JOURNAL OF MAGNETIC RESONANCE 95, 221–234 (1991)

Investigation of Analysis Techniques for Complicated NMR Relaxation Data

KENNETH P. WHITTALL

Department of Geophysics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

AND

MICHAEL J. BRONSKILL AND R. MARK HENKELMAN

Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada

Received May 10, 1990; revised May 9, 1991

Precision NMR relaxation measurements of biological tissue frequently show complex multicomponent behavior. This paper presents methods for generating information about these relaxation spectra even when the original data are nonideal. A variety of acceptable solutions with differing types of simplicity are investigated. Ways of inferring trends in relaxation spectra that are independent of the particular analysis model are illustrated with examples using linear programming techniques. The question of the number, spacing, and signal-to-noise ratio of data for optimal experiments is also addressed.

Measurements of NMR relaxation times in biological tissues have been used for a number of years to probe the organizational structure of tissue (*1*), the molecular dynamics of water in tissues (*2*), and the nature of malignant changes (*3*). Water proton relaxation in tissue occurs through spin exchange with macromolecular protons, by chemical-shift anisotropy, by restricted motion of water associated with the macromolecular–water interface, by paramagnetic ions, and by microscopic perturbation of the local magnetic field, to name just a few of the mechanisms. It is evident, therefore, that a clear understanding of tissue relaxation can be achieved only from highly precise and complete measurements.

NMR relaxation times of tissues now produce contrast in clinical magnetic resonance images but image contrast is usually interpreted only qualitatively. Greater exploitation of the diagnostic potential of relaxation times requires a better understanding of the biophysics of relaxation times in both normal and abnormal tissues as well as more quantitative and accurate methods of measuring relaxation *in vivo*. The successful accomplishment of these goals would lead to significant diagnostic and therapeutic applications.

It is the purpose of this paper to provide a practical guide for the analysis of experimental relaxation data to derive the maximum amount of specific information from any relaxation experiment.

0022-2364/91 $3.00

MATHEMATICAL BASIS

A basic equation describing the relaxation of magnetization in an NMR experiment is

$$y(t_i) = y_i = \int_{T_{\min}}^{T_{\max}} s(T)e^{-t_i/T}dT, \qquad i = 1, \ldots, N, \qquad [1]$$

where the N decay-curve data y_i are measured at times t_i, and $s(T)$ is the unknown spectrum amplitude as a function of the relaxation time T. The spectrum $s(T)$ may consist of discrete components or a continuous distribution. T represents T_2 or, with a simple transformation of the inversion-recovery data, it represents T_1. The limits on the integral of $T_{\min}$ and $T_{\max}$ are chosen to encompass the values of T expected for the physical system being analyzed.

One way to make Eq. [1] discrete for computer implementation is to assume that the spectrum is a sum of M δ functions with unknown positive weights s_j at known relaxation times T_j. That is,

$$s(T) = \sum_{j=1}^{M} s_j\delta(T - T_j). \qquad [2]$$

Typically M is chosen to lie between 100 and 200 so as not to bias the solution into a small number of relaxation times. Substituting Eq. [2] into Eq. [1] yields

$$y_i = \sum_{j=1}^{M} s_j e^{-t_i/T_j}, \qquad i = 1, \ldots, N. \qquad [3]$$

This linear system of N equations in M unknowns is solved using least-squares or linear programming algorithms to find a variety of relaxation spectra s_j, $j = 1, \ldots, M$, which fit the data.

An adequate fit is usually defined in terms of the χ^2 statistic. That is,

$$\chi^2 = \sum_{i=1}^{N} (y_i - y_i^{\mathrm{p}})^2/\sigma_i^2 \qquad [4]$$

where y_i^{p} are the data corresponding to the constructed spectrum, and σ_i are the standard deviations of the measurements y_i. The expected value of χ^2 is N and its standard deviation is $(2N)^{1/2}$. Typically, spectra with misfits less than one or two standard deviations above the expected value are classed as acceptable.

Effects which distort the data from the idealized representation given in Eq. [1] should be incorporated if they have a known form. The following sections describe two additions to the basic equation: baseline offsets and nonlinear effects.

BASELINE OFFSETS

Baseline offsets are implicit in inversion-recovery sequences and can be present in CPMG sequences for various reasons including electronic offsets, nonideal 180° pulses, and RF ringdown (*4, 5*). To account for offsets the mathematical formulation must be changed to

$$y_i = \int_{T_{min}}^{T_{max}} s(T)e^{-t_i/T}dT + B, \qquad i = 1, \ldots, N, \qquad [5]$$

where B is the constant baseline offset. Nonconstant offsets of a known mathematical form are also tractable.

