JOURNAL OF MAGNETIC RESONANCE, Series A **120,** 155–189 (1996)
ARTICLE NO. 0113

Nonlinear-Least-Squares Analysis of Slow-Motion EPR Spectra in One and Two Dimensions Using a Modified Levenberg–Marquardt Algorithm

DAVID E. BUDIL,* SANGHYUK LEE,† SUNIL SAXENA, AND JACK H. FREED

Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853

Received September 6, 1995; revised February 14, 1996

The application of the "model trust region" modification of the Levenberg–Marquardt minimization algorithm to the analysis of one-dimensional CW EPR and multidimensional Fourier-transform (FT) EPR spectra especially in the slow-motion regime is described. The dynamic parameters describing the slow motion are obtained from least-squares fitting of model calculations based on the stochastic Liouville equation (SLE) to experimental spectra. The trust-region approach is inherently more efficient than the standard Levenberg–Marquardt algorithm, and the efficiency of the procedure may be further increased by a separation-of-variables method in which a subset of fitting parameters is independently minimized at each iteration, thus reducing the number of parameters to be fitted by nonlinear least squares. A particularly useful application of this method occurs in the fitting of multicomponent spectra, for which it is possible to obtain the relative population of each component by the separation-of-variables method. These advantages, combined with recent improvements in the computational methods used to solve the SLE, have led to an order-of-magnitude reduction in computing time, and have made it possible to carry out interactive, real-time fitting on a laboratory workstation with a graphical interface. Examples of fits to experimental data will be given, including multicomponent CW EPR spectra as well as two- and three-dimensional FT EPR spectra. Emphasis is placed on the analytic information available from the partial derivatives utilized in the algorithm, and how it may be used to estimate the condition and uniqueness of the fit, as well as to estimate confidence limits for the parameters in certain cases.

INTRODUCTION

In recent years, there has been a dramatic increase in the power and flexibility of electron paramagnetic resonance methods, mostly as a result of significant advances in EPR instrumentation. Two of the most important advances include the introduction of two-dimensional Fourier-transform methods (*1–4*), and the extension of both continuous-wave (*5, 6, 7*) and pulsed (*8, 9*) EPR methods to high frequencies requiring superconducting magnets. By analogy with similar developments in NMR, these enhancements have greatly increased the amount of structural and dynamic information that is available from spectra of spin labels and intrinsic paramagnetic species.

One important application of EPR, and magnetic resonance in general, is to study the molecular dynamics of an appropriate label in isotropic fluids, ordered phases such as liquid crystals or biological membranes, on surfaces, or attached to macromolecules. Many details of the dynamics are discernible in the slow-motion regime (*10*), where the characteristic time scale of the motion is on the order of the inverse spectral bandwidth. However, the analysis of slow-motion spectra is complicated by the fact that the relationship between the spectrum and the physical parameters of interest is rather indirect. The partial averaging of EPR spectra by molecular motion or spin dynamics can produce very complicated and irregular lineshapes requiring detailed spectral simulation to extract the desired information.

The increase in the number of spectral dimensions and the resolution now available in EPR has been accompanied by a corresponding increase in the computational demands of spectral simulation. The computational requirements are particularly stringent for slow-motion problems, which typically require the solution of large matrix equations (*11–13*) for the calculation of a single spectrum.

Although the literature on computer simulation of EPR spectra is quite extensive (*14*), most of the methods for fitting simulations to experimental EPR spectra have been applied either in the motionally narrowed regime or in the rigid limit. Only relatively recently have nonlinear-least-squares methods been used to fit slow-motion EPR spectra (*15–22*). One feature of slow-motion spectra that makes them particularly problematic to fit is the much larger num-

* Present address: Department of Chemistry, Northeastern University, Boston, Massachusetts 02115.

† Present address: Department of Chemistry, Ewha Womens University, Seoul, Korea, 120-750.

1064-1858/96 $18.00

ber of parameters that may in principle enter into the calculation of the spectrum relative to motionally narrowed or rigid limit spectra. Aside from inhomogeneous broadening, which can be present in any motion regime, fast-motion spectra may be fitted with just the isotropic magnetic parameters and the Lorentzian linewidths which relate to the rotation of the probe. In the specific case of ^{14}N nitroxides, there are just three linewidth parameters which can relate to the rotational diffusion rate. At the other extreme, rigid limit spectra may require all of the magnetic tensor components, but do not require any dynamic parameters. In contrast, slow-motion spectra generally require full information about the magnetic tensors and rotational dynamics, which frequently entail additional parameters to characterize the motional model or molecular ordering.

The large number of variable parameters and the computational power required for fitting slow-motion spectra limit the choices from among the many algorithms that are available for nonlinear-least-squares minimization. One of the more popular methods in widespread use is the downhill simplex method, which has been applied to fit rigid-limit (powder) samples by Duling (*23*) and Fajer *et al.* (*24*), who also outline the use of that method for analyzing distributions of oriented assemblies, and multiexponential kinetics in EPR. For general application to slow-motion problems, however, this method becomes unwieldy because of the relatively large number of function evaluations it requires.

Among the more highly recommended methods for computation-intensive minimization problems are the so-called quasi-Newton methods, which include the Levenberg–Marquardt algorithm and its various modifications. These methods have the drawback that they require the partial derivatives of the spectrum with respect to each of the nonlinear parameters at each iteration of the procedure. For functions such as the slow-motion EPR spectral intensity, which have no simple analytical expression, each partial derivative must be calculated by a forward-difference approximation, which requires a recalculation of the spectrum with a small step in each parameter. To balance this additional computational burden, however, is the very useful feature that the partial derivative information may be used to estimate the condition of the fit at each iteration of the minimization procedure. Such a diagnostic capability allows one to improve the overall efficiency and reliability of the fit by eliminating indeterminate parameters or by imposing additional constraints amongst them. This is a particularly critical advantage for models with a large number of variable parameters, which may not all be measurable from a given experimental data set. Other advantages of quasi-Newton methods include global convergence, and more flexible control over convergence criteria, which, when properly applied, allows excessive iterations to be avoided. Finally, under appropriate conditions, the partial derivative information can be used to obtain confidence intervals for the fitting parameters and to carry out statistical tests to evaluate different models for the molecular dynamics.

