Inferring Lifetime Distributions from Kinetics by Maximizing Entropy Using a Bootstrapped Model

Peter J. Steinbach*

Center for Molecular Modeling, Center for Information Technology, National Institutes of Health, Building 12A Room 2051, 12 South Drive, Bethesda, Maryland 20892-5624

Received June 6, 2002

A bootstrapped model is used to improve the lifetime distribution recovered using the maximum entropy method from kinetics that involves overlapping exponential and distributed phases. The model defaulted to in the limit of low signal-to-noise is iteratively derived from the data to counter the tendency of regularization methods to over-smooth sharp features while under-smoothing broad ones. Upon each revision, some of the lifetime distribution is focused and the rest is blurred. This differential blurring can produce distributions that are virtually free of artifacts. The change in the result obtained upon a reasonable change in the default model provides a useful measure of the uncertainty in the lifetime distribution. In particular, the widths of peaks may not be well determined.

INTRODUCTION

The recovery of lifetimes from kinetics is an important step in many biophysical investigations. This problem is known to be ill-posed; many lifetime distributions will fit noisy data equally well. In an attempt to include only the essential kinetic features in a unique solution, regularization methods optimize a desirable trait of the lifetime distribution at a given quality of fit. For example, the maximum entropy method (MEM),[1] as commonly used, chooses the most uniform distribution. The program CONTIN[2] favors the distribution with the smallest integrated curvature.

Although each of these regularizers often performs satisfactorily, neither works well in all cases.[3] Many physically reasonable lifetime distributions change discontinuously in some places (delta functions corresponding to exponential kinetics) and gradually in others. Such distributions possess neither maximum entropy nor minimal integrated curvature. Moreover, they will not be recovered optimally by any traditional method that regularizes all regions of the lifetime distribution equivalently. The distribution obtained from such data, using either the MEM with a uniform default distribution or CONTIN, may exhibit excessively broadened sharp features and small ripples superimposed on broader features.[3]

Recently, the ‘differential blurring’ of an intermediate MEM result was introduced in the program MemExp[4] as a way to derive the model distribution from kinetics that involves one or more quasi-exponential phases that are well separated from distributed phases. Here, the bootstrapping of the default model is extended to analyze data with coincident exponential and distributed phases. This functionality is available online in MemExp, version 2.

COMPUTATIONAL METHODS

MemExp can be used to interpret kinetics that possess rising and decaying phases and a slowly varying baseline in terms of lifetime distributions and discrete exponentials.[4] For simplicity, the current analysis is restricted to data that decrease monotonically to zero. Assume the kinetics is recorded at N time values t_i and that the uncertainties can be approximated as normally distributed. The fit $\mathcal{F}_i$ to datum D_i can be written as

$$\mathcal{F}_i = D_0 \int_{-\infty}^{\infty} d\log\tau f(\log\tau) e^{-t_i/\tau} \qquad (1)$$

where D_0 is a normalization constant. In the discretized form of eq 1, the lifetime distribution is represented as a vector **f**. From all distributions **f** that produce fits with a given value of chi-square

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\mathcal{F}_i - D_i}{\sigma_i}\right)^2 \qquad (2)$$

the MEM selects the distribution possessing maximum entropy S:[5]

$$S(\boldsymbol{f},\boldsymbol{F}) = \sum_{j=1}^{M}[f_j - F_j - f_j \ln(f_j/F_j)] \qquad (3)$$

Maximizing S with respect to f_j yields $f_j = F_j$. Thus, **F** is the model distribution that is defaulted to in regions of low signal-to-noise. Here, data were simulated with uniform standard errors, $\sigma_i = 0.001$. However, all results were obtained using the errors estimated with MemExp.[4] A traditional MEM calculation sets all F_j equal and maximizes S while constraining χ^2 and, optionally, the normalization of **f**. The uniform **F** is conceptually appealing given its simplicity and its apparent supposition that maxima resolved in the lifetime distribution should be demanded by the data, not assumed a priori. However, distributions obtained with a uniform model in test cases can contain artifactual ripples, as when recovering an exponential close in time to a

* Corresponding author phone: (301)496-1100; e-mail: steinbac@helix.nih.gov.