A test was made using a synthetic decay curve of 100 points contaminated with zero-mean, Gaussian noise with a standard deviation equal to 0.2% of the maximum amplitude. Many new data sets were constructed by adding positive and negative offsets to these original data.

The relaxation spectra were calculated using a nonnegative least-squares (NNLS) algorithm (*6*). Figure 1 plots the least-squares misfit assuming $B = 0$ versus the actual data baseline offset. For large negative offsets the misfit is very high, reflecting the inability to describe negative data. For positive offsets the least-squares misfit rises more slowly because a peak appears near the maximum allowed relaxation time T_{max} and generates data similar to a constant positive offset. As T_{max} increases, the approximation to B improves and the misfit decreases. For this figure, the expected value of χ^2 was 100 (horizontal dashed line). Erring on the side of a positive offset enables a safer interpretation.

NONLINEAR EFFECTS

Known distortions due to nonlinearity in the electronics should be incorporated into the mathematical description. Failure to do so results in errors in peak positions and amplitudes (*7, 8*). For T_1 experiments, nonlinearity flattens the initial recovery and depresses the final stages such that fast relaxation components are diminished and slow components are shifted toward faster times. For T_2 experiments short relaxation components are diminished or eliminated. A typical representation of nonlinear effects is

$$y_i = \int_{T_{min}}^{T_{max}} s(T)[e^{-t_i/T} - \Delta(e^{-t_i/T})^3]dT, \qquad i = 1, \ldots, N, \qquad [6]$$

where Δ is the degree of nonlinearity embodied by the $(e^{-t_i/T})^3$ term.

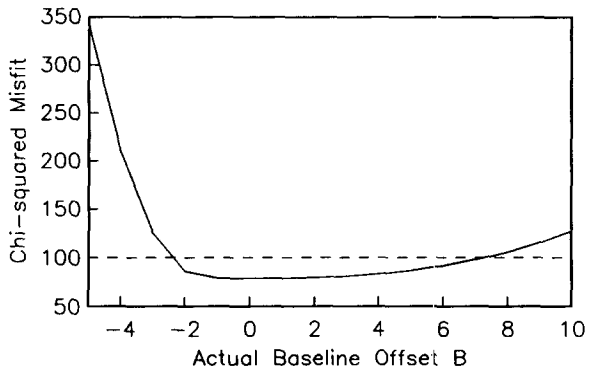

FIG. 1. Misfit assuming zero baseline. The solid line is the minimum χ^2 achievable by NNLS spectra assuming zero offset for data with actual offsets B between −5 and 10. The horizontal dashed line is the expected value of χ^2, which is 100 for these data.

Figure 2 illustrates that incorporating a description of the nonlinearity increases the reliability of the interpretation. We used 200 CPMG data from a spectrum with discrete amplitudes of 512 at $T = 0.030$ and 0.100 s (dashed lines). The data were contaminated with 0.4% noise and distorted by nonlinear effects as in Eq. [6] to produce a new data set with $\Delta = 0.20$. Figure 2a shows the corresponding NNLS spectrum (solid lines) derived from these new data without allowing for the nonlinearity. Increasing the nonlinear distortion of the data increases the χ^2 misfit, and the peaks shift to longer times before eventually merging. Figure 2b shows the NNLS spectrum obtained after incorporating the correct degree of nonlinearity. The χ^2 misfit is acceptable and the peak positions and amplitudes are well determined (given the noisy data).

If the actual degree of nonlinearity is not known then a series of Δ values may be tested. For Δ less than the true value, peaks are shifted to longer times and χ^2 is slightly larger than its expected value. As Δ increases beyond the true value of 0.20 the peak positions remain well reproduced but their amplitudes become progressively more distorted. In addition, spurious peaks begin to appear at short times although