In this work, we present a general approach that we have developed in our studies over the past several years, for fitting both one-dimensional and two-dimensional slow-motion EPR data to spectra calculated using the stochastic Liouville equation. Several strategies for increasing the efficiency of the fitting procedure will be discussed. The first is the separation of variables into a set of nonlinear fitting parameters containing the magnetic and/or dynamic parameters of interest, and a set of separately optimizable (usually linear) parameters. The resulting reduction in the size of the parameter space that must be searched significantly enhances the speed of the overall fitting procedure. Additional efficiency is obtained with a modified version of the Marquardt–Levenberg method. We also demonstrate how, by varying different linear combinations of parameters (e.g., spherical vs Cartesian components of the magnetic and diffusion tensors) the effects of correlation amongst these parameters may be reduced. These advantages, combined with improvements in the computational methods used to solve the SLE, have made it possible to carry out interactive, real-time fitting of large-scale problems on a laboratory workstation with a graphical interface. We achieve order-of-magnitude reductions in computing time compared with earlier versions of our program (*15–18*).

Also, the earlier versions of the SLE calculation (*11–13*) have been upgraded to include a fully anisotropic rotational diffusion tensor and the specification of all three Euler angles describing the principal axes of molecular diffusion relative to the principal axes of the magnetic tensors. We also allow the specification of all three Euler angles describing the relative orientation of the A and g tensors, and include the isotropic nuclear Zeeman interaction.

The dynamic models available for slow-motion spectral simulation are summarized under Methods below, where we also briefly review the numerical methods used to calculate one- and two-dimensional spectra, describe algorithms used in the nonlinear minimization and optimization of the separated variables, and review methods for statistical analysis of EPR fitting parameters based on the Levenberg–Marquardt and related quasi-Newton methods.

Under Examples and Discussion, we demonstrate the method for selected examples of both synthetic and experimental 1D and 2D slow-motion data. Emphasis will be placed on the analytic information available from the partial derivative information utilized in the algorithm, as well as statistical analysis of the fitting parameters, and the condition of the fit. The first example illustrates the effects of fitting highly correlated variable parameters and discusses methods for reducing parameter correlation. In the second example, we demonstrate global fits to a series of spectra taken for

different values of an experimental parameter (in this case, the orientation of an ordered phase), while the third example shows the fitting of CW spectra with multiple independent spectral components having different dynamic or ordering properties and illustrates how the spectral resolution and quantification of the individual components can be assessed by statistical analysis. A method is described for estimating parameter uncertainty bounds from fits to experimental data in the usual case where the residual distribution violates the assumptions underlying the standard error expressions. Finally, we illustrate fitting of 2D-FT spectra such as COSY spectra and 2D-ELDOR spectra obtained at a series of mixing times, which are then collectively equivalent to a three-dimensional experiment. The section concludes with a description of how the programs developed in this work may be obtained.

METHODS

Slow-Motion Models

Table 1 summarizes the model parameters that may enter into a slow-motion spectral calculation. These are identified by the labels used in the interactive fitting program and are divided into four categories common to both CW and 2D-FT models: (i) *magnetic and structural parameters* that describe the magnetic tensors (including an inhomogeneous linewidth tensor) and their relative geometries; (ii) *dynamic parameters,* which include the rates of rotational diffusion for various types of motional models; (iii) *ordering potential coefficients* that describe the orienting influence of anisotropic fluids such as liquid crystals and membranes; and (iv) *integer parameters* that define the type of model or calculation used. In addition, the 2D-FT models include rotation-independent electronic and nuclear spin relaxation times that influence the 2D spectral lineshape, and the experimental delay times required to specify a 2D or 3D experiment fully.

The typical approach in applying slow-motion least-squares fitting is first to determine the magnetic parameters from rigid-limit spectra and then to fix these parameters, varying only the dynamic and ordering parameters in the fitting procedure. In fact, we have found that the slow-motion lineshape program can also be efficiently used to obtain magnetic parameters from rigid-limit spectra by performing calculations in the limit of very small rotational rates. For many applications, particularly 2D and high-field spectra, we find that the spherical harmonic expansion utilized in the slow-motion calculation is competitive with standard approaches that utilize some variation of discretized integration over the unit sphere.

Magnetic and Structural Parameters

The programs model the reorientation of a one-electron, one-nucleus system (typically a nitroxide) that is described by (a) the electron spin Zeeman interaction, including an orientation-dependent g factor; (b) the electron-nuclear hyperfine interaction tensor A, and (c) the isotropic nuclear spin Zeeman interaction.

Several different coordinate systems are employed to represent the full problem: (a) the *magnetic tensor frame* (x_m, y_m, z_m) fixed in the molecule; (b) the *rotational diffusion tensor frame* (x_R, y_R, z_R), which is also a molecule-fixed frame; (c) the *director frame* (x_d, y_d, z_d), which is a laboratory axis system relative to which the molecule orients in anisotropic fluids (cf. below); and (d) the *laboratory frame* (x_L, y_L, z_L), in which the dc magnetic field lies along z_L.

