



Published in final edited form as:

Anal Biochem. 2012 August 1; 427(1): 102–105. doi:10.1016/j.ab.2012.04.008.

Filtering artifacts from lifetime distributions when maximizing entropy using a bootstrapped model

Peter J. Steinbach

Center for Molecular Modeling, Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA

Peter J. Steinbach: steinbac@mail.nih.gov

Abstract

The maximum entropy method (MEM) has been used in many studies to reliably recover effective lifetimes from kinetics, whether measured experimentally or simulated computationally. Here, recent claims made by Mulligan et al. regarding MEM analyses of kinetics (*Anal. Biochem.* 421 (2012) 181–190) are shown to be unfounded. Their assertion that their software allows “analysis of datasets too noisy to process by existing iterative search algorithms” is refuted with a MEM analysis of their triexponential test case with increased noise. In addition, it is shown that lifetime distributions recovered from noisy kinetics data with the MEM can be improved by using a simple filter when bootstrapping the prior model. When deriving the bootstrapped model from the lifetime distribution obtained using a uniform model, only the slower processes are represented as Gaussians in the bootstrapped model. Using this new approach, results are clearly superior to those of Mulligan et al. despite the presence of increased noise. In a second example, ambiguity in the interpretation of Poisson kinetics in the presence of scattered excitation light is resolved by filtering the prior model.

Keywords

Kinetics; Inverse Laplace transform; Rate constants; Lifetime distributions; Maximum entropy method

Rooted in information theory, the maximum entropy method (MEM)¹ has been used to interpret a variety of noisy measurements for many years. Since its early applications to astronomical images [1,2], the method has proven to be valuable in biophysical studies, as when interpreting time-dependent signals in terms of a distribution of effective lifetimes [3–8]. The MEM has also been used to infer a two-dimensional distribution of activation enthalpy and entropy from ligand rebinding kinetics monitored at low temperatures [9]. The program MemExp [7,8] can be used to analyze kinetics with phases of opposite sign and a slowly varying baseline in terms of lifetime distributions and discrete exponentials. Experiments (e.g., see Ref. [10,11]) and simulations [12,13] have been interpreted using MemExp.

Here, a simple way of bootstrapping the prior model that defines the entropy of the lifetime distribution is shown to suppress unwarranted features. The approach is applied first to a test case proposed recently involving exponentials of opposite sign and then to simulated Poisson kinetics convolved with a known instrument response function. For simplicity, we

¹Abbreviation used: MEM, maximum entropy method.

assume that neither analysis requires a baseline correction and that the Poisson data involve kinetics that decay monotonically. The fit \mathcal{F}_i to datum D_i at time t_i can then be written as

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

for the first test case and as

$$\mathcal{F}_i = D_0 \int_{-\infty}^{\infty} d\log \tau f(\log \tau) \int_{t_0}^{\min(t_i, t_f)} dt' R(t') e^{-(t_i - t')/\tau} + \xi R(t_i) \quad (2)$$

for the second application. Each distribution is represented on a discrete set of points that are equally spaced in $\log \tau$. D_0 is a normalization constant, and scattered excitation light can be accounted for with positive values of ξ . The response function R is peaked at zero time and is appreciable only on the interval $[t_0, t_f]$, with $t_0 < 0$. The zero-time shift [14,15] between the measured kinetics and instrument response is assumed to be negligible. The data are fit using a smoothed and renormalized response function.

Because the goal of a MEM calculation is to fit the data well using an “image” with maximal entropy, the function maximized is $Q \equiv S - \lambda C$, where S is the entropy of the image, C is a goodness-of-fit statistic, and λ is a Lagrange multiplier. No additional Lagrange multiplier is used here to constrain normalization. For data with Gaussian noise (first test case), C is the familiar chi-square:

$$C(\mathbf{D}, \mathcal{F}) = \chi^2 = \frac{1}{N} \sum_{i=1}^N \left(\frac{\mathcal{F}_i - D_i}{\sigma_i} \right)^2 \quad (3)$$

where σ_i is the standard error of the mean. For data with Poisson noise (second test case), C is the Poisson deviance:

$$C(\mathbf{D}, \mathcal{F}) = \frac{2}{N} \sum_{i=1}^N [D_i \ln(D_i / \mathcal{F}_i) + \mathcal{F}_i - D_i] \quad (4)$$

The entropy S of image \mathbf{f} is given by [16]

$$S(\mathbf{f}, \mathbf{F}) = - \sum_{j=1}^M [f_j \ln(f_j / F_j) + F_j - f_j] \quad (5)$$

where \mathbf{F} is the prior model used to incorporate existing knowledge into the MEM solution. Unconstrained maximization of S without regard for the fit ($\lambda = 0$) yields $f_j = F_j$.