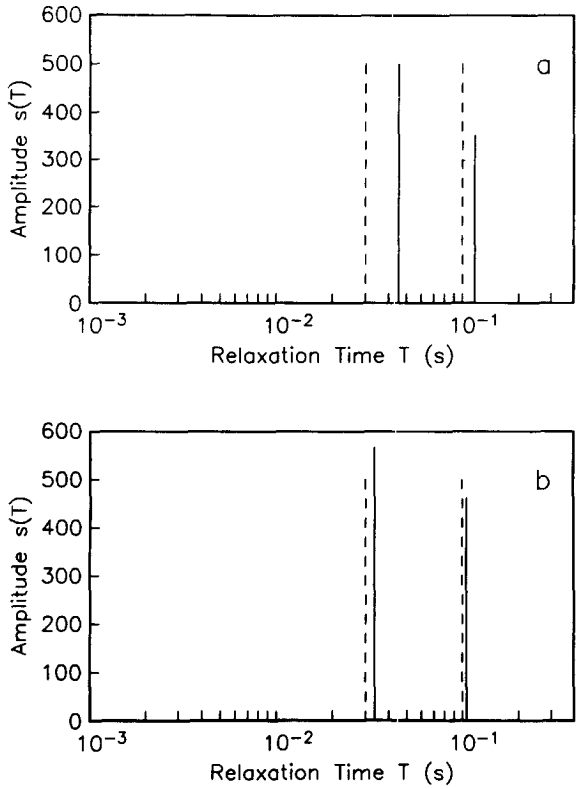

FIG. 2. (a) A NNLS spectrum (solid lines) derived from nonlinearly distorted data ($\Delta = 0.20$) without allowing for the nonlinearity. The dashed lines represent the true spectrum. (b) A NNLS spectrum incorporating the nonlinear effects (solid lines). The noisy data prohibit an exact recovery of the spike amplitudes and positions; however, the derived spectra and their misfits are improved.

the χ^2 misfit remains at an acceptable level. Thus, overestimating the degree of nonlinearity results in an acceptable solution but not one which resembles the true spectrum.

Numerical schemes which attempt to determine a spectrum and Δ automatically such that χ^2 is minimized are doomed to fail because χ^2 decreases monotonically as Δ increases. The marginally better fit achieved by setting Δ larger than its true value may be due to the extra degree of flexibility to represent the noisy data provided by the $(e^{-t_i/T})^3$ term. Thus nonlinearity in the amplitude of the data can be handled explicitly to achieve accurate relaxation spectra if the nonlinearity is known. However, attempts to estimate nonlinearities give unreliable results unless the nonlinearity is extreme.

MINIMUM-STRUCTURE SOLUTIONS

In practice, a finite number of noisy relaxation data cannot uniquely determine $s(T)$. Many solutions exist which adequately fit the data. Moreover, the possible forms of these acceptable spectra are very different, ranging from δ function representations, to piecewise-constant models, to smooth profiles. This is important because the correct underlying explanation for tissue relaxation cannot be determined uniquely from relaxation measurements alone (*6, 8*). Models of tissue relaxation based on several identifiable groups of spins which lead to discrete relaxation spectra cannot be distinguished experimentally from models based on continuous size distributions of compartments which lead to continuous spectra. Because of this nonuniqueness, it is essential to analyze relaxation data using algorithms which fit the data with diverse spectra. Each solution reveals different aspects of the data so that more complete interpretations are possible. Our philosophy is to construct a variety of minimum-structure solutions which are likely to have the essential features required by the data and be consistent with physical reality. Minimum structure may be defined in several ways. Three examples are a few isolated δ functions, a piecewise-constant spectrum with few changes in level, or a smooth curve with few oscillations. This is the principle of parsimony (*9–12*).

The minimum-misfit algorithms generate spectra composed of a few isolated δ functions which correspond to the first definition of minimum structure. Sometimes it is appropriate to relax the fit from the minimum possible to a larger misfit near the expected value. A linear programming (LP) algorithm (*6*) is able to do this while still generating spectra composed of a few δ functions.

The second definition of minimum structure is realized by using LP to minimize the one norm of the spectrum variation (*13*). That is, the algorithm finds that spectrum which minimizes

$$\sum_{j=1}^{M-1} |s_{j+1} - s_j| \qquad [7]$$

subject to fitting the data constraints. Figure 3a (solid line) gives the minimum one-norm variation spectrum derived using 100 noisy data corresponding to the true spectrum given by the dashed lines. The one norm allows large jumps locally, but the total variation is the minimum possible for the misfit level achieved. Allowing a larger misfit results in even less variation.

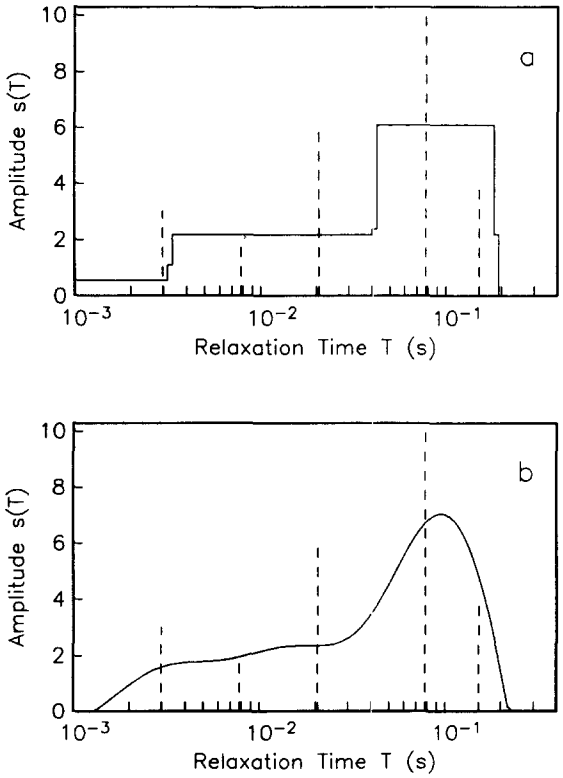

FIG. 3. The solid lines in (a) and (b) are the spectra corresponding to the minimum one-norm and minimum two-norm variations, respectively. The dashed lines are the true spectra. The χ^2 misfits are both near the expected value of 100 despite the very different forms of the calculated spectra. The small steps on the flanks of the spectrum in (a) are artifacts of the finite partition of the relaxation-time axis.