The relationships amongst these coordinate systems are specified by different sets of tilt angles. The diffusion tilt angles $\Omega_D = (\alpha_D, \beta_D, \gamma_D)$ are the Euler angles that specify the rotation taking the magnetic axes into those of the rotational diffusion tensor. The rotation is accomplished by the following series of operations: (i) a rotation about z_R through the angle γ_D; (ii) a rotation about y_R through the angle β_D; and (iii) a rotation about z_R through the angle α_D, where a positive angle produces a counterclockwise rotation when viewed along the axis of rotation looking from the positive side. In the case that the A tensor and g tensor are not coaxial, it is possible to specify a set of magnetic tilt angles $\Omega_m = (\alpha_m, \beta_m, \gamma_m)$ that rotate the A-tensor axes (x_A, y_A, z_A) into the g-tensor (magnetic) axes using the same convention. It should be noted that, according to these definitions, the α angle is not required if the tensor(s) in the reference frame is (are) axially symmetric; for example, $\alpha_D \neq 0$ has no effect if $R_x = R_y$. Finally, the director frame is related to the laboratory frame by the angles $\Psi = (0, \psi, 0)$; only one angle is required for this rotation because of the axial symmetry of the orienting potential. The director tilt angle ψ is thus the angle between the director of the anisotropic fluid and the magnetic field.

In addition to the explicit magnetic interactions of the electronic and nuclear spins, it is possible to specify two types of orientation-dependent inhomogeneous broadening: (a) a linewidth tensor W associated with the magnetic frame that specifies an additional Lorentzian linewidth, and (b) the quantities $\Delta^{(0)}$ and $\Delta^{(2)}$ that specify an added Gaussian inhomogeneous linewidth Δ that has been chosen to relate to the director frame according to the formula $\Delta = \Delta^{(0)} + \Delta^{(2)}\cos^2\psi$. (The present 2D implementation of these models does not utilize the orientation-dependent term).

Dynamic Parameters

The spectral calculation incorporates several different models for rotational diffusion, including: (a) *Brownian* rotational diffusion; (b) *non-Brownian* diffusion, including different types of jump diffusion models; (c) *anisotropic viscosity* for motion in oriented fluids; and (d) *discrete jump*

TABLE 1
List of Model Parameters Used in Nonlinear-Least-Squares Simulation of One- and Two-Dimensional EPR Spectra in the Slow-Motion Regime

Parameter symbol	Parameter Name	Description
g_{xx}, g_{yy}, g_{zz}	`gxx, gyy, gzz`	g tensor
A_{xx}, A_{yy}, A_{zz}	`axx, ayy, azz`	A (hyperfine) tensor (G)
W_{xx}, W_{yy}, W_{zz}	`wxx, wyy, wzz`	Lorentzian linewidth tensor (molecular frame, p-p width in gauss)
γ_N	`gamman`	Nuclear gyromagnetic ratio (rad s^{-1} G^{-1})
α_D, β_D, γ_D	`alphad, betad, gammad`	Diffusion tilt angles (°)
α_m, β_m, γ_m	`alpham, betam, gammam`	Magnetic tilt angles (°)
ψ	`psi`	Director tilt angle (°)
$\Delta^{(0)}$, $\Delta^{(2)}$	`gib0, gib2`	Director-frame-dependent Gaussian inhomogeneous broadening (p-p width, G)
$\{c_K^L\}$	`c20, c22, c40, c42, c44`	Coefficients for orienting potential
ϕ	`phase`	Phase of spectrum (°); +90° = pure dispersion
B_0	`b0`	Static field (G)
	Dynamic parameters	
P_l, P_{xy}, P_{zz}	`pl, pkxy, pkzz`	Log_{10} of model parameters for non-Brownian diffusion
R_{xx}, R_{yy}, R_{zz}	`rxx, ryy, rzz`	Log_{10} of rotational diffusion tensor (s^{-1})
D_j	`djf`	Log_{10} of rate for jumping among symmetry-related sites (s^{-1})
D_j, $D_{j,\perp}$	`djf, djfprp`	Log_{10} of effective diffusion tensor for anisotropic viscosity (s^{-1})
ω_{ss}	`oss`	Log_{10} of Heisenberg spin exchange rate (s^{-1})
	Integer parameters	
$2I$	`in2`	Twice the nuclear spin, I.
	`ipdf`	Model parameter (Brownian, non-Brownian, anisotropic viscosity)
n_{jump}	`ist`	Number of symmetry-related sites in discrete-jump model
m_l, m_{xy}, m_z	`ml, mxy, mzz`	Non-Brownian diffusion model flags
L_{emx}, L_{omx}, K_{mn}, K_{mx}, M_{mn}, M_{mx}, P^1_{mx}	`lemx, lomx, kmn, kmx, mmn, mmx, ipnmx`	Basis set truncation indices
n_{MOMD}	`nort`	Number of orientations in MOMD model
	Additional parameters for 2D models	
Δ_g	`gib`	Gaussian inhomogeneous broadening (`gib0` in CW program), p-p width (G)
Δ_l	`lib`	Lorentzian inhomogeneous broadening, p-p width (G)
	`hwid, mwid`	Constant and linear terms for T_{mix}-dependent broadening (Lorentzian) in 2D ELDOR, p-p width (G)
$(T_{2e}^{-1})^{(0)}$, $(T_{2e}^{-1})^{(2)}$	`t2edi, t2efi`	Log_{10} of angular dependent homogeneous T_{2e}^{-1} tensor (s^{-1})
$2W_e$	`t1edi`	Log_{10} of longitudinal electron spin-relaxation rate (s^{-1})
$2W_n$	`t1ndi`	Log_{10} of longitudinal nuclear spin-relaxation rate (s^{-1})
t_1, t_2	`init1, init2`	Initial t_1 and t_2 times in a 2D experiment (ns)
Δt_{1d}, Δt_{2d}	`stept1, stept2`	Step sizes for t_1 and t_2 in a 2D experiment (ns)
T_{mix}	`tfix`	Mixing time in a 2D ELDOR experiment (ns)
	Additional integer parameters for 2D models	
	`iexp`	Experiment type: FID, 2D SECSY, stimulated SECSY, 2D COSY, 2D ELDOR, echo-ELDOR
	`icomb`	Linear combination: S_{c+} or S_{c-}
	Levenberg–Marquardt parameters	
	`ftol`	Relative tolerance for change in residuals at convergence
	`xtol`	Relative tolerance for change in parameters at convergence
	`gtol`	Relative tolerance for change in residual gradient at convergence
	`scale`	Scaling factor for step size along each parameter dimension
	`fdstep`	Relative forward-difference step for calculating partial derivative for each parameter
	`bound`	Scale for initial trust region boundary

motion for hopping amongst a set of symmetry-related sites. Different diffusional models may be used for different molecular axes, but only when the motions occur on very different time scales.

