Exponential Parameter Estimation (in NMR) Using Bayesian Probability Theory

G. LARRY BRETTHORST,[1] WILLIAM C. HUTTON,[1] JOEL R. GARBOW,[1,2] JOSEPH J.H. ACKERMAN[1–3]

[1] *Department of Radiology, Washington University, One Brookings Drive, St. Louis, MO 63130, USA*
[2] *Department of Chemistry, Washington University, One Brookings Drive, St. Louis, MO 63130, USA*
[3] *Department of Internal Medicine, Washington University, One Brookings Drive, St. Louis, MO 63130, USA*

ABSTRACT: Data modeled as sums of exponentials arise in many areas of science and are common in NMR. However, exponential parameter estimation is fundamentally a difficult problem. In this article, Bayesian probability theory is used to obtain optimal exponential parameter estimates. The calculations are implemented using Markov chain Monte Carlo with simulated annealing to draw samples from the joint posterior probability for all of the parameters appearing in the exponential model. Monte Carlo integration is then used to approximate the marginal posterior probabilities for each of the parameters. We give numerical examples taken from simulated data and NMR relaxation experiments to illustrate the calculations and the effect of prior information on the parameter estimates.
 Concepts Magn Reson Part A 27A: 55–63, 2005

KEY WORDS: exponential data analysis; rate constant; time constant; biexponential

INTRODUCTION

Magnetic resonance experiments often generate data that are modeled as a sum of exponentials (*1–11*). Experiments relying on NMR to probe reaction kinetics, diffusion, molecular dynamics, and xenobiotic metabolism are only some of the applications where parameter estimates from exponential models provide insight into chemical and biological processes. Recent reviews discuss the analysis and methods commonly used to estimate exponential parameters (*12*, *13*). Exponential parameter estimation can be challenging (*14*, *15*), and low signal-to-noise ratios or sparse data sets, common in magnetic resonance, only exacerbate the difficulties (*16*). In this article, an extension of a previous brief exposition (*17*), we address the exponential parameter estimation problem by applying Bayesian probability theory. Because the parameter

Received 14 June 2005; revised 19 August 2005; accepted 19 August 2005

Correspondence to: G. Larry Bretthorst; E-mail: gbretthorst@wustl.edu

Concepts in Magnetic Resonance Part A, Vol. 27A(2) 55–63 (2005)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/cmr.a.20043

estimates are made using Bayesian probability theory, one is able to make well-informed judgments about the uncertainties associated with the parameter estimates.

We model the data as a sum of exponentials

$$d_i = C + \sum_{j=1}^{m} A_j \exp\{-\alpha_j t_i\} + n_i, \qquad [1]$$

where m is the number of exponentials and d_i is a data value sampled at t_i. The parameters of interest are the decay rate constants, α_j, the amplitudes, A_j, and, if present, the constant offset C. The error parameters, n_i, are commonly referred to as noise. In the frequency interpretation of probability one must make assumptions about the sampling frequency distribution of the noise. In Bayesian probability theory no such assumptions are made; rather, one assigns probabilities to represent what is actually known about the noise.

In any given problem, one may or may not know the number of exponentials or whether the constant offset is present. In Bayesian probability theory, when these are known, the problem is one of parameter estimation. When the value of m or the presence of C is not known, the problem is one of model selection. Both problems are solved using Bayes' theorem (*18*) and the rules of probability theory (*19*). However, the parameter estimation and model selection problems have different solutions. In this article, we address the parameter estimation problem. The model selection problem is addressed in a companion article (*20*).

THEORY

Our goal is to estimate all of the decay rate constants, amplitudes, and offsets appearing in Eq. [1]. In Bayesian probability theory, everything known about a parameter is summarized in a probability density function. If the data are biexponential, there are four different probability density functions to compute. All of these probability density functions can be computed from the joint posterior probability for all of the parameters given the data, D, and the prior information, I, by application of the sum rule. For example, the posterior probability for α_1 is computed as

$$P(\alpha_1|DI) = \int d\alpha_2 dA_1 dA_2 dC P(\alpha_1\alpha_2 A_1 A_2 C|DI). \qquad [2]$$

In a similar manner, the posterior probability for A_1 is given by

$$P(A_1|DI) = \int d\alpha_1 d\alpha_2 dA_2 dC P(\alpha_1\alpha_2 A_1 A_2 C|DI). \qquad [3]$$

The process of integrating over unneeded or nuisance parameters to obtain the probability for a parameter of interest is called marginalization. If the model contains more than two exponentials, one must evaluate more integrals to obtain a marginal posterior probability, but the principle remains unchanged.