Note that for Poisson data, $C(\mathbf{D}, \mathcal{F})$ and $S(\mathbf{f}, \mathbf{F})$ share the same functional form to within a multiplicative constant. The symmetry in Q is apparent in Eqs. (4) and (5); the kinetics data represent a set of N points with Poisson expectations \mathcal{F}_i , whereas the MEM image is a set of M values with Poisson expectations F_j . Thus, the MEM calculation drives the fit toward the data and the image toward the prior model by minimizing the corresponding Poisson deviances.

In the first application (Eq. (1)), \mathbf{f} includes the two distributions of lifetimes, \mathbf{g} and \mathbf{h} , that describe kinetics of opposite sign. In the second application (Eq. (2)), \mathbf{f} includes the distribution of lifetimes and the scattered-light parameter ξ . For simplicity, \mathbf{f} is referred to as

the lifetime distribution. The ability to adapt \mathbf{F} to the measured signal, to bootstrap the prior model, lends flexibility to the MEM that is lacking in other regularization methods such as those that minimize the sum of squared amplitudes [17]. This flexibility inherent in the MEM (Eq. (5)) is rather advantageous in practice because the bias introduced by regularization can be tuned by changing \mathbf{F} . Consequently, various bootstrapping approaches have been developed [8]. By contrast, MEM analyses that define the entropy simply as $S = - \int \ln f$ assume a uniform prior model and may suffer from artifacts as a result. Similarly, regularization methods that optimize a function other than the entropy of Eq. (5) introduce a bias that cannot be reduced conveniently.

The prior model \mathbf{F} can be derived from the data in different ways, and it is wise to assess the sensitivity of the recovered distribution \mathbf{f} to reasonable changes in \mathbf{F} . Beginning with a uniform \mathbf{F} , the distribution recovered by the MEM can be blurred by convolution with a Gaussian to produce the bootstrapped prior model for the remainder of the calculation. This approach has been shown to improve results in image restoration [18] and in the analysis of kinetics [7]. When sharp and broad features coexist in the underlying distribution, so-called differential blurring can improve the results, as monitored by the goodness of fit and the autocorrelation of the residuals [7,8]. In this case, relatively sharp peaks in \mathbf{f} are made even sharper before incorporating them into the prior model, whereas the remainder of \mathbf{f} is blurred. Differential blurring can suppress artifactual ripples that persist when the entire lifetime distribution is blurred uniformly. Another way to modify \mathbf{F} so as to reduce the bias of regularization in a reasonable and intuitive way is introduced in the next section.

Methods

A new means of bootstrapping the prior model was implemented in MemExp, version 4.0 (available at <http://www.cmm.cit.nih.gov/memexp>), as follows. A simple filter is used to suppress any potentially artifactual peak that arises at short lifetimes, for example, when scattered excitation light can be mistaken for a fast process in the kinetics. The MEM calculation is initiated using a uniform prior. Once C drops below a specified value (~ 1.0), the mean, area, and full width at half maximum are estimated numerically for each peak in \mathbf{f} that is to be represented in the bootstrapped prior, that is, any feature peaked at a lifetime longer than a user-specified cutoff. Here, the cutoff was set at $\log \tau = -1.5$ for the first application and at $\log \tau = -1.2$ for the second, values chosen to exclude the earliest peak observed when using a uniform prior. The area of each peak to be included is estimated by integrating over the distribution between the local minima immediately preceding and following the peak. Of course, the \mathbf{g} and \mathbf{h} distributions are treated independently. For peaks that are not resolved well enough to directly calculate the half width at half maximum on either side, the half width is simply approximated as the difference between the locations of the peak and the adjacent minimum. If the peak is well separated on only one side, that half width is doubled to approximate the full width. The uniform prior is then replaced by a sum of Gaussians with the same means and areas as the peaks resolved in \mathbf{f} (at lifetimes longer than the cutoff), with all peak widths multiplied by a user-specified constant (1.5 used here). The MEM calculation is then resumed with the entropy redefined in terms of this new prior distribution \mathbf{F} . This ability to change \mathbf{F} (and S) in reasonable ways helps the user to assess which features in $f(\log \tau)$ are truly established by the kinetics being analyzed and which are dependent on the choice of model \mathbf{F} .