10.1021/ci025551i Published 2002 by the American Chemical Society
Published on Web 10/15/2002

distributed process. Similar discrepancies have been reported with CONTIN.[3]

Improved results can be obtained by adapting the model **F** to features recovered in **f**. Often, small wiggles can be suppressed by blurring **f** uniformly[6] to redefine the model. That is, **f** is convoluted with a Gaussian, and the calculation is resumed with this blurred **f** used as **F**. However, when sharp (e.g., exponential) and broad features are proximal in time, the differential blurring[4] of **f** may be needed to avoid excessive broadening of sharp features and under-smoothing of slowly varying features. When **F** is revised by differential blurring, sharp features identified in **f** are made even sharper before being added to **F** and all remaining features are first blurred by convolution. This peak-dependent manipulation is subjective, but it can dramatically suppress artifacts that result when using a uniform model.[4]

The differential blurring of **f** into **F** was improved in MemExp to deal with overlapping sharp and broad features as follows. The area A_i and mean M_i of the ith peak resolved in **f** are approximated by numerical integration. When **F** is to be revised by differential blurring, all relatively small and/or broad features in **f** are incorporated into **F** after being blurred by convolution with a Gaussian. Any peak in **f** that is sufficiently large and well resolved is accounted for in **F** in one of three ways. For each such sharp peak, a fraction of A_i is incorporated into **F** as a Gaussian having a full width at half-maximum (fwhm) that is reduced from that of the MEM peak. For a well-separated peak, all of A_i is represented in **F** as a narrowed Gaussian having area A_i and mean M_i.[4]

For relatively sharp peaks in **f** that are not well separated from neighboring features, the user specifies which of two methods is to partition A_i into the fraction to be sharpened and the fraction to be blurred. The first approach interprets **f** in the vicinity of the ith peak as representing a sharp peak of relatively large area that only slightly overlaps a broad feature. The second corresponds to a smaller sharp peak that sits atop the broad feature. If interpreted as slightly overlapping (Figure 1), the sharp peak is folded over from its most resolved side, and the symmetric peak obtained is represented in **F** as a narrowed Gaussian while the remainder of the ith peak is added to **F** after convolution with a Gaussian.

If interpreted as significantly overlapping (Figures 2 and 3), the sharp and broad features are separated by a line. This baseline is subtracted from the **f** values above it to obtain the area attributed to the sharp feature. The line is initially determined by the points on either side of the local maximum where the second derivative of **f** is positive and the first derivative is minimal in absolute value. The user can increase the intercept of the line to reduce the area to be sharpened and increase the area to be blurred. This upward shift of the baseline counters the suppression of shoulders in **f** that can result from the MEM's bias favoring a uniform distribution. The portion of the distribution above the shifted line is sharpened; the density below is blurred.

RESULTS AND DISCUSSION

Building on previous work,[3,4] the differential blurring of **f** into **F** was tested on kinetics simulated as the sum of a distributed process and an exponential of amplitude 1/6 located at $\log(\tau/\mathrm{s}) = -5$ (Figure 1), -4 (Figure 2), and -3 (Figure 3). Each of the three MEM calculations was started

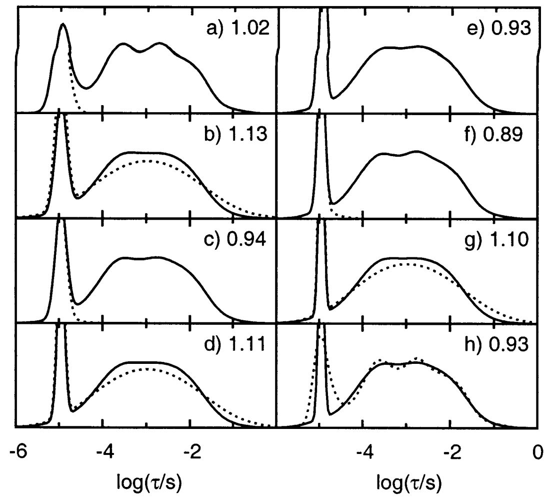

Figure 1. MemExp analysis of kinetics simulated from the sum of a distributed process and an exponential at $\log(\tau/\mathrm{s}) = -5$. Convergence proceeds from (a) to (h); each panel shows the current distribution **f** (solid) and value of χ^2. Upon each differential blurring of **f** (panels a, c, and f), the symmetric peak obtained by folding over is drawn (dotted), and the resultant **F** is plotted in the following panel (dotted; panels b, d, and g). The distribution recommended by MemExp is shown in (h), along with the distribution (dotted) obtained at $\chi^2 = 0.95$ with a uniform **F**.