If we designate all of the decay rate constants by $\alpha \equiv \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, the amplitudes by $A \equiv \{A_1, A_2, \ldots, A_m\}$ and the constant, if present, by C, then the joint posterior probability for the parameters is represented symbolically by $P(\alpha AC|DI)$. We write this joint posterior probability as if C is known to be present; if C is not present, then any notation involving C is absent. To compute the joint posterior probability for all the parameters, one applies Bayes' theorem:

$$P(\alpha AC|DI) = \frac{P(\alpha AC|I)\, P(D|\alpha ACI)}{P(D|I)}, \qquad [4]$$

where $P(\alpha AC|I)$ is the joint prior probability for the parameters and $P(D|\alpha ACI)$ is the direct probability for the data. In parameter estimation problems the marginal posterior probability for the data, $P(D|I)$, is a normalization constant and is dropped from the calculation. However, in model selection problems, $P(D|I)$ is the quantity of interest. Here, we drop the normalization constant and factor the joint prior probability for the parameters to obtain

$$P(\alpha AC|DI) \propto P(C|I) \prod_{j=1}^{m} \times [P(\alpha_j|I)\, P(A_j|I)] P(D|\alpha ACI), \qquad [5]$$

where terms of the form $P(\cdot|I)$ are prior probabilities, (i.e., they depend only on the prior information I). At this point in the calculation, all of the prior probabilities can be assigned.

However, the direct probability for the data, $P(D|\alpha ACI)$ in Eq. [5], does not include the error parameters, n_i, nor does it, in any obvious way, include information about the noise characteristics. This direct probability cannot be assigned because these noise characteristics were removed by marginaliza-

Table 1 Parameter Estimates Obtained from Simulated Biexponential Data

	Peak α_1	Mean α_1	Peak A_1	Mean A_1	Peak α_2	Mean α_2	Peak A_2	Mean A_2	Mean f_2:f_1	Mean $A_1 + A_2$
True value	1.85	—	32.5	—	13.09	—	217.5	—	13:87	250
S:N = 400	1.82	1.83 ± .08	32.6	32.7 ± 1.3	13.14	13.16 ± .15	218.3	218.3 ± 1.3	13:87	251 ± .9
S:N = 40	1.8	1.9 ± .5	43	45 ± 11	13.1	13.4 ± 1.4	203	202 ± 11	18:82	247 ± 8
S:N = 20	1.7	2.2 ± .8	53	67 ± 22	12.9	15.6 ± 4.9	189	181 ± 22	27:73	249 ± 19

The simulated data contain 16 data points uniformly sampled over a 700-ms period. A Markov chain Monte Carlo simulation was used to draw samples from the joint posterior probability for the parameters. Monte Carlo integration was then used to obtain samples from the marginal posterior probability for each of the four parameters. The columns labeled "Peak" are parameters from the Markov chain Monte Carlo simulation with the highest overall joint posterior probability. The mean and standard deviations are the mean and standard deviations of the samples computed for each of the four parameters. The column labeled f_2:f_1 are the mean fractional amplitudes at $t = 0$.

tion. Appendix A provides further details on the Bayesian calculations, the assignment of all the probabilities, and the marginalizations necessary to obtain $P(D|\alpha ACI)$.

To compute the marginal posterior probability for each parameter, one must integrate $P(\alpha AC|DI)$ over all of the parameters except the one of interest. No analytic solution exists for these integrals so they must be evaluated numerically. We evaluate these integrals by first using Markov chain Monte Carlo with simulated annealing (*21*, *22*) to draw samples from the joint posterior probability for the parameters. We usually run 50 to 100 Markov chain Monte Carlo simulations in parallel. After the annealing phase, we draw 50 to 100 independent samples from each simulation, which generates 2,500 to 10,000 samples from the joint posterior probability. Quantitative evaluation of Markov chain convergence remains an active research area, and no universal solutions are known. We assess Markov chain convergence visually by plotting the logarithm of the posterior probability for each simulation as a function of the sample repeat number. When converged, the individual trajectories consistently cross each other and appear stationary. Monte Carlo integration is then used to compute means, standard deviations, and histograms for each parameter of interest. The histograms serve only as visual representations of the marginal distributions and are not used in any calculations. See our companion article in this issue and a recent review for details on the use of Markov chain Monte Carlo in Bayesian probability theory (*20*, *23*).

RESULTS AND DISCUSSION

We begin by considering the results from a Bayesian analysis of simulated data. The simulation mimics water relaxation experiments with human-brain gray matter in vivo. Biexponential relaxation is reported in stimulated echo spectra from localized voxels along the Sylvian fissure (*6*). The parameters used to generate the simulated data are from these empirical observations (see Table 1). The only parameter varied was the noise level. Three data sets with signal-to-noise ratios of 400:1, 40:1, and 20:1 were analyzed. Table 1 lists the parameter estimates from these three analyses. Each mean and standard deviation in Table 1 was computed from 10,000 Markov chain Monte Carlo samples (100 simulations sampled 100 times each). A typical analysis completes in less than 90 s on a Sun Microsystems Ultra 60, dual 450 MHz workstation.