The third minimum structure definition is achieved by adding extra constraints to the basic NNLS formulation (*6*). One version of these constraints requires a joint minimization of the χ^2 misfit plus the two-norm variation

$$\sum_{j=1}^{M-1} |s_{j+1} - s_j|^2. \qquad [8]$$

These constraints discriminate more strongly against large jumps in the spectrum than the one-norm version (Eq. [7]). The result is the very smooth profile shown in Fig. 3b (solid line). This solution has an acceptable misfit which is nearly the same as that in Fig. 3a. Figure 3 illustrates the nonuniqueness by showing that broad distributions of relaxation times fit data from spectra composed of a few δ functions.

INFERENCE

The linear programming and least-squares tools described in the previous sections provide a flexible framework for the interpreter to design algorithms to return useful information from the data. In this section, we explore some possible design strategies.

Diverse spectra may fit the same noisy data equally well. This inherent nonuniqueness cannot be overcome using methods which construct only a single solution because the space of acceptable solutions is infinite in extent. Since an exhaustive search is impossible, it is preferable to infer information about features or properties common to all acceptable spectra. This approach usually generates extreme values of some property along with the spectrum that achieves this extreme. For a particular misfit level, all other acceptable solutions have values bounded by this extreme.

The best-fit solutions can be thought of as examples of this technique. All other spectra must have a larger misfit. These solutions are important because they determine whether any spectrum which adequately fits the data exists.

The minimum-structure solutions in Fig. 3 are extreme in the sense that every other spectrum must have more one-norm or two-norm variation for the given misfit level. They are not extreme in the sense of barely fitting the data. Minimum-structure solutions are important because features on these spectra are likely required by the data and not simply artifacts of the noise or the construction algorithm. They are conservative representations of the data.

Linear programming is a powerful tool for generating bounds on any linear combination of the spectrum amplitudes, s_j, by minimizing or maximizing the objective function

$$\phi = \sum_{j=1}^{M} w_j s_j \qquad [9]$$

subject to fitting the data. The weights w_j are chosen such that in combination with s_j they represent the property to be made extreme.

Figure 4a shows an application of the LP inference approach. In this case, the signal is maximized within a series of narrow, nonoverlapping intervals, subject to achieving an acceptable misfit level. The maximum value as a percentage of the total spectrum integral is plotted as a horizontal bar spanning each interval. The maxima themselves do not form a solution which fits the data because each one is realized by a different spectrum; however, the maxima closely follow the form of the true spectrum percentages (dashed lines). Note that this information is independent of any construction algorithm and is information about the entire space of acceptable solutions. Hence, the true solution (if it does not radically misfit the noisy data) must have values within these extremes. An example inference is that all spectra (with a one-norm misfit less than that used in generating the figure) must have less than 31% of their signal within the interval $0.0185 \leqslant T \leqslant 0.0216$ s. Examining a single spectrum from a particular construction algorithm does not permit rigorous inferences of this type.

The corresponding lower bound over such narrow intervals is zero because the signal is easily moved just outside the region without significantly changing the misfit. However, as the intervals widen, larger shifts become more difficult and eventually some signal must appear within the interval. For these data, a region near 0.100 s is the first to have a nonzero minimum. This is consistent with the form of the true spectrum, which has most of its signal near 0.100 s. For example, when the range is from 0.040 to 0.185 s, the minimum signal is 38% of the total. That is, these data do not permit any spectrum with a similar or smaller misfit to have less than 38% of its signal over this range. The true spectrum has 56% of its signal within this range.