Results

Test case I: triexponential kinetics with amplitudes that sum to zero

The kinetics recently purported to expose inadequacies [17] in the program MemExp, and in iterative methods in general, is revisited here. Gaussian noise with a time-independent standard error of $\sigma = 0.2$, a value greater than that used by Mulligan et al. [17], was added to the kinetics simulated at 512 points given by $D(t) = e^{-t} - 2e^{-3.333t} + e^{-10t}$ (Fig. 1A). Lifetime distributions obtained with the program MemExp, version 3.0, using a uniform prior and MemExp, version 4.0, using a filtered prior are shown in Fig. 1B. D_0 was set to 1, and the “allin1” mode of MemExp was used to estimate the standard errors. Although the results of both MemExp runs compare favorably with results obtained using ALIA [17], those obtained with the adapted prior are clearly superior. Not only is the small peak at $\log \tau \sim -2$ obtained with a uniform prior suppressed when using the filtered prior, but the means of the two outer peaks improve considerably. The areas of the three peaks resolved using the filtered prior are 1.039, -2.020 , and 0.951, respectively.

Test case II: Poisson kinetics, deconvolution of instrument response, and scattered light

Five different data sets were simulated as a sum of three exponentials with equal amplitudes (1000) and lifetimes equal to 0.1, 0.3, and 1.0. The kinetics was convolved with a Gaussian instrument response with a full width at half maximum of 0.1, and scattered light was added with $\xi = 150$. Poisson noise was added to the kinetics and to the instrument response. The interval between successive time points was $\Delta t = 0.005$. Because the inversion of Eq. (2) is performed in MemExp after renormalizing the known instrument response, the normalization can be initialized to a value approximating the maximum value of the data. Here, D_0 was set to 3000. The integral of the recovered \mathbf{f} distribution, multiplied by D_0 , ultimately determines the normalization of the fit.

Each of the five data sets was numerically inverted using a uniform prior of $F_j = 0.001$ and using a bootstrapped prior that was filtered to omit peaks resolved at $\log \tau < -1.2$ when using the uniform prior. For three of the five kinetic traces, the use of a uniform prior resulted in lifetime distributions with three phases, in accord with the simulated data (not shown). For the other two data sets, an artifactual peak appeared at short lifetimes in the inversion using a uniform prior. Results are shown in Fig. 2 for one of these two data sets. Distributions recovered at a Poisson deviance of 1.041 are shown for the uniform and bootstrapped priors. The integral of the \mathbf{f} distribution is 1.692 for the inversion using a uniform prior and 1.091 for the inversion using the filtered prior. The areas of the four peaks resolved using a uniform prior are 0.701, 0.322, 0.313, and 0.356. The areas of the three peaks resolved using the filtered prior are 0.387, 0.355, and 0.348. Analysis of the kinetics using the two priors leads one to correctly conclude that the additional fast process recovered with a uniform prior is not needed and that three phases of approximately equal weight describe the data. Moreover, the scattered light is estimated more accurately with the adapted prior ($\xi = 141.2$) than with the uniform prior ($\xi = 125.1$); the correct value is 150. By incorporating only the slower peaks in \mathbf{F} , the fit at short times is influenced by the data at long times where ambiguities due to scattered light and finite pulse width are much smaller. By emphasizing the slower processes, as opposed to an assumed or estimated normalization, this approach can filter artifacts from the distribution of lifetimes.

Discussion

In their recent article, Mulligan et al. [17] advocated minimization of the sum of squared amplitudes using their program ALIA, rather than the maximization of entropy, to infer rate distributions from noisy kinetics. They stated,

“Analysis can also be hindered if the normalization of the signal must be estimated from the data given that inversion by the maximum entropy method used by MemExp is dependent on this information. Because ALIA requires no information about signal normalization, MemExp was also not provided with the normalization for this test. We found that in the case of this test dataset, MemExp was unable to produce an accurate rate spectrum.”