When the signal-to-noise ratio is high, 400:1, the Bayesian parameter estimates are essentially identical to the simulation inputs. In the moderate signal-to-noise ratio data, 40:1, all of the parameter estimates are less certain by about a factor of 10. Parameter estimate uncertainties from the low signal-to-noise ratio data, 20:1, are about twice those from the 40:1 data. The direct dependence of the width of the posterior probability (the parameter estimate uncertainty) on the noise standard deviation is a general feature in probability theory and is found in virtually all Bayesian calculations. Despite the large uncertainties in the low signal-to-noise data, in all cases the estimated values of the parameters are within plus or minus one or two standard deviations of their true values.

The simulated data in Fig. 1 illustrates a common problem in magnetic resonance data, low signal-to-noise ratios. Although the peak signal-to-RMS noise level of the data is 40:1, the second data point's signal-to-noise ratio drops to 25:1, and more than half the data have signal-to-noise ratios of 5:1 or less.

Figure 2 displays the marginal posterior probability density for the small decay rate constant (α_1, upper panel) and its amplitude (A_1, lower panel) computed from the simulated 400:1, 40:1, and 20:1 signal-to-noise ratio data. The hypotheses that best agree with the data and the prior information are those with the

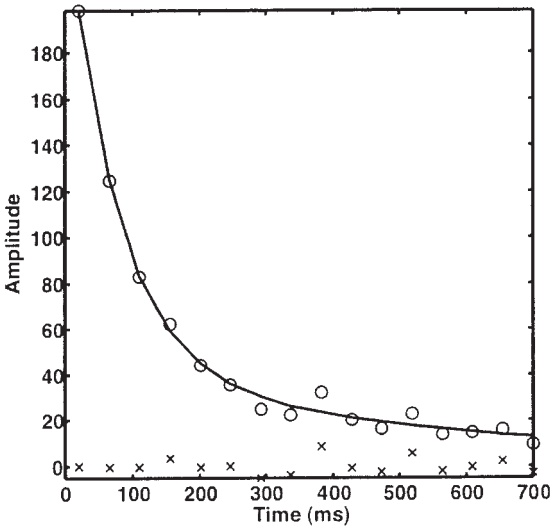

Figure 1 A comparison of simulated biexponential data, (○) and the model decay curve (solid line) generated from the maximum joint posterior probability for the parameters. The residuals, the difference between the data and the model, are shown at the bottom (x). The parameters used to generate the simulated data are given in Table 1.

highest posterior probabilities. A probability density function's width is a natural measure of how uncertain one is of the true, but unknown, value. Visual inspection of the functions' widths efficiently communicates the relative quality of the estimates.

We next illustrate how Bayesian parameter estimates are affected by model choice and prior probabilities. Table 2 summarizes analysis results from a ^{23}Na transverse relaxation study of rat brain in vivo (James Goodman, personal communication). The relaxation data were collected with a surface coil using slice-selection methods and a LASER pulse sequence (*24*). We assume the ^{23}Na transverse relaxation is biexponential and use three different sets of priors: loose, tight, and flat to analyze the data. In all cases, we use bounded Gaussian prior probabilities (see Appendix A for assignment details) and adjust their bounds, means, and standard deviations. The tight prior bounds the parameters to within one order of magnitude of their maximum likelihood estimates, whereas the loose prior restricts them to within two orders of magnitude. The flat prior bounds are extremely wide and are so uninformative that the analysis essentially uses a uniform prior probability around the high likelihood region of parameter space.

The results of the biexponential analyses are found in the first three rows of Table 2. Note that changing the bounds, means, and standard deviations of the prior probabilities have almost no effect on the parameter estimates. Unless the width of the joint prior probability density function is comparable with the width of the direct probability for the data, the joint prior probability has essentially no effect on the results. Though the tight prior is much more informative than the others, it is still uninformative (wide) compared with the direct probability.

The joint prior probability for the parameters does affect the Markov chain Monte Carlo simulation used in the calculation. The uninformative (flat) joint prior probability increases the size of the four-dimensional parameter space and requires a threefold increase in the number of Markov chain annealing steps to achieve convergence. However, the resulting parameter estimates are essentially the same as those from the other two analyses.