Brownian diffusion. Brownian diffusion occurs as a series of infinitesimal reorientational steps. This model is most appropriate for intermediate-sized spin probes and spin-labeled macromolecules, and is therefore the most widely used. The program allows a fully anisotropic rotational diffusion tensor to be specified in the case of Brownian diffusion.

Non-Brownian diffusion. In non-Brownian diffusion, the probe undergoes a discontinuous step motion. The two non-Brownian models supported by the program include jump diffusion and approximate free diffusion (*10*). In jump diffusion, the molecule remains stationary for an average time τ after which it instantaneously jumps to a new orientation, specified by some angle of rotation about the specified axis. The root-mean-square reorientation angle for axis i (in radians) is $\theta_{\text{avg}} = \langle \theta^2 \rangle = 6R_i\tau_i$. In an approximate free-diffusion model, the molecule rotates freely about axis i with a rotational rate R_i, but instantaneously reorients at an average interval τ_i, and continues its free rotation.

The residence time τ_i in each of these models is specified by means of a ''model parameter'' $P_i = R_i\tau_i$. In the present implementation, the non-Brownian rotational diffusion tensor must possess at least axial symmetry; thus, i may refer only to the axes perpendicular (i = xy) or parallel (i = z) to z_R. However the model does allow for *internal* non-Brownian diffusion around z_R, in addition to the overall rotation. The residence time τ_l corresponding to this internal diffusion is specified by the model parameter P_l using the relation $P_l = R_\perp\tau_l$. This is useful for describing non-Brownian behavior of a spin probe moiety that is flexibly attached to a larger diffusing molecule (*25*).

Discrete jump motion. In addition to either Brownian or non-Brownian diffusion, one may specify a discrete jump motion amongst n sites having n-fold symmetry about z_R, with a rate given by D_j. It is only appropriate to use this model in combination with other rotational diffusion when the jump motion is much faster than the overall molecular rotation around z_R. If the jump motion is the only important feature of the system, the R tensor elements should be set to very small values.

Anisotropic viscosity. This model is intended for use with ordered fluids (cf. below) in which the rotational diffusion rate of the probe may depend on its orientation with respect to the director frame. The specific model used applies to the case in which the director lies along the spectrometer field (i.e., $\psi = 0$) and includes rates for diffusion perpendicular and parallel to the field direction (called $\hat{R}_\perp$ and $\hat{R}_\parallel$) *in addition to* the molecular diffusion tensor. This model cannot be used with either non-Brownian or discrete jump motion.

Heisenberg spin exchange. In addition to any of the diffusional models mentioned above, it is possible to account approximately for the effects of isotropic Heisenberg spin exchange between probe molecules with a rate specified by ω_{SS}.

All dynamic parameters are specified as $\log_{10}$ of the parameter, divided by units of seconds^{-1} or seconds as appropriate. (Ratios such as rotational anisotropies N and N_{xy} and non-Brownian model parameters P_i do not require this cancellation of units). The convention applies to the rotational diffusion tensor, the Heisenberg spin-exchange rate, and the jump rate in the discrete jump model, as well as the residence times in the non-Brownian diffusion models. Logarithmic scaling has been adopted because it adjusts the dynamic parameters to be of the same order of magnitude as the other fitting parameters, which can substantially improve the convergence properties of the least-squares algorithm.

Orienting potential. For diffusion of a probe in a medium with microscopic molecular ordering such as a liquid crystal, membrane, or polymer, the tendency of the probe to order is modeled by a restoring potential that is defined relative to the director axes of the liquid. The potential is expanded in a series of spherical harmonics $D^L_{0K}(\Omega) \equiv Y^L_K(\beta, \gamma)$

$$U(\Omega) = -k_B T \sum_{L,K} c^L_K D^L_{0K}(\Omega), \qquad [1]$$

where $\Omega = (\alpha, \beta, \gamma)$ represents the angles relating the diffusion axes to the director frame (β and γ are the polar angles of the director z axis in the diffusion frame), and the dimensionless coefficients c^L_K have the property that $c^L_K = c^L_{-K} = c^L_K{}^*$. In the present implementation, the summation in Eq. [1] is restricted to even values of K and L up to 4 ($K \leqslant L$). Non-Brownian diffusion models may not be used with an orienting potential.

The programs allow for microscopic order with macroscopic disorder (MOMD), i.e., a case in which there are domains with local ordering described by the potential in Eq. [1], but whose directors are isotropically distributed so that there is no macroscopic ordering (*26*). This case is modeled by integrating the spectral lineshape over the director tilt angle ψ using a specified number of ψ values.

Tensor Representations

All of the tensor properties used in the spectral calculation, including the *g*-matrix, the *A*-tensor, the linewidth tensor W, and the rotational diffusion tensor, R, may be expressed in three different forms: (a) *Cartesian form* $\{M_x, M_y, M_z\}$, the representation that is most frequently used in spectral fitting procedures; (b) axial form $\{M_\perp, M_\parallel\}$, which also refers to the Cartesian axes, but is restricted to axial symmetry; and

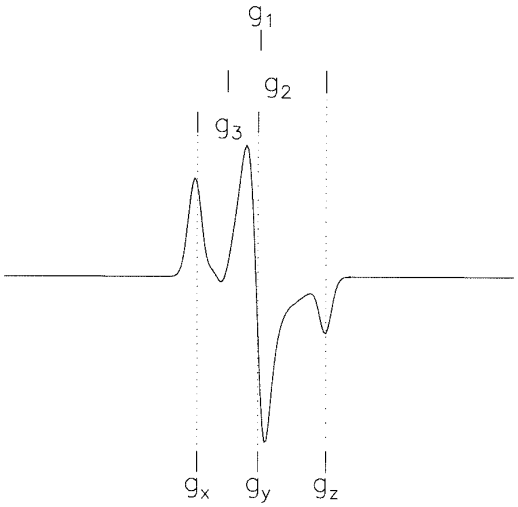

FIG. 1. Illustration of Cartesian versus ''spherical'' component specification of the g anisotropy for a $S = \frac{1}{2}$, $I = 0$ system. The spherical components correspond to (i) the isotropic average, $(x + y + z)/3$, (ii) the axial component, $z - (x + y)/2$, and (iii) the rhombic component, $(x - y)$. The hyperfine, electronic Zeeman, rotational diffusion, and inhomogeneous linewidth tensor components may all be specified in either form.