These claims are clearly unsubstantiated. It is more accurate to say that because MemExp was written with the goal of generality and broad applicability, the MemExp user has options, and not all options are appropriate in all cases. In particular, the user may prompt the program to estimate the normalization of the signal, D_0 in Eqs. (1) and (2). The documentation warns that this estimation is not recommended for all applications. Alternatively, the user can set D_0 to specify normalization that is known independently. Of course, the user can also choose to set $D_0 = 1$ to perform the inversion analogous to that performed using ALIA. With $D_0 = 1$, it is easily shown that the MemExp user can produce an accurate rate spectrum for the triexponential test case. Moreover, one can use MemExp to successfully obtain the rates, even when the Gaussian noise added is increased substantially (Fig. 1). Therefore, the assertion by the authors that ALIA can be used to analyze “datasets too noisy to process by existing iterative search algorithms” has not been established. The strange results they obtained using MemExp have nothing to do with the regularization choice: Eq. (5) versus χ^2 . Rather, they wrote their program precluding a user-specified normalization, which can be useful in some applications, and they used MemExp in an unreasonable way. Although it is true that iterative methods are relatively slow computationally, speed is typically not a concern in the analysis of kinetics.

There are several advantages to maximizing the entropy, instead of minimizing the sum of squared amplitudes, that have been well documented for some time. As was discussed in 1990, in the absence of prior knowledge, maximizing the entropy, $S = -\sum \ln f_i$, imposes no correlation in simple examples such as the “kangaroo problem,” whereas minimizing χ^2 does impose correlation [19]. In the presence of prior knowledge, the entropy can be generalized to good effect by defining the entropy in terms of a prior model (Eq. (5)). The adaptation of this model to the data being analyzed can be useful in reducing artifacts and/or resolving ambiguities (Figs. 1B and 2C). In addition, evaluation of the entropy requires the natural logarithm of each amplitude, ensuring that amplitudes remain positive. Analyses that minimize the sum of squared amplitudes will be plagued by artifactual lobes in the rate spectrum of the wrong sign.

The two examples analyzed here highlight the advantage of the MEM’s flexibility to alleviate bias by adapting the prior model to the data being analyzed. Improved results are obtained in both cases simply by discarding peaks at very short lifetimes and incorporating the others into the definition of S . In the second example, qualitatively different lifetime distributions, one with three peaks and one with four peaks, can describe the noisy data equally well. The corresponding fits differ only slightly at short times (Fig. 2B) because the presence of scattered excitation light introduces ambiguity. The incorporation of the incorrect fast process (and greater signal normalization) in the distribution recovered with a uniform prior is compensated for by its incorrect (reduced) estimate of the scattered light. The incorrect fast process is clearly undesirable given that it is peaked near $\log \tau = -2$ and the data are sampled at intervals of $\Delta t = 0.005$. When faced with such an ill-posed problem, multiple interpretations of the data are possible, and use of the adaptable prior automated here can help to conveniently identify the simplest interpretation.

The recovery of a distribution of effective lifetimes from noisy kinetics data is clearly complicated whenever there is uncertainty in the value of the signal at short times. Distributions recovered by the MEM using a uniform prior ($F_j = \text{constant}$) can include

unwarranted features if the normalization is underestimated [20]. The method introduced here is attractive because the fit at short times is guided by the data at longer times. The prior is bootstrapped so as to suppress artifacts that arise from ambiguities at short times, for example, due to scattered light. Whenever use of a filtered prior eliminates a peak at short lifetimes, as in the current examples, a simpler interpretation of the data has been found. It is always important to assess the sensitivity of the MEM result to the choice of prior distribution, and the current approach should be among those methods routinely applied.

Acknowledgments

This study was supported by the Intramural Research Program of the National Institutes of Health (NIH), Center for Information Technology (CIT). I thank Wesley Asher for stimulating discussions and Sergio Hassan and Charles Schwieters for comments on the manuscript.