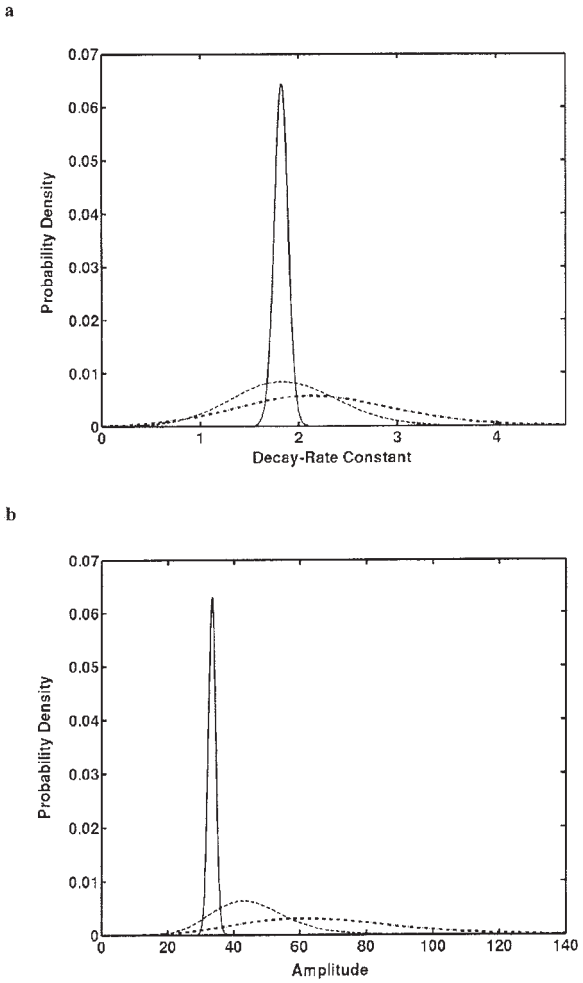

Figure 2 The marginal posterior probability density for the decay rate constant α_1 (top); the marginal posterior probability for the amplitude A_1 (bottom). The three posterior probabilities shown in each of these plots are computed from the simulated 400:1 (—), 40:1 (- - -), and 20:1 (•–•–•) signal-to-noise ratio data.

Table 2 Parameter Estimates Obtained from ^{23}Na Transverse Relaxation Data (Rat Brain in vivo at 4.7 T)

m	Priors	α_1	A_1	α_2	A_2	f_2:f_1
2	Loose	23.3 ± 1.0	36 ± 2	106 ± 25	19 ± 4	35:65
2	Tight	23.1 ± 1.1	36 ± 2	101 ± 24	20 ± 3	36:64
2	Flat	23.2 ± 1.1	36 ± 2.4	106 ± 29	20 ± 4	36:64
1	Loose	27.9 ± 1.3	46 ± 1	—	—	—

Nine data points were collected over 140 ms. A Markov chain Monte Carlo simulation was used to draw samples from the joint posterior probability for the parameters given a one- or two-component exponential model. The estimated parameters are the mean and standard deviation of these samples. The column labeled f_2:f_1 are the mean fractional amplitudes at $t = 0$. The loose analysis used the prior bounds set based on the total acquisition time. These bounds span a little more than two orders of magnitude. In the tight analysis the priors were much more informative, spanning only one order of magnitude. The flat analysis continues to use a Gaussian prior but with the bounds so wide that the prior is essentially flat over the region of parameter space supported by the likelihood, i.e., a flat, uniform prior. The final analysis uses a monoexponential model and the priors span about two orders of magnitude. Note the monoexponential α estimate differs from the biexponential model α estimates by about five standard deviations.

Reports of mono- and biexponential ^{23}Na transverse relaxation in vivo are found in the literature (*25–29*). In the biexponential cases, the reported ^{23}Na decay rate constants range from 200 to 1,400 sec^{-1} for the rapidly decaying component, and 25 to 145 sec^{-1} for the slowly decaying component. The reported monoexponential decay rate constants fall in the range of 18 to 55 sec^{-1}. What happens when we assume the model for our ^{23}Na data is a monoexponential? The results for the single exponential model with loose priors are shown as the last line in Table 2. The monoexponential's decay rate constant and amplitude estimates are significantly different (more than five standard deviations) from the parameter estimates computed with a biexpoential model.

The ^{23}Na data analysis points out the importance of the model selection in obtaining meaningful parameter estimates. An analysis based on a model where either the number of exponentials or the presence of a constant does not correspond to the data will give incorrect results. The results are incorrect in the sense that none of the parameter estimates accurately reflect the unknown true signal parameters. The adverse consequence of analyzing exponentially decaying data using the wrong model cannot be overemphasized.