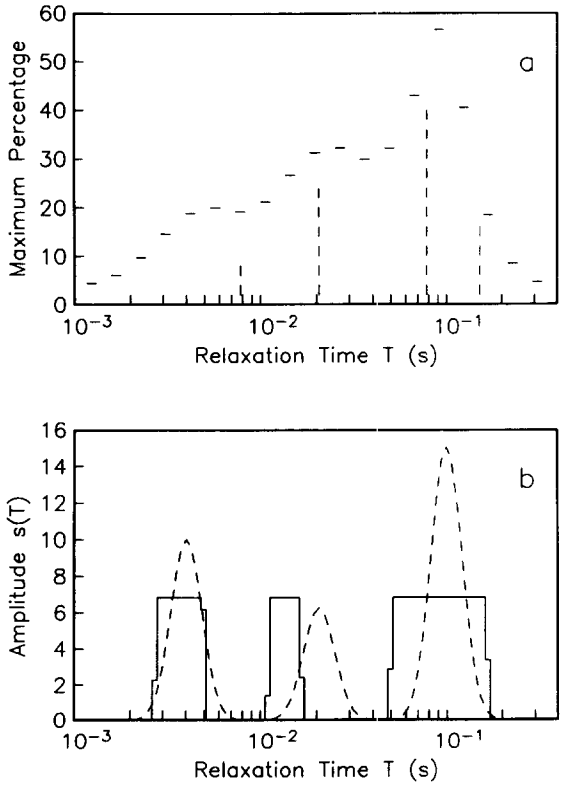

FIG. 4. In (a) the horizontal bars define a series of narrow intervals over which the spectrum amplitudes were maximized. The height of the bar is the resulting maximum percentage of signal in that interval. The dashed lines are the percentages of the components of the true spectrum. In (b) the least-upper-bound spectrum is plotted with solid lines and the true spectrum with dashed lines. The small steps on the flanks of the spectra are artifacts of the finite partition of the relaxation axis.

The least-upper-bound solution is a second example of an inference using LP. In this case, subject to fitting the data and the extra constraints

$$s_j \leq S, \qquad j = 1, \ldots, M, \qquad [10]$$

the LP algorithm minimizes the objective function

$$\phi = S, \qquad [11]$$

where S is an additional variable. The result is the least upper bound S and a spectrum which achieves this bound. Figure 4b shows an example of this approach. The true spectrum is plotted as dashed lines. This smooth, continuous spectrum was selected to be representative of distributions of relaxation times arising from tissues with a wide range of proton environments. The inference is that all other acceptable spectra must meet or exceed the least upper bound S somewhere within the given interval from 0.001 to 0.400 s. Note that these spectra have quite an unusual form. They are

essentially either zero or the bound S. The minor steps on the sides of the blocks are due to the finite partition of the T axis. Yet these block-like spectra fit the data well, with χ^2 equal to its expected value.

One final example of an inference using LP follows the observation by Kroeker *et al.* (*14*) that spectrum peaks from KHT tumors in mice redistribute toward higher relaxation times as the tumors grow. The first moment quantifies this shift and in discrete form it is

$$\sum_{j=1}^{M} T_j s_j. \qquad [12]$$

This is equivalent to the LP objective function of Eq. [9] with weights $w_j = T_j$. The first moment of individual peaks within a spectrum is calculated by altering the upper and lower limits of the summation over j. Minimizing and maximizing the objective function [12] subject to the data constraints gives the extremes allowed by the data. Comparing the maximum first moment at one tumor weight with the minimum first moment at a greater weight shows whether the increase in first moment is rigorously required by the data. Figure 5 gives an example of this method applied to noisy data

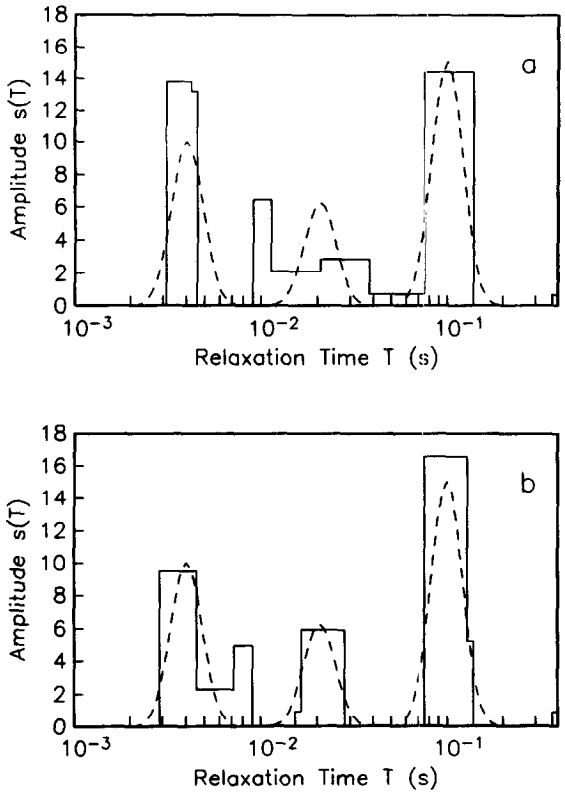

FIG. 5. The solid lines in (a) and (b) are spectra with minimum and maximum first moment, respectively, within an interval from 0.001 to 0.009 s. The dashed line is the true spectrum.