References

1. Gull SF, Daniell GJ. Image reconstruction from incomplete and noisy data. *Nature*. 1978; 272:686–690.
2. Skilling J, Bryan RK. Maximum entropy image reconstruction—general algorithm. *Mon. Not. R. Astron. Soc.* 1984; 211:111–124.
3. Livesey AK, Brochon JC. Analyzing the distribution of decay constants in pulse fluorimetry using the maximum entropy method. *Biophys. J.* 1987; 52:693–706. [PubMed: 19431708]
4. Lavalette D, Tetreau C, Brochon J-C, Livesey A. Conformational fluctuations and protein reactivity: determination of the rate constant spectrum and consequences in elementary biochemical processes. *Eur. J. Biochem.* 1991; 196:591–598. [PubMed: 2013282]
5. Steinbach PJ, Ansari A, Berendzen J, Braunstein D, Chu K, Cowen BR, Ehrenstein D, Frauenfelder H, Johnson JB, Lamb DC, Luck S, Mourant JR, Nienhaus GU, Ormos P, Philipp R, Xie A, Young RD. Ligand binding to heme proteins: connection between dynamics and function. *Biochemistry*. 1991; 30:3988–4001. [PubMed: 2018767]
6. Steinbach PJ, Chu K, Frauenfelder H, Johnson JB, Lamb DC, Nienhaus GU, Sauke TB, Young RD. Determination of rate distributions from kinetic experiments. *Biophys. J.* 1992; 61:235–245. [PubMed: 1540692]
7. Steinbach PJ, Ionescu R, Matthews CR. Analysis of kinetics using a hybrid maximum-entropy/nonlinear-least-squares method: application to protein folding. *Biophys. J.* 2002; 82:2244–2255. [PubMed: 11916879]
8. Steinbach PJ. Inferring lifetime distributions from kinetics by maximizing entropy using a bootstrapped model. *J. Chem. Inf. Comput. Sci.* 2002; 42:1476–1478. [PubMed: 12444746]
9. Steinbach PJ. Two-dimensional distributions of activation enthalpy and entropy from kinetics by the maximum entropy method. *Biophys. J.* 1996; 70:1521–1528. [PubMed: 8785309]
10. Abbruzzetti S, Faggiano S, Bruno S, Spyrikis F, Mozzarelli A, Dewilde S, Moens L, Viappiani C. Ligand migration through the internal hydrophobic cavities in human neuroglobin. *Proc. Natl. Acad. Sci. USA*. 2009; 106:18984–18989. [PubMed: 19850865]
11. Lepeshkevich SV, Biziuk SA, Lemeza AM, Dzhagarov BM. The kinetics of molecular oxygen migration in the isolated alpha chains of human hemoglobin as revealed by molecular dynamics simulations and laser kinetic spectroscopy. *Biochim. et Biophys. Acta*. 2011; 1814:1279–1288.
12. Hassan SA. Computer simulation of ion cluster speciation in concentrated aqueous solutions at ambient conditions. *J. Phys. Chem. B*. 2008; 112:10573–10584. [PubMed: 18680338]
13. Zimmermann J, Romesberg FE, Brooks CL III, Thorpe IF. Molecular description of flexibility in an antibody combining site. *J. Phys. Chem. B*. 2010; 114:7359–7370. [PubMed: 20455589]
14. Isenberg I. On the theory of fluorescence decay experiments: I. Nonrandom distortions. *J. Chem. Phys.* 1973; 59:5696–5707.
15. Bajzer Z, Zelic A, Prendergast FG. Analytical approach to the recovery of short fluorescence lifetimes from fluorescence decay curves. *Biophys. J.* 1995; 69:1148–1161. [PubMed: 8519969]

16. Skilling, J. Maximum Entropy and Bayesian Methods. Kluwer, Dordrecht, The Netherlands; 1989. Classic maximum entropy; p. 45-52.
17. Mulligan VK, Hadley KC, Chakrabarty A. Analyzing complicated protein folding kinetics rapidly by analytical Laplace inversion using a Tikhonov regularization variant. *Anal. Biochem.* 2012; 421:181–190. [PubMed: 22119751]
18. Gull, SF. Maximum Entropy and Bayesian Methods. Kluwer, Dordrecht, The Netherlands; 1989. Developments in maximum entropy data analysis; p. 53-71.
19. Sivia DS. Bayesian inductive inference maximum entropy and neutron scattering. *Los Alamos Sci.* 1990; 19:180–206.
20. Kumar ATN, Zhu L, Christian JF, Demidov AA, Champion PM. On the rate distribution analysis of kinetic data using the maximum entropy method: applications to myoglobin relaxation on the nanosecond and femtosecond timescales. *J. Phys. Chem. B.* 2001; 105:7847–7856.