To get some impression of which model is correct, we plotted the residuals computed from the maximum posterior probability parameter estimates for the mono- (dotted line) and biexponential (solid line) models (Fig. 3). The biexponential residuals look random; they oscillate above and below zero. By contrast, the monoexponential residuals are first above zero, then below, and then back above zero in a systematic way. This type of systematic behavior is characteristic of residuals when the model does not match the data and is significant, but qualitative, evidence that these data are biexponential. Bayesian probability theory solves the model selection problem by computing the posterior probability for the number of exponentials in the data and whether a constant is present. In our companion article in this issue, a Bayesian exponential model selection calculation is described and is applied to these data (*20*).

SUMMARY AND CONCLUSION

Bayesian probability theory provides a rigorous theoretical framework for the analysis of data modeled as a sum of exponentials. Bayesian posterior probability density functions provide optimal exponential parameter estimates as well as a quantitative assessment of the estimates' uncertainties. We have implemented these Bayesian calculations on single and parallel computational architectures. The software is fully integrated into a commercial NMR data analysis pack-

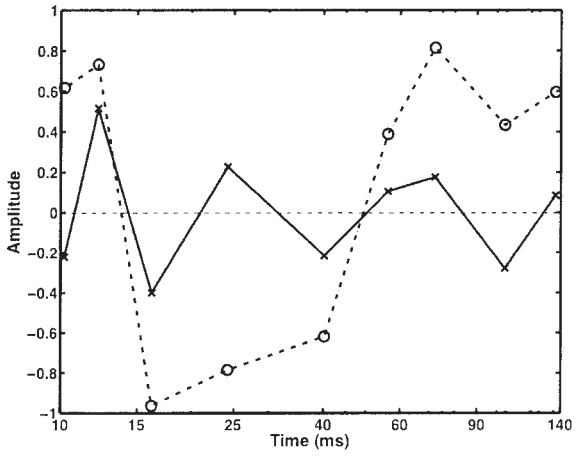

Figure 3 The residuals, the difference between the data and the model, for the ^{23}Na relaxation analysis using the loose joint prior probability density function for mono- (- - - -) and biexponential (—) models.

age (VnmrJ, Varian NMR Systems, Palo Alto, CA). Both the Vnmrj implementation and standalone executables for manual use are available for free download at http://BayesianAnalysis.wustl.edu.

ACKNOWLEDGMENTS

We thank Dr. Jeffrey J. Neil for his encouragement, support, and comments. We thank James Goodman for supplying the ^{23}Na relaxation data. This work was supported by a contract with Varian NMR Systems, Palo Alto, CA; by the Small Animal Imaging Resource Program (SAIRP) of the National Cancer Institute, grant R24 CA83060; and by the National Institute of Neurological Disorders and Stroke, grants NS35912 and NS41519.

APPENDIX

A. The Calculations

In the theory section of this article, we factored the joint posterior probability for the parameters, but we did not explain how the resulting probabilities were assigned. In addition, Eq. [5] is a marginal probability from which the standard deviation of the noise prior probability was removed. Marginal probabilities cannot be assigned; rather, they must be computed using the rules of probability theory. In this appendix we detail how this joint marginal posterior probability, $P(\alpha AC|DI)$, is computed.

To compute Eq. [5] we reintroduce the standard deviation of the noise prior probability, σ. The joint posterior probability for the parameters, Eq. [5], is given by

$$P(\alpha AC|DI) \propto \int d\sigma P(C|I) \prod_{j=1}^{m} \times [P(\alpha_j|I)\, P(A_j|I)] P(D\sigma|\alpha ACI), \quad [6]$$

where the sum rule was used to remove the standard deviation of the noise prior probability. Factoring the last term in Eq. [6] using the product rule gives

$$P(\alpha AC|DI) \propto \int d\sigma P(C|I)\, P(\sigma|I) \prod_{j=1}^{m} \times [P(\alpha_j|I)\, P(A_j|I)] P(D|\alpha AC\sigma I), \quad [7]$$

where $P(\sigma|I)$ is the prior probability for σ, and $P(D|\alpha AC\sigma I)$ is the direct probability for the data given the parameters. These two probabilities are the only terms in this equation that depend on σ. If we separate these two probabilities in preparation for evaluating the integral over the standard deviation, we obtain

$$P(D|\alpha ACI) = \int d\sigma P(\sigma|I)\, P(D|\alpha AC\sigma I). \quad [8]$$

The exponential model equation, Eq. [1], contains a set of parameters, the n_i, along with the parameters of interest, α, A, and C. But the n_i do not appear in Eq. [8], nor do they appear in Eq. [5]. Consequently, the direct probability for the data, Eq. [8], must be a marginal probability from which both σ and the n_i were removed using the sum rule. To show how these hypotheses were removed, we reintroduce the n_i. If we designate $n \equiv n_1, n_2, \ldots, n_N$, the direct probability for the data, Eq. [8], is given by

$$P(D|\alpha ACI) = \int dn d\sigma P(\sigma|I)\, P(nD|\alpha AC\sigma I). \quad [9]$$