(c) ''*modified spherical*'' *form* $\{M_1, M_2, M_3\}$, the components of which are defined in terms of the Cartesian components as

$$\begin{aligned} M_1 &= \tfrac{1}{3}(M_x + M_y + M_z) \\ M_2 &= M_z - \tfrac{1}{2}(M_x + M_y) \\ M_3 &= M_x - M_y. \end{aligned} \qquad [2]$$

In this representation, M_1, M_2, and M_3 represent the isotropic, axial, and rhombic components of the tensor, respectively. These components differ from the usual second-order spherical tensor components, $M(0, 0)$, $M(2, 0)$, and $M(2, 2)$ by constant factors, but are defined in such a way that they may be directly correlated with and/or estimated from features of an experimental ESR spectrum. This is illustrated in Fig. 1, which depicts a typical ''powder pattern'' spectrum for a randomly oriented electron spin with rhombic g anisotropy and no coupled nucleus. The three turning points in the spectrum correspond to the standard Cartesian components x, y, and z, of the g tensor. The modified spherical components correspond to (i) the isotropic average, $(x + y + z)/3$, (ii) the axial component, $z - (x + y)/2$, and (iii) the rhombic component, $(x - y)$, which may be determined from the spectral features as indicated in the figure. A transformation similar to that given in Eq. [2] has been applied previously by us (*15–18*) for varying components of the g and A tensors, and by Hustedt *et al.* (*19*) specifically for varying components of the g tensor. The hyperfine, inhomogeneous linewidth, and rotational diffusion tensor components may all be specified in a similar way.

The rotational diffusion tensor is a special case, since the transformations in Eq. [2] are applied to the $\log_{10}$ of the rate constant (per seconds^{-1}). Thus, the average rotational diffusion rate constant $\bar{R} \equiv \sqrt[3]{R_x R_y R_z}$ is the *geometric* mean of the rate constants for rotational diffusion about each of the axes of the molecular diffusion frame; similarly, the axial rotational anisotropy $N \equiv R_z/\sqrt{R_x R_y}$ *is the ratio* of the rate constants for diffusion about axes perpendicular and parallel to the z_D axis. It should be noted that, in the case of axial rotational symmetry ($R_x = R_y$), this definition of N just given is consistent with the conventional usage (*10*); however, that for $\bar{R}$ differs from the more commonly employed definition, $\bar{R} = \sqrt{R_\perp R_\parallel}$ (*10*).

There are several reasons for introducing alternative representations for the various tensor quantities used in the fitting procedure. It is often convenient to constrain a tensor to maintain a constant trace or axial symmetry during a minimization, based on some prior knowledge of the system (e.g., the isotropic g value and hyperfine splitting from motionally narrowed spectra). This is easily accomplished using the modified spherical tensor components: for example, in order to maintain axial symmetry, one would set M_3 to zero and vary M_1 and M_2; in order to maintain a constant trace, one would fix M_1 at the desired value and vary M_2 and M_3. In addition, it frequently occurs that small deviations from spherical or rhombic symmetry of a tensor cannot be resolved from a given EPR spectrum. This situation is identified by the appearance of strong correlations amongst the elements of a given tensor when they are varied in a minimization. Such correlations can sometimes be avoided by using a different representation of the tensor (cf. Examples and Discussion below).

Numerical Calculation of Slow-Motion EPR Spectra

The computational methods used to calculate 1D CW and 2D FT slow-motion magnetic resonance spectra using the stochastic Liouville equation have been described in detail (*10–13, 27, 28*). Only a brief overview of those features of the spectral calculation that are most relevant to the nonlinear-least-squares minimization procedure will be presented here.

One-Dimensional CW Spectra

The unsaturated CW EPR intensity as a function of frequency, $I(\omega)$, may be represented in compact ''matrix element'' notation (*10–12*) as

$$I(\omega) = \langle\langle \nu | (\tilde{\Gamma} - i\mathcal{L}) + i\omega \mathbf{1} | \nu \rangle\rangle, \qquad [3]$$

where $\langle\langle \nu |$ is a vector representing the observable magneti-

zation, including the $\mathbf{S}_+$ spin operator and the equilibrium probability distribution function for the orientation of the radicals, $\tilde{\Gamma}$ is the ''symmetrized'' diffusion operator used to model the classical motion of the radicals, and L is the Liouville superoperator matrix derived from the orientation-dependent spin Hamiltonian. The matrix equation [3] is constructed using a basis set with a typical dimension of from 30 to 20,000, depending upon the motional rates, degree of probe ordering, and the desired spectral resolution, which increases with EPR frequency and the number of spectral dimensions. The basis set is constructed as a product of generalized spherical harmonics $D^L_{MK}(\Omega)$ and spin transition indices p and q for the electronic (S) and nuclear (I) spins:

$$|L, M, K, p^{\mathrm{S}}, q^{\mathrm{S}}, p^{\mathrm{I}}, q^{\mathrm{I}}\rangle \equiv |L, M, K\rangle \otimes |p^{\mathrm{S}}, q^{\mathrm{S}}, p^{\mathrm{I}}, q^{\mathrm{I}}\rangle$$

$$|L, M, K\rangle \equiv \sqrt{\frac{2L+1}{8\pi}}\, D^L_{MK}(\Omega), \qquad [4]$$

where $p^{\mathrm{S}} = m''_{\mathrm{S}} - m'_{\mathrm{S}}$ and $q^{\mathrm{S}} = m''_{\mathrm{S}} - m'_{\mathrm{S}}$ for the transition between electronic spin states $|S, m'_{\mathrm{S}}\rangle$ and $|S, m''_{\mathrm{S}}\rangle$, and p^{I} and q^{I} are similarly defined for the nuclear spin states. In practice, the stochastic Liouville operator is constructed using a basis set that is obtained by forming certain symmetry-derived linear combinations of the vectors defined in Eq. [4], which permits the matrix to be expressed in complex symmetric form, and also allows reduction of the basis set in many cases. The details of these transformations are given elsewhere (*11, 28*).