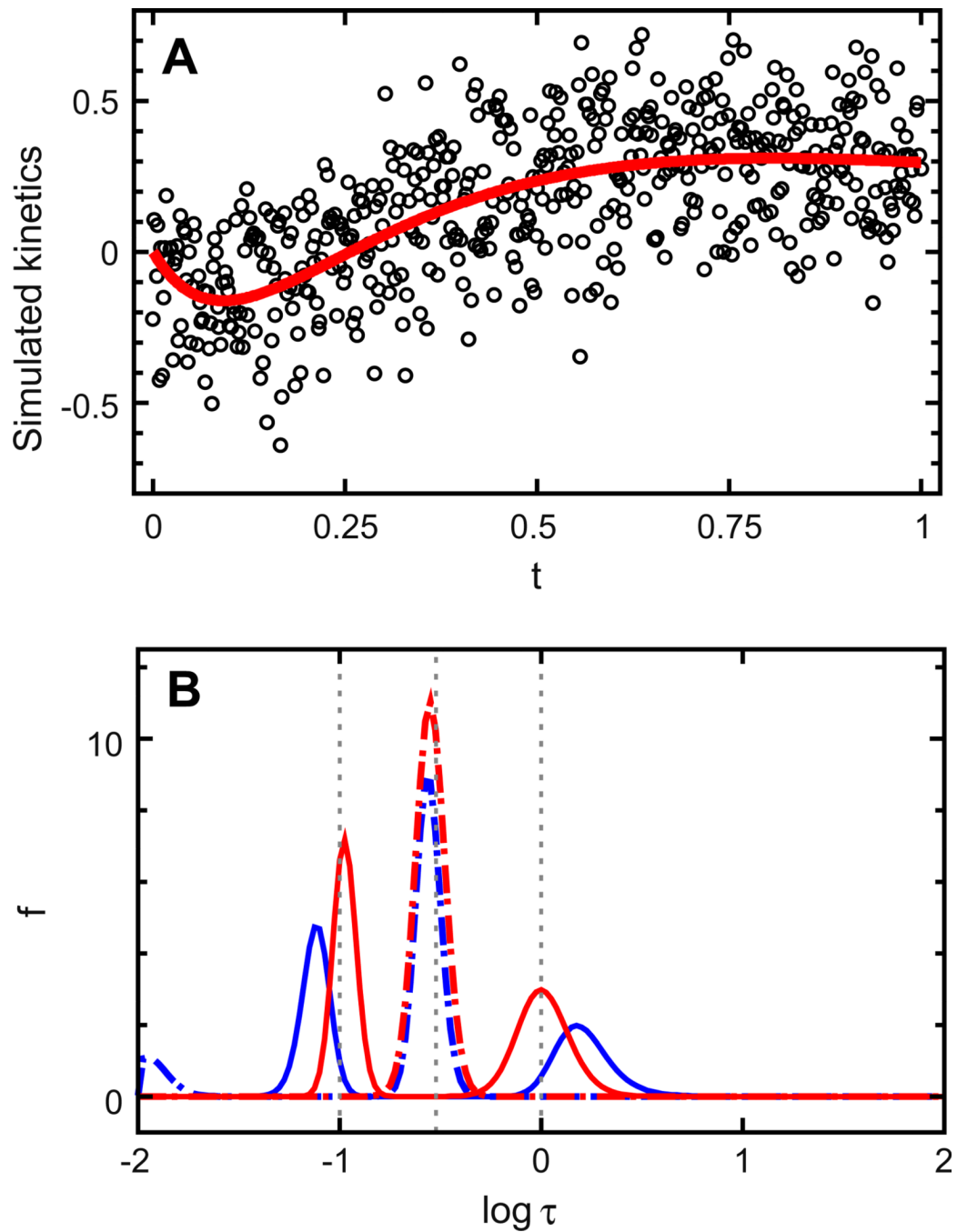


Fig. 1. MemExp analysis of kinetics simulated as a sum of three exponentials with amplitudes that sum to zero. (A) Simulated data with (circles) and without (red line) Gaussian noise added. (B) Lifetime distributions recovered at $\chi^2 = 0.850$ using a uniform prior model (blue) and using a model bootstrapped from the data by omitting features resolved with the uniform model peaked at $\log \tau < -1.5$ (red). Peaks shown as dot-dashed lines correspond to negative amplitudes. Vertical dotted lines indicate the “true” lifetimes.

