathway probability distributions may arise for any combination of reaction orders (due to the influence of distributed rate constants), but this is more likely for systems with an equal number of each reaction type (due to the additional influence of distributed reaction orders). As complexity increases, therefore, the likelihood that a system will exhibit first-order behavior also increases even if the system does not have a large N_{rxns} . Convergence toward a central limit with two independent measures of complexity, N_{rxns} and S_{pthwy} , indicates that this result may apply to more complicated reaction systems that involve, for example, reversible steps and/or multiple reactants.

Acknowledgments

This research was supported primarily by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SER-DP). Additional support was provided by the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program. The work presented here has not been reviewed by either agency, so no official endorsement of the results or conclusions should be inferred.

References

- [1] A.M. Mood, T.A. Graybill and D.C. Boes, *Introduction to the Theory of Statistics*, 3rd ed. (McGraw-Hill, New York, 1974).
- [2] L.E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, TX, 1980).
- [3] D.T. Gillespie, *J. Phys. Chem. A* 106 (2002) 5063.
- [4] H.A. Simon, *Proc. Am. Philos. Soc.* 106 (1962) 467.
- [5] L. Poon and C. Grebogi, *Phys. Rev. Lett.* 75 (1995) 4023.
- [6] P. Grassberger, *Int. J. Theor. Phys.* 25 (1986) 907.
- [7] K. Lindgren and M.G. Nordahl, *Complex Syst.* 2 (1988) 409.
- [8] D.P. Feldman and J.P. Crutchfield, *Phys. Lett. A* 238 (1998) 244.
- [9] P. Szépfalussy and G. Györgyi, *Phys. Rev. A* 33 (1986) 2852.
- [10] J.P. Crutchfield and K. Young, *Phys. Rev. Lett.* 63 (1989) 105.
- [11] J.P. Crutchfield, *Physica D* 75 (1994) 11.
- [12] C.R. Shalizi and J. P. Crutchfield, *J. Stat. Phys.* 104 (2001) 817.
- [13] J.S. Shiner, M. Davison and P.T. Landsberg, *Phys. Rev. E* 59 (1999) 1459.
- [14] C.E. Shannon, *Bell Syst. Tech. J.* 27 (1948) 379.
- [15] W. Li, *Complex Syst.* 5 (1991) 381.
- [16] W.H. Press, B.P. Flannery and S.A. Teukolsky et al., *Numerical Recipes in C. The Art of Scientific Computing* (Cambridge University, Cambridge, England, 1988).
- [17] E.T. Urbansky, *J. Chem. Educ.* 78 (2001) 921.