It is convenient to divide the basis set into subspaces corresponding to different p^{S} indices. These subspaces are coupled only through the $\mathbf{S}_+$ or $\mathbf{S}_-$ operators, which appear only in the nonsecular terms in the spin Hamiltonian or in the interaction with a strong irradiating microwave field (cf. below). Thus, in the high-field limit where nonsecular terms may be neglected, the unsaturated CW EPR lineshape calculation requires only the subspace upon which the vector $|\nu\rangle\rangle$ has a projection, namely the $p^{\mathrm{S}} = 1$, $q^{\mathrm{S}} = 0$ subspace (*11, 12*). The L matrix does not need to be fully diagonalized in order to obtain the CW spectrum; the least-squares calculation uses a Lanczos tridiagonalization of the matrix followed by a continued-fraction approximation to the spectral lineshape using the elements of the tridiagonal matrix (*11–13*). The conjugate gradients version of the Lanczos algorithm (*27*) is employed so that the number of Lanczos steps required for convergence of the spectrum may be determined during each calculation. Thus, the number of steps taken can vary during the course of a minimization procedure. Proper choice of a basis set for the spectral calculation is critical to efficient nonlinear-least-squares analysis, particularly for the fitting of 2D spectra. Too large a basis set can result in excessive computation for each spectral evaluation, significantly diminishing the efficiency of the procedure. For more demanding problems, the basis set truncation method described by Vasavada *et al.* (*27*) is typically employed to identify and eliminate elements of the basis set that are unimportant for the solution of the problem at hand. This procedure can dramatically reduce the time required to compute the spectrum, particularly for 2D calculations. However, too small a basis set may introduce unacceptable deviations from the ''true'' spectral lineshape, even to the point of producing physically unreasonable spectra. In addition, an insufficient basis set can have more subtle adverse effects during the course of nonlinear-least-squares minimization. First, since the minimization must carry out spectral evaluations over a range of parameter values, the search may extend beyond the region for which a given truncated set is valid. Second, the calculation of partial derivatives of the spectrum with respect to fitting parameters by the forward-difference approximation (cf. below) requires a higher degree of accuracy in the spectral intensities. Thus, a basis set which is sufficient to produce a given spectrum with acceptable accuracy may not be large enough to permit accurate calculation of its partial derivatives. The qualitative consequences of inaccuracies in the partial derivative spectra are described in greater detail under Examples and Discussion.

Two-Dimensional Fourier-Transform Spectra

Two basic types of two-dimensional Fourier-transform (2D FT) experiments have been demonstrated to have major utility in the analysis of spin-label dynamics. The first is the two-pulse 2D COSY (correlation spectroscopy) experiment, and the second is the three-pulse 2D ELDOR (electron-electron double resonance) experiment, which is analogous to 2D exchange experiments in NMR. In both types of experiments, a dual-quadrature signal is detected; that is, the signal has a real, absorptive part and an imaginary, dispersive part with respect to each of the two frequencies defining the 2D spectrum. In both the 2D COSY and 2D ELDOR experiments, these signals can be combined in appropriate linear combinations to yield an FID-like signal called the S_{c+} signal and an echo-like signal called S_{c-}. Because of partial cancellation of inhomogeneous broadening in the echo-like signal, the S_{c-} spectra are, in general, substantially sharper than the S_{c+}, so that a comparison of the two aids in the discrimination of the different broadening contributions to the lineshape (*21, 22, 28*).

In general, 2D FT experiments contain pulses that couple the different p^{S} manifolds, and thus spectral calculation requires the $p^{\mathrm{S}} = 0$ (''diagonal'') subspace corresponding to longitudinal magnetization as well as the $p^{\mathrm{S}} = \pm 1$ ''off-diagonal'' subspaces, which are conjugate to each other and correspond to counter-rotating transverse magnetization. For slow-motion 2D FT spectral calculations, it is generally nec-

essary to diagonalize the stochastic Liouville matrix in each of the three subspaces, in order to obtain the eigenmodes that fully describe the evolution of the spin system in the absence of any microwave pulse in the high field limit, i.e.,

$$\mathsf{O}_1^{\mathrm{tr}}\mathsf{L}_1\mathsf{O}_1 = \mathbf{\Lambda}_1 \quad \mathsf{O}_0^{\mathrm{tr}}\mathsf{L}_0\mathsf{O}_0 = \mathbf{\Lambda}_0 \quad \mathsf{O}_{-1}^{\mathrm{tr}}\mathsf{L}_{-1}\mathsf{O}_{-1} = \mathbf{\Lambda}_{-1}. \tag{5}$$

Here, the O and $\mathbf{\Lambda}$ matrices are respectively the orthogonal transformation matrices composed of the eigenvectors and the diagonal matrices of eigenvalues for each subspace with the p^{S} values for the subspace specified as a subscript. Detailed expressions for L and a discussion of the properties of O and $\mathbf{\Lambda}$ in each manifold are given elsewhere (*28*).