Factoring $P(nD|\alpha AC\sigma I)$ using the product and sum rules yields,

$$P(D|\alpha ACI) = \int dn d\sigma P(\sigma|I)\, P(n|\sigma) \times P(D|\alpha AC\sigma nI). \quad [10]$$

The three terms on the right-hand side of Eq. [10] are, respectively, the prior probability for the standard deviation, the prior probability for the errors given the standard deviation, and the direct probability for the data given all of the parameters, including the standard deviation and the n. But this last term is a delta function because of the model, Eq. [1]. Either the error plus the model is equal to the data, in which case the probability is 1, or the error plus the model do not add up to the data, in which case the probability is 0,

$$P(D|\alpha AC\sigma nI) = \prod_{i=1}^{N} \delta\left(d_i - n_i - C - \sum_{j=1}^{m} A_j e^{-\alpha_j t_i} \right). \quad [11]$$

The prior probability for the errors, given the standard deviation, must now be assigned. If all we know is the

standard deviation of the errors, and if we use the principle of maximum entropy to assign this prior probability, then maximum entropy will assign a 0 mean Gaussian of standard deviation σ,

$$P(n|\sigma) = (2\pi\sigma^2)^{-N/2}\exp\left\{-\sum_{i=1}^{N}\frac{n_i^2}{2\sigma^2}\right\}. \quad [12]$$

Equation [12] explicitly demonstrates that σ is the standard deviation of the prior probability used to represent what is known about the errors, and is not the standard deviation of the noise. No assumptions are made about the sampling frequency distribution of the errors. When maximum entropy is used to assign a Gaussian prior probability for the errors, all of the noise characteristics except the mean and second moments cancel from the calculations. Any noise sample having the same mean and second moments will yield exactly the same inference. See (*19*, chap. 7) and (*31*) for more on the near irrelevance of sampling-frequency distributions.

The prior probability for the standard deviation of the noise prior is typically assigned using a Jeffreys' prior (*30*),

$$P(\sigma|I) \propto \frac{1}{\sigma}. \quad [13]$$

Strictly speaking, the Jeffreys' prior is not a probability because it is not normalizable. If one adheres to the rules of probability theory, this Jeffreys' prior must be bounded and normalized, and the bounds should only be allowed to go to infinity as a limit at the end of the calculations. However, here the use of a Jeffreys' prior is harmless because the likelihood is so strongly peaked.

The integrals over the n_i are trivial, one obtains

$$P(D|\alpha ACI) \propto \int d\sigma\sigma^{-N-1}\exp\left\{-\frac{Q}{2\sigma^2}\right\}, \quad [14]$$

where

$$Q \equiv \sum_{i-1}^{N}\left(d_i - C - \sum_{j=1}^{m} A_j e^{\alpha_j t_i}\right)^2. \quad [15]$$

Equation [14] is a Gaussian likelihood and is usually assigned directly because it is taken as axiomatic. However, learning how these probabilities are assigned to represent states of knowledge is an important step in understanding how Bayesian probability theory should be used.

The integral over σ can be transformed into a Gamma function integral, and its evaluation yields

$$P(D|\alpha ACI) \propto \left[\frac{Q}{2}\right]^{-N/2}. \quad [16]$$

This equation is of the form of Student's *t*-distribution and is substituted into Eq. [5]. It is this *t*-distribution that is actually used in the Markov chain Monte Carlo simulations. Making this substitution, one obtains

$$P(\alpha AC|DI) \propto P(C|I)\prod_{j=1}^{m}[P(\alpha_j|I)\,P(A_j|I)]\left[\frac{Q}{2}\right]^{-N/2}, \quad [17]$$

and all that remains is to assign the other prior probabilities.

Though the prior probability for the parameters may have little or no effect on the final parameter estimates, they can affect the efficiency and convergence of the Markov chain Monte Carlo simulations. In these calculations the prior probabilities' primary function is to provide order-of-magnitude estimates for the parameters. These estimates are ascertained from known factors, such as the sampling time and the magnitude of the data. Because a bounded Gaussian correctly describes an order of magnitude estimate, we assign the prior probabilities for the amplitudes, the constant offset, and decay rate constants using broad uninformative Gaussians. These priors are assigned by specifying a low and high parameter range. From this range a mean value is computed, Mean = (low + high)/2. The standard deviation of the Gaussian prior probability is set so the entire interval, low to high, represents a three standard deviation interval, standard deviation = (high−low)/3. If L, H, M, and σ_ν represent the lower bound, the upper bound, the mean, and the standard deviation of the Gaussian prior probability, then the prior probability for parameter ν is given by

$$P(\nu|LH) \propto \begin{cases} e^{-[(\nu-M)^2]/2\sigma_\nu^2} & (L \le \nu \le H) \\ 0 & \text{otherwise} \end{cases} \quad [18]$$

where ν is any one of the parameters of interest. The prior probabilities, represented symbolically by Eq. [18], are substituted into Eq. [17] and the resulting equation is used in the numerical computation.