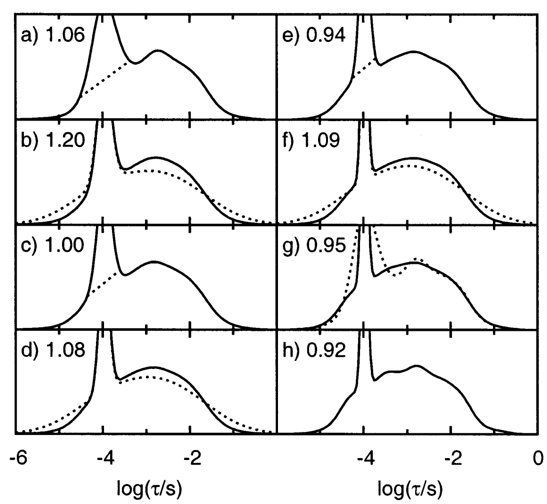

Figure 2. MemExp analysis of kinetics simulated from the sum of a distributed process and an exponential at $\log(\tau/\mathrm{s}) = -4$. Convergence proceeds from (a) to (h); each panel shows the current distribution **f** (solid) and value of χ^2. Upon each differential blurring of **f** (panels a, c, and e), the shifted linear baseline is drawn (dotted), and the resultant **F** is plotted in the following panel (dotted; panels b, d, and f). The distribution recommended by MemExp is shown in (g), along with the distribution (dotted) obtained at $\chi^2 = 0.96$ with a uniform **F**.

with a uniform **F** and continued until χ^2 reached 1.1, where **f** was differentially blurred for the first time (Figures 1a, 2a, 3a). In each case, the focused portion of **f** was represented in **F** by a Gaussian with a fwhm multiplied by 0.7, and the remainder of **f** was convoluted with a Gaussian having a fwhm of 1.5. In Figure 1 the tall peak was folded over to determine the portion of **f** to be focused; in Figures 2 and 3, the linear baseline originally determined (as described above) was shifted up so that 85% of the area above it was focused. The differential blurring was repeated following each subsequent decrease in χ^2 by 5% until convergence. For each calculation, the lifetime distribution recommended by MemExp (Figures 1h, 2g, 3g) was the last one stored after the

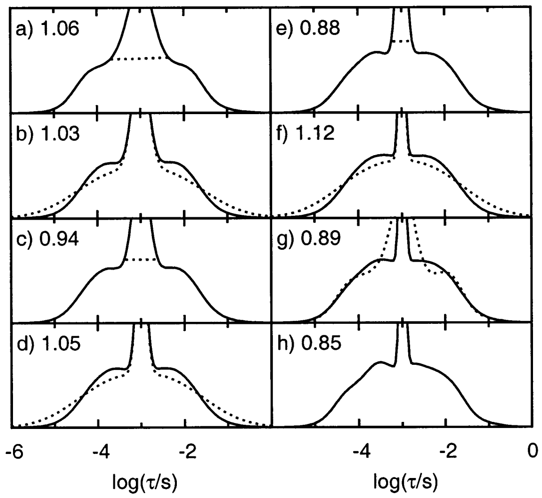

Figure 3. MemExp analysis of kinetics simulated from the sum of a distributed process and an exponential at log (τ/s) = −3, plotted as in Figure 2. The distribution recommended by MemExp is shown in (g), along with the distribution (dotted) obtained at $\chi^2 = 0.88$ with a uniform **F**.