from the three-peak test spectrum shown as dashed lines. An upper bound was imposed on the one-norm variation of the solution (Eq. [7]) to give the blocky spectra shown as solid lines. An interval from 0.001 to 0.009 s was chosen to encompass the first peak. The minimization result in Fig. 5a shows a shift to shorter times with respect to the true spectrum. The maximization result in Fig. 5b shows a redistribution to longer times. The average relaxation time of a peak is given by

$$T_{\text{ave}} = \sum_{j=j_{\text{lo}}}^{j_{\text{hi}}} T_j s_j \Big/ \sum_{j=j_{\text{lo}}}^{j_{\text{hi}}} s_j. \qquad [13]$$

For the true spectrum $T_{\text{ave}} = 4.1$ ms. The minimum and maximum first-moment spectra in Fig. 5 have $T_{\text{ave}} = 3.6$ and 4.6 ms, respectively. These extreme values can be compared with other relaxation times of individual probes to see if systematic change with tumor progression is resolvable. Such strategies also cope with the biological heterogeneity inherent in such systems.

EXPERIMENTAL DESIGN

The resolution and accuracy of spectra derived from relaxation data improve with the number of points, N, and the signal-to-noise ratio, SNR. The ideal experiment yields an infinite number of precise data so that $s(T)$ is uniquely determined; however, compromises must be made in practice. For T_1 experiments especially, there are usually strict time constraints which prohibit sampling a large number of points. In this section, we examine how to choose N, SNR, and the measurement times t_i to maximize the information in the data. Backus and Gilbert (*15, 16*) presented a method which is suited to this analysis. Parker (*17*) and Oldenburg (*18*) reviewed the Backus–Gilbert (BG) method.

Consider a linear combination of both sides of Eq. [1],

$$\sum_{i=1}^{N} a_i(T_0) y_i = \sum_{i=1}^{N} a_i(T_0) \int_{T_{\min}}^{T_{\max}} s(T) e^{-t_i/T} dT$$

$$= \int_{T_{\min}}^{T_{\max}} s(T) \sum_{i=1}^{N} a_i(T_0) e^{-t_i/T} dT, \qquad [14]$$

where $a_i(T_0)$ are coefficients for a particular $T = T_0$. Define the averaging function $A(T, T_0)$ as a linear combination of the exponentials

$$A(T, T_0) = \sum_{i=1}^{N} a_i(T_0) e^{-t_i/T}. \qquad [15]$$

Substituting Eq. [15] into Eq. [14] gives

$$\langle s(T_0) \rangle = \int_{T_{\min}}^{T_{\max}} s(T) A(T, T_0) dT, \qquad [16]$$

where

$$\langle s(T_0) \rangle = \sum_{i=1}^{N} a_i(T_0) y_i \qquad [17]$$

is an average of the spectrum near $T = T_0$. For accurate data these averages are unique for a given set of $a_i(T_0)$ because every solution reproducing the data, y_i, has the same linear combination in Eq. [17]. Thus the averaging function is a window through which all acceptable spectra are viewed, and the average value is information common to these spectra and independent of any construction algorithm. For inaccurate data, the averages have a known variance derived from the data covariance matrix.

The BG method achieves a tradeoff between the averaging function resolution and the variance of the average value by finding those $a_i(T_0)$ which minimize

$$\psi(T_0) = V(T_0) + \mu W(T_0), \qquad [18]$$

where

$$W(T_0) = \int_{T_{\min}}^{T_0} \left(\frac{T_0}{T}\right)^8 A^2(T, T_0)dT + \int_{T_0}^{T_{\max}} \left(\frac{T}{T_0}\right)^8 A^2(T, T_0)dT \qquad [19]$$

is a measure of the averaging function width,

$$V(T_0) = \sum_{i=1}^{N} a_i^2(T_0)\sigma_i^2 \qquad [20]$$

is the variance of the average estimate, and μ is a positive tradeoff parameter. When $\mu = 0$ the variance is minimized at the expense of the resolution. When μ is large the resolution width is small but the variance is increased. The definition of $W(T_0)$ used here is a modification of the spread criterion of (*15*). We chose this form because it usually produces nonnegative averaging functions centered on T_0 and symmetric on a logarithmic scale. The resolving width of the averaging function is defined as the ratio of the relaxation time at the upper half-maximum point to that at the lower.

Figure 6a shows three typical averaging functions for a relative error in the average estimate of 2%. The locations are $T_0 = 0.004$, 0.020, and 0.100 s. These functions are rather wide, suggesting that components separated by less than a ratio of about 4 are not fully resolved by the 100 noisy, logarithmically spaced data. The three averaging functions from left to right have widths of 4.8, 5.8, and 4.4. However, these averaging functions are wide partly because no positivity constraints on the spectrum can be applied. Hence, this extra information cannot be used to improve the resolution.