The pulses used in a 2D FT EPR experiment are expressed in the form of pulse propagator matrices that depend only on the p^{S} and q^{S} indices, and whose matrix elements in the $|p^{\mathrm{S}}, q^{\mathrm{S}}\rangle$ basis may be expressed in terms of the pulse tipping angle θ and phase ϕ (*28–30*). Given the eigenvectors and eigenvalues in Eq. [5] and a sequence of appropriate pulse propagators, it is straightforward to calculate the 2D EPR signal for the 2D correlation spectroscopy (COSY) and 2D exchange (ELDOR) experiments utilized in this work. The initial density matrix is the equilibrium density matrix which has nonzero elements only for zeroth-order coherence ($p^{\mathrm{S}} = 0$). The initial $\pi/2$ pulse in each sequence converts this magnetization to first-order coherence which is represented as the two counter-rotating transition moment vectors $|\nu_{\pm 1}\rangle$ that specify the transverse magnetization corresponding to density matrix elements with $p^{\mathrm{S}} = \pm 1$. During the detection period, only the $|\nu_{-1}\rangle$ transition moment contributes to the observable magnetization $\langle \mathbf{S}_+ \rangle$. Thus, 2D ESR signals for the coherence-transfer pathways of interest may be written as a product of $|\nu_{\pm 1}\rangle$ with the appropriate combination of transformation matrices, exponentiated eigenvalue matrices, and pulse propagator matrices, which is then projected onto $\langle \nu_{-1}|$.

In the COSY sequence, the coherence-transfer pathways $p^{\mathrm{S}} = 0 \rightarrow \mp 1 \rightarrow -1$ are responsible for the $S_{\mathrm{c}+}$ and $S_{\mathrm{c}-}$ signals defined by Gamliel and Freed (*31*), which are respectively known as ''P-type'' and ''N-type'' signals in the NMR literature. Those signals can be written as

$$\begin{aligned} S_{\mathrm{c}\pm}^{\mathrm{COSY}} &= \langle \nu_{-1}|\mathsf{O}_{-1}\exp(-\mathbf{\Lambda}_{-1}t_2)\mathsf{O}_{-1}^{\mathrm{tr}}\mathsf{P}_{(-1\leftarrow\mp 1)}\mathsf{O}_{\mp 1} \\ &\quad \times \exp(-\mathbf{\Lambda}_{\mp 1}t_1)\mathsf{O}_{\mp 1}^{\mathrm{tr}}|\nu_{\mp 1}\rangle, \end{aligned} \tag{6}$$

where $\mathsf{P}_{(p_2^{\mathrm{S}}\leftarrow p_1^{\mathrm{S}})}$ denotes the pulse propagator that transforms the density-matrix elements from the p_1^{S} subspace into the p_2^{S} subspace. The pulse propagators $\mathsf{P}_{(p_2^{\mathrm{S}}\leftarrow p_1^{\mathrm{S}})}$ in Eq. [6] are proportional to the unit matrix.

The coherence-transfer pathways for the 2D ELDOR sequence can be written as $p^{\mathrm{S}} = 0 \rightarrow \mp 1 \rightarrow 0 \rightarrow -1$ for the $S_{\mathrm{c}\pm}$ signals. The signal following the third pulse is given by

$$\begin{aligned} S_{\mathrm{c}\pm}^{\mathrm{ELDOR}} &= \langle \nu_{-1}|\mathsf{O}_{-1}\exp(-\mathbf{\Lambda}_{-1}t_2)\mathsf{O}_{-1}^{\mathrm{tr}}\mathsf{P}_{(-1\leftarrow 0)}\mathsf{O}_0\exp(-\mathbf{\Lambda}_0 T) \\ &\quad \times \mathsf{O}_0^{\mathrm{tr}}\mathsf{P}_{(0\leftarrow\mp 1)}\mathsf{O}_{\mp 1}\exp(-\mathbf{\Lambda}_{\mp 1}t_1)\mathsf{O}_{\mp 1}^{\mathrm{tr}}|\nu_{\mp 1}\rangle. \end{aligned} \tag{7}$$

Because the diagonal space is larger, the pulse propagators between the diagonal space and the off-diagonal space in this equation are rectangular. Detailed expressions for the matrix elements of L in each of the p^{S} manifolds, and the methods for efficient diagonalization of the stochastic Liouville equation are given elsewhere (*28*).

Spectral Fitting Algorithm

''Model Trust Region'' Levenberg–Marquardt Algorithm

The specific algorithm used in this work closely follows the implementation found in the MINPACK program library (*32*), which utilizes one of the quasi-Newton methods known as the ''model trust region'' modification of the Levenberg–Marquardt approach (*33, 34*). Previous versions of our slow-motion fitting programs (*15–18*) utilized the basic Levenberg–Marquardt algorithm available in Version 4 of the IMSL library (*35*).

The experimental data are represented as a column vector $\mathbf{h}$ consisting of a list of values h_i containing the intensity of the spectrum at each point in the spectral domain. The vector $\mathbf{h}$ may include a sequence of spectra that are to be fitted simultaneously, a spectrum with multiple dimensions, or a sequence of multidimensional spectra. The experimental data values are always represented as a one-dimensional vector regardless of the actual number or dimensions of the spectra. The two- and three-dimensional spectra considered in this work are specified in leading-dimension order (i.e., spectra depending upon the frequencies ω_1 and ω_2 are stored so that points corresponding to the same ω_1 value are contiguous). A similar scheme has been utilized in the nonlinear-least-squares analysis of motionally averaged 2D NMR spectra (*36*).