The model, Eq. [5], is completely symmetric under exchange of labels on the exponential model components. Nothing in the model differentiates which exponential model component goes with which exponential signal component. This symmetry manifests

itself in the probability density function. For example, in the biexponential case if there is a peak in the joint posterior probability function for the decay rate constants at ($\alpha_1 = \hat{\alpha}_1$, $\alpha_2 = \hat{\alpha}_2$), then there is an identical peak at ($\alpha_1 = \hat{\alpha}_2$, $\alpha_2 = \hat{\alpha}_1$). To break this symmetry we order the decay rate constants as follows ($L \leq \alpha_1 < \alpha_2 \leq H$). This condition leaves a single unique global maximum in the joint posterior probability.

REFERENCES

1. Brown SP, Wimperis S. 1994. NMR measurement of spin-3/2 transverse relaxation in an inhomogeneous B_1 field. Chem Phys Lett 224:508–516.
2. Narayanan A, Hartman JS, Bain AD. 1995. Characterizing nonexponential spin-lattice relaxation in solid-state NMR by fitting to the stretched exponential. J Magn Reson A 112:58–65.
3. Schaefer J, Sefcik MD, Stejskal EO, McKay RA, Dixon WT, Cais RE. 1984. Molecular motion in glassy polystyrenes. Macromolecules 17:1107–1118.
4. Garbow JR, Schaefer J. 1987. Effect of annealing on the main-chain motions of poly(butylene terephthalate). Macromolecules 20:819–822.
5. Woessner DE, Bansal N. 1998. Temporal characteristics of NMR signals from spin 3/2 nuclei of incompletely disordered systems. J Magn Reson 133:21–35.
6. Knight-Scott J, Farace E, Simnad VI, Siragy HM, Manning CA. 2002. Constrained modeling for spectroscopic measurement of bi-exponential spin-lattice relaxation of water in vivo. Magn Reson Imaging 20:681–689.
7. Asllani I, Shankland E, Pratum T, Kushmerick M. 2001. Double quantum filtered ^{1}H NMR spectroscopy enables quantification of lactate in muscle. J Magn Reson 152:195–202.
8. Venkatasubramanian PN, Shen Y-J, Wyrwicz AM. 1995. In vivo ^{19}F NMR spectroscopic study of halothane uptake in rabbit brain. Biochimica et Biophysica Acta 1245:262–268.
9. Seo Y, Murakami M, Suzuki E, Kuki S, Nagayama K, Watari H. 1990. NMR characteristics of intracellular K in the rat salivary gland: a ^{39}K NMR study using double-quantum filtering. Biochemistry 29:599–603.
10. Cohen Y, Assaf Y. 2002. High b-value q-space analyzed diffusion-weighted MRS and MRI in neuronal tissues—a technical review. NMR Biomed 15:516–542.
11. Beaulieu C, Fenrich FR, Allen PS. 1998. Multicomponent water proton transverse relaxation and T_2-discriminated water diffusion in myelinated and nonmyelinated nerve. Magn Reson Imaging 16:1201–1210.
12. Istratov AA, Vyvenko OF. 1999. Exponential analysis in physical phenomena. Rev Sci Instrum 70:1233–1257.
13. Holmstrom K, Petersson J. 2002. A review of the parameter estimation problem of fitting positive exponential sums to empirical data. Appl Math Comput 126:31–36.
14. Bromage GE. 1983. A quantification of the hazards of fitting sums of exponentials to noisy data. Comput Phys Commun 30:229–233.
15. Shrager RI, Hendler RW. 1998. Some pitfalls in curve-fitting and how to avoid them: a case in point. J Biochem Biophys Methods 36:157–173.
16. Bretthorst GL. 2005. How accurately can parameters from exponential models be analyzed. Concepts Magn Reson Part A 27A:73–83.
17. Neil JJ, Bretthorst GL. 1993. On the use of Bayesian probability theory for analysis of exponential decay data: an example taken from intravoxel incoherent motion experiments. Magn Reson Med 29:642–647.
18. Bayes Rev. T. 1763. An essay toward solving a problem in the doctrine of chances. Philos Trans R Soc London 53:370–418; reprinted in 1958. Biometrika 45:293–315; and 1963. Facsimiles of two papers by Bayes with commentary by Deming WE. New York: Hafner.
19. Jaynes ET. 2003. Probability theory: the logic of science. Bretthorst GL, editor. Cambridge: Cambridge University Press.
20. Bretthorst GL, Hutton WC, Garbow JR, Ackerman JJH. 2005. Exponential model selection (in NMR) using Bayesian probability theory. Concepts Magn Reson A, in this issue.
21. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953. Equations of state calculations by fast computing machines, J Chem Phys 21:1087–1091.
22. Gilks WR, Richardson S, Spiegelhalter DJ. 1996. Markov chain Monte Carlo in practice. London: Chapman & Hall.
23. Neal RM. 1993. Probabilistic inference using Markov chain Monte Carlo methods. Technical Report CRG-TR-93-1. University of Toronto Department of Computer Science, Toronto, Ontario.
24. Garwood M, DelaBarre L. 2001. The return of the frequency sweep: designing adiabatic pulses for contemporary NMR. J Magn Reson 153:155–177.
25. Perman WH, Thomasson DM, Bernstein MA, Turski PA. 1989. Multiple short-echo (2.5-ms) quantitation of in-vivo sodium T_2 relaxation. Magn Reson Imag 9:153–160 .
26. Hashimoto T, Ikehira H, Fukuda H, Yamaura A, Watanabe O, Tateno Y, et al. 1991. In vivo sodium-23 MRI in brain tumors: evaluation of preliminary clinical experience. Am J Physiol Imaging 6:74–80.
27. Kohler S, Preibisch C, Nittka M, Haase A. 2001. Fast three-dimensional sodium imaging of human brain. Magma 13:63–69.
28. Ra JB, Hilal SK, Cho ZH. 1986. A method for in vivo MR imaging of the short T_2 component of sodium-23. Magn Reson Med 3:296–302.
29. Hirai H, Yamasaki K, Kidena H, Kono M. 1992. Quantitative analysis of sodium fast and slow component in in-vivo human brain tissue using MR Na image. Kaku Igaku [Jpn J Nucl Med] 29:1447–1454.
30. Jefferys H. 1983. Theory of probability. Oxford: Clarendon.
31. Bretthorst GL. 1999. The near-irrelevance of sampling