Figure 6b shows the effect on resolution of a different data sampling. The decay curve was resampled with a linear instead of logarithmic spacing over the same interval. The three new averaging functions now have widths of 10.1, 6.4, and 4.1, for the same 2% relative error in the average estimate. Resolution of fast relaxing components near 0.004 s is degraded because of the relatively few data at short times. Resolution of the long components near 0.100 s is enhanced because more decay-curve data now fall near this time. In general, to resolve a component at a particular relaxation time, this method shows that the decay curve must be sampled often at measurement times near the target relaxation time.

The effects of various N and SNR on the resolution can be assessed by a tradeoff curve. A tradeoff curve at a particular T_0 is formed by plotting the standard deviations of the averages versus the widths of the averaging functions obtained by minimizing Eq. [18] for many values of μ. This calculation requires only the sampling times t_i

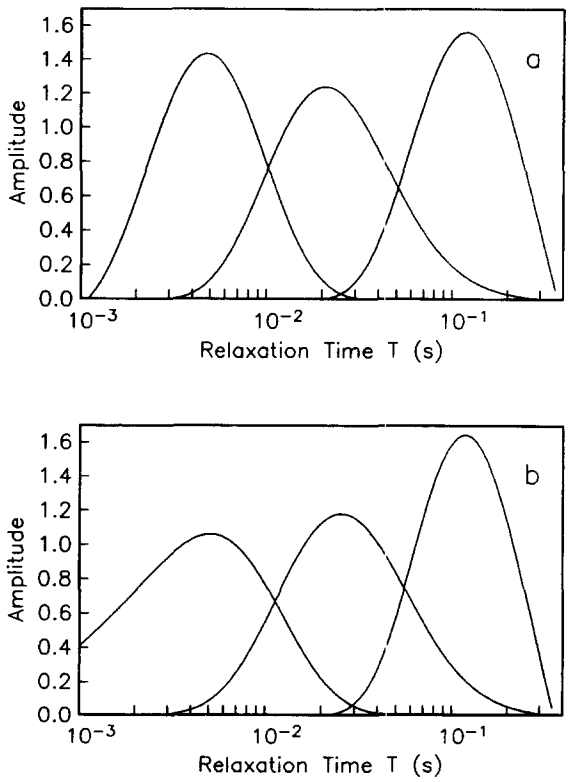

FIG. 6. (a) The three averaging functions are derived from 100 logarithmically spaced decay-curve data covering the interval 0.001 to 0.500 s. Their locations T_0 are 0.004, 0.020, and 0.100 s, and their ratios of upper to lower half-maximum point are 4.8, 5.8, and 4.4. (b) The three averaging functions are derived from 100 linearly spaced data covering the same interval. Their locations are also the same. However, the resolving widths are now 10.1, 6.4, and 4.1.

and the data standard deviations σ_i. Thus, the tradeoff curve is independent of the actual data y_i and so the analysis is general rather than specific to a particular set of measurements.

We generated data sets with 30, 60, and 120 points spaced logarithmically between 0.001 and 0.500 s. For each of these, the SNR was set at 250, 125, and 62.5. Thus there are nine data sets in total. The SNR is defined as the ratio of the decay-curve maximum at zero time to the standard deviation of the added noise. Five of the nine tradeoff curves for $T_0 = 0.020$ s are plotted in Fig. 7a. Curves closer to the lower left corner indicate better resolution because, at a given standard deviation of the average estimate, the width is smaller. As expected, increasing the number of points at a particular SNR (top three curves) or increasing the SNR for a fixed number of points (bottom three curves) improves the resolution. Figure 7b shows two pairs of curves (indistinguishable) which indicate a rule for choosing the number of data N and the SNR. The top pair corresponds to 30 points with a SNR of 125 and 120 points with a SNR of 62.5. The lower two curves result from 30 points with SNR = 250 and 120