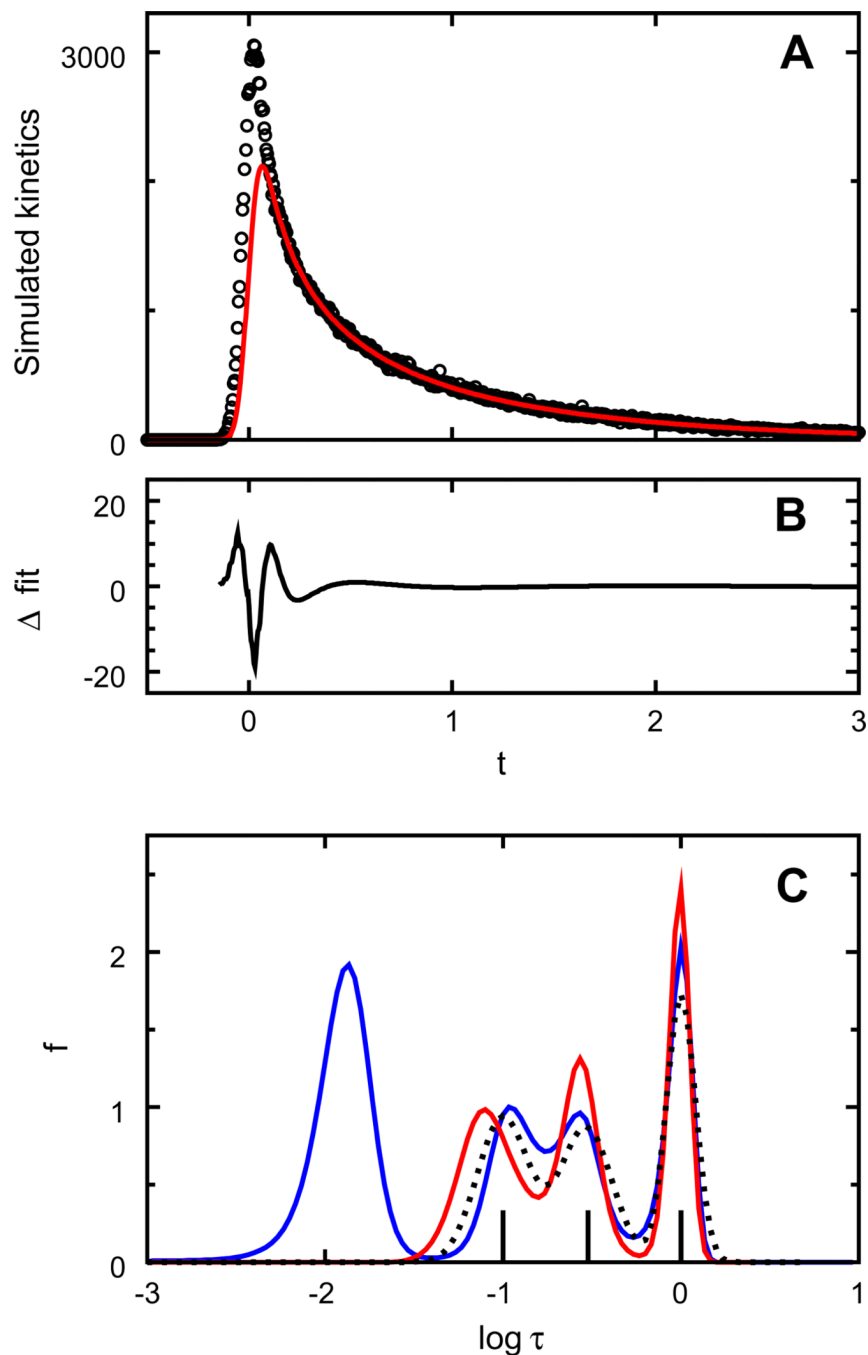


Fig. 2. MemExp analysis of kinetics simulated as a sum of three exponentials convolved with a Gaussian instrument response. (A) Simulated data with (circles) and without (red line) scattered light and Poisson noise added. (B) Difference in the two fits to the noisy data obtained at a Poisson deviance of 1.041. (C) Corresponding lifetime distributions recovered using a uniform prior model (blue) and using a model bootstrapped from the data by omitting features resolved with the uniform model peaked at $\log \tau < -1.2$ (red). Also shown are the bootstrapped model (dotted line) and the “true” exponentials (vertical bars).