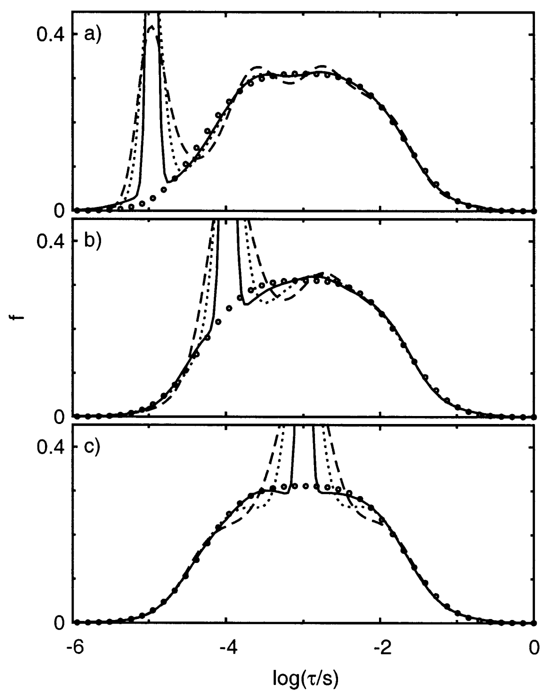

Figure 4. Convergence of lifetime distributions obtained from kinetics simulated as the sum of a distributed process (circles) and an exponential of amplitude 1/6 at log (τ/s) = −5 (a), −4 (b), and − 3 (c). Plotted in each panel are the first distribution **f** to be differentially blurred (dashed), an intermediate iteration (dotted), and the distribution recommended by MemExp following the final update to **F** (solid).

final revision to **F** for which both χ^2 and the correlation length of the residuals had decreased by more than 1% from the previously stored distribution.[4]

Figure 4 summarizes the convergence of the MemExp calculations and shows the true distributed process (circles) for comparison. For each of the three simulated data sets, the iterative bootstrapping of **F** by differential blurring produced a distribution closer to the correct distribution than was obtained at comparable values of χ^2 using a uniform **F** (Figures 1h, 2g, 3g). Use of an adaptive model suppressed the artifacts that result from the use of a uniform **F**, the latter representing a more traditional regularization in which all regions of the lifetime distribution are biased equivalently. Moreover, the change in the lifetime distribution observed upon a reasonable change in the model **F** is a good measure of the uncertainty in the result. Clearly, each of these data sets can be fit well using a range of widths for the peak corresponding to the simulated exponential phase.

Because the recovery of **f** from noisy kinetics is ill-posed, it is difficult to avoid user interaction in deciding how to partition the area under the distribution among overlapping features. Ambiguity in interpretation is fundamental to ill-posed problems, and user input is often invoked in other signal processing applications, such as image reconstruction.

Instead of seeking a unique solution to an ill-posed problem, it is informative to explore the range of reasonable solutions by interpreting kinetics from multiple perspectives. The ability to change **F**, and hence the bias imposed by maximizing *S*, should be exploited. MemExp allows the use of uniform default models and models adapted to a particular data set by the uniform or differential blurring of **f**. In addition, the MEM calculation performed by MemExp can be used to guide a series of fits by discrete exponentials.[4] This diversity of approaches, developed for the analysis of kinetics data, could expedite the interpretation of other types of experiments as well.

CONCLUSIONS

When seeking the minimally structured lifetime distribution, it is important to analyze kinetics with different default models to determine which features are required by the data. Small undulations in the lifetime distribution may reflect true kinetic phases or simply a negative side effect of regularization. When a model obtained by differential blurring leads to fewer local maxima or diminished ripples in the lifetime distribution at a given value of χ^2, progress has been made in identifying the necessary kinetic features. Also, fine-tuning of the model is needed to assess the variability in the lifetime distribution afforded by the signal-to-noise of the data. In particular, this analysis demonstrates the considerable uncertainty in the widths of peaks in lifetime distributions.

ACKNOWLEDGMENT

I thank Charles Schwieters for his comments on the manuscript.

REFERENCES AND NOTES

(1) Livesey, A. K.; Brochon, J. C. Analyzing the distribution of decay constants in pulse-fluorimetry using the maximum entropy method. *Biophys. J.* **1987**, *52*, 693−706.
(2) Provencher, S. W. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comput. Phys. Comm.* **1982**, *27*, 213−227.
(3) Provencher, S. W. In *Laser Light Scattering in Biochemistry*; Harding, S. E.; Sattelle, D. B.; Bloomfield, V. A., Eds.; The Royal Society of Chemistry: Cambridge, U.K., 1992; pp 92−111.
(4) Steinbach, P. J.; Ionescu, R.; Matthews, C. R. Analysis of kinetics using a hybrid maximum-entropy/nonlinear-least-squares method: application to protein folding. *Biophys. J.* **2002**, *82*, 2244−2255.
(5) Skilling, J. In *Maximum Entropy and Bayesian Methods*; Skilling, J., Ed.; Kluwer Academic: Norwell, MA, 1989; pp 45−52.
(6) Gull, S. F. In *Maximum Entropy and Bayesian Methods*; Skilling, J., Ed.; Kluwer Academic: Norwell, MA, 1989; pp 53−71.

CI025551I