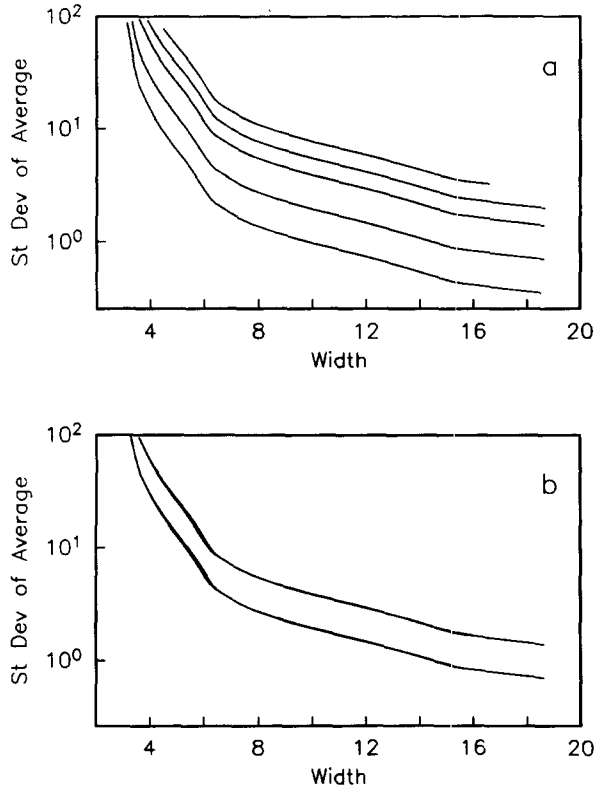

FIG. 7. The tradeoff curves in (a) and (b) quantify the decrease in standard deviation of the average estimate with the increase in width of the averaging function. (a) The curves from top to bottom correspond to 30 data with SNR = 62.5, 60 data with SNR = 62.5, 120 data with SNR = 62.5, 120 data with SNR = 125, and 120 data with SNR = 250. (b) The top curve corresponds to 30 data with SNR = 125, as well as 120 data with SNR = 62.5. The lower curve is 30 data with SNR = 250 and 120 data with SNR = 125. These two pairs of curves overlap and suggest that the resolving power follows the product SNR $\sqrt{N}$.

points with SNR = 125. The overlap of the individual pairs suggests that the resolving power follows the product

$$\text{SNR}\sqrt{N}, \qquad [21]$$

with higher values indicating better resolution. Hence, to achieve the same resolution at half the SNR it is necessary to measure four times as many points. This rule is intuitively obvious for monoexponential data, where the decay curve is a straight line on a semilogarithmic plot. Measuring at four times as many locations is equivalent to repeating the measurements at the original locations four times and so doubling the SNR. Because the BG method is independent of the actual data values, it demonstrates the general applicability of Eq. [21]. This result aids the design of experiments where time constraints may be different for different SNR and N.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the Natural Sciences and Engineering Research Council of Canada, the Medical Research Council of Canada, and the National Cancer Institute of Canada.

REFERENCES

1. P. S. BELTON AND R. G. RATCLIFFE, *Magn. Reson. Med.* **17,** 241 (1985).
2. J. ZHONG, J. C. GORE, AND I. M. ARMITAGE, *Magn. Reson. Med.* **21,** 295 (1989).
3. W. R. INCH, J. A. MCCREDIE, R. R. KNISPEL, R. T. THOMPSON, AND M. M. PINTAR, *J. Natl. Cancer Inst.* **52,** 353 (1974).
4. R. L. VOLD, R. R. VOLD, AND H. E. SIMON, *J. Magn. Reson.* **11,** 283 (1973).
5. D. G. HUGHES AND G. LINDBLOM, *J. Magn. Reson.* **26,** 459 (1977).
6. K. P. WHITTALL AND A. L. MACKAY, *J. Magn. Reson.* **84,** 134 (1989).
7. R. J. S. BROWN, *J. Magn. Reson.* **82,** 539 (1989).
8. C. H. NEWCOMB, S. J. GRAHAM, AND M. J. BRONSKILL, *J. Magn. Reson.* **90,** 279 (1989).
9. S. W. PROVENCHER, *Comput. Phys. Commun.* **27,** 213 (1982).
10. OCCAM, WILLIAM OF, Summa Totius Logical, Paris, 1488.
11. S. W. PROVENCHER, *Comput. Phys. Commun.* **27,** 229 (1982).
12. S. C. CONSTABLE, R. L. PARKER, AND C. G. CONSTABLE, *Geophysics* **52,** 289 (1987).
13. S. E. DOSSO AND D. W. OLDENBURG, *Geophys. J. Int.* **99,** 483 (1989).
14. R. M. KROEKER, C. A. STEWART, M. J. BRONSKILL, AND R. M. HENKELMAN, *Magn. Reson. Med.* **6,** 24 (1988).
15. G. E. BACKUS AND J. F. GILBERT, *Geophys. J. R. Astron. Soc.* **16,** 169 (1968).
16. G. E. BACKUS AND J. F. GILBERT, *Phil. Trans. R. Soc. A* **266,** 123 (1970).
17. R. L. PARKER, *Annu. Rev. Earth Planet. Sci.* **5,** 35 (1977).
18. D. W. OLDENBURG, *IEEE Trans. Geosci. Remote Sensing* **GE-22,** 664 (1984).