frequency distributions. In: von der Linden W, Dose V, Fischer R, Preuss R, editors. Maximum entropy and Bayesian methods, vol. 105. Dordrecht: Kluwer p 21–46.

BIOGRAPHIES

G. Larry Bretthorst is a senior scientist at Washington University. He has written extensively on parameter estimation and model selection using Bayesian probability theory, including a book titled *Bayesian Spectrum Analysis and Parameter Estimation* and more recently as editor of *Probability Theory: The Logic of Science*.

William C. Hutton received his B.A. in chemistry from Maryville College in Tennessee in 1971. In 1972, he joined the staff at the University of Virginia in the Department of Chemistry. At Virginia, he conducted research in phospholipid membrane biophysics, metal binding in biological molecules, and natural product structure elucidation. In 1983, Mr. Hutton joined Monsanto Company. At Monsanto he conducted research in life sciences NMR and laboratory automation. In 1994 he was appointed a Monsanto Fellow. Mr. Hutton joined the Washington University Medical School in 2003, where he is currently a research associate.

Joel R. Garbow, Ph.D., is a senior scientist in chemistry and radiology working in the Biomedical MR Laboratory at Washington University in St. Louis. Dr. Garbow received his undergraduate degree from the University of Illinois, Urbana-Champaign, and his Ph.D. in chemistry from the University of California, Berkeley. He joined Washington University in 2000 following a 17-year career at Monsanto Company where he directed a solid-state NMR spectroscopy-based research program. At Washington University, his current research interests include the development of MR imaging techniques and their application to the study of small-animal models of cancer.

Joseph J.H. Ackerman, Ph.D., is the William Greenleaf Eliot Professor of Chemistry and chair of the Department of Chemistry at Washington University in St. Louis where he also holds appointments as professor of radiology and research professor of chemistry in medicine. Dr. Ackerman and collaborators pursue research on the development and application of magnetic resonance spectroscopy and imaging for study of intact biological systems, from single cells to genetically engineered mice to man. A primary focus is the use of water diffusion-sensitive MR methods to probe tissue architecture and microstructure at the micron length scale, far less than the actual voxel resolution of the image itself. Dr. Ackerman is a Fellow and Gold Medal winner of the (International) Society for Magnetic Resonance in Medicine.