The theoretical spectrum calculated at each point in the spectral domain is similarly represented as the vector $\phi(\mathbf{x})$, where $\mathbf{x}$ is a vector containing the set of parameters to be determined by the fitting procedure. The problem is then to find the $\mathbf{x}$ that best matches $\phi(\mathbf{x})$ with $\mathbf{h}$ according to some criterion, most commonly, according to the minimum in the squared differences between the vectors. More explicitly, given the residual vector $\mathbf{f}(\mathbf{x}) \equiv \mathbf{h} - \phi(\mathbf{x})$, we may seek to minimize the *unweighted* sum of squared residuals

$$s^2 = |\mathbf{f}(\mathbf{x})|^2 = \mathbf{f}^{\mathrm{T}}(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}) \tag{8a}$$

It is also possible to weight the residuals using some *a priori* knowledge of the experimental error, which is taken to be the uncertainty in the measured spectral intensity at a given field value, i.e., the experimental noise level. The standard deviation of the noise, σ, is estimated from the experimental points along a sufficient length of baseline (or area of baseplane). In practice, σ is obtained from the variance of a linear fit to the two baseline segments at either end of the spectrum that should comprise about 10–20% of the spectral sweep width. Similarly, a plane may be fitted to a 10–20% edge around the border of a two-dimensional spectrum to estimate the spectral noise. The noise level is assumed to be constant across a given spectrum; however, for global fits involving multiple spectra, σ may be different for each spectrum. The sum of weighted residuals, which should be the ''true'' χ^2 of the fit, is calculated as

$$\chi^2 = \mathbf{f}^{\mathrm{T}}(\mathbf{x}) \cdot \boldsymbol{\sigma}^{-1} \cdot \mathbf{f}(\mathbf{x}), \qquad [8b]$$

where $\boldsymbol{\sigma}$ is a diagonal matrix whose elements contain the experimental σ_i for each element of $\mathbf{f}$.

The minimization procedure takes an initial estimate for the parameters being searched, $\mathbf{x}_0$, to calculate a refined estimate of the parameters, $\mathbf{x}_+$, according to the formula

$$\mathbf{x}_+ = \mathbf{x}_0 - [\mathsf{J}^{\mathrm{T}}(\mathbf{x}_0)\mathsf{J}(\mathbf{x}_0) + \lambda \mathsf{D}]^{-1}\mathsf{J}(\mathbf{x}_0)\mathbf{f}(\mathbf{x}_0), \qquad [9]$$

where $\mathbf{J}(\mathbf{x}_0)$ represents the *Jacobian matrix* of $\mathbf{f}$ (or $\boldsymbol{\sigma}^{-1} \cdot \mathbf{f}$ for weighted residuals) with respect to the parameters $\mathbf{x}$, evaluated at $\mathbf{x}_0$

$$\mathbf{J}(\mathbf{x}_0) = \left(\frac{\partial \mathbf{f}}{\partial x_1}\right)_{x_0} \oplus \left(\frac{\partial \mathbf{f}}{\partial x_2}\right)_{x_0} \oplus \cdots = \begin{bmatrix} \left(\dfrac{\partial f_1}{\partial x_1}\right)_{x_0} & \left(\dfrac{\partial f_1}{\partial x_2}\right)_{x_0} & \cdots \\ \left(\dfrac{\partial f_2}{\partial x_1}\right)_{x_0} & \left(\dfrac{\partial f_2}{\partial x_2}\right)_{x_0} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \qquad [10]$$

and D is a diagonal matrix with D_{ii} containing a scaling factor controlling the step length for the parameter x_i. For weighted residuals, the ith row of J is also multiplied by a factor of σ_i^{-1}. In practice, each column of J is calculated by a forward-differences approximation

$$\left(\frac{\partial \mathbf{f}(\mathbf{x})}{\partial x_i}\right) = \frac{1}{\delta x_i}[\mathbf{f}(\mathbf{x} + \hat{\mathbf{x}}_i \delta x_i) - \mathbf{f}(\mathbf{x})]. \qquad [11]$$

where $\hat{\mathbf{x}}_i$ is the unit vector corresponding to x_i in $\mathbf{x}$.

User control is provided over some of the parameters in Eqs. [9] and [11]. The relative forward-differences step δx_i (typically 10^{-5}) is adjustable for each search parameter, which provides a means of checking for roundoff errors in the computation of J. The accuracy of the calculation of J is critical to the proper performance of the algorithm and correct statistical analysis of the results. The user may also specify a multiplicative factor for each of the scaling factors D_{ii}, which are initially set to the norms of the corresponding column of J. This effectively increases or decreases the search step size for the given parameter.

The scalar λ in Eq. [9] is the Levenberg–Marquardt parameter. In the simplest and most commonly used implementation of the Levenberg–Marquardt method (*37*), the steps along each parameter dimension are not scaled (i.e., $\mathsf{D} = \mathsf{I}$), and λ is either increased or decreased by a constant factor at each iteration, depending on whether $|\mathbf{f}(\mathbf{x}_+)|^2 > |\mathbf{f}(\mathbf{x}_0)|^2$ or $|\mathbf{f}(\mathbf{x}_+)|^2 < |\mathbf{f}(\mathbf{x}_0)|^2$. This approach has the drawback that the optimal value of λ is not determined at each iteration, and certain pathologies in the J matrix may arise if the scales of individual parameters are too different. In our previous versions (*15–18*), this problem was reduced by the use of a log scale for parameters such as the R tensor, and also by further rescaling of the parameters.

More sophisticated implementations carry out a line search in order to minimize $|\mathbf{f}(\mathbf{x}_+)|^2$ with respect to λ at each iteration, but this approach requires additional function evaluations. However, if one defines a step bound δ_c (the trust region) around $\mathbf{x}_0$ within which the residual function is assumed to be well approximated by the equation

$$|\mathbf{f}(\mathbf{x})|^2 \cong |\mathbf{f}(\mathbf{x}_0)|^2 - [\mathsf{J}^{\mathrm{T}}(\mathbf{x}_0)\mathbf{f}(\mathbf{x}_0)] \cdot \mathbf{x} + \tfrac{1}{2}\mathbf{x} \cdot [\mathsf{J}^{\mathrm{T}}(\mathbf{x}_0)\mathsf{J}(\mathbf{x}_0)] \cdot \mathbf{x}, \qquad [12]$$

then the $\mathbf{x}_+$ corresponding to the optimal λ may be found by a line search that utilizes the information in $\mathsf{J}(\mathbf{x}_0)$ without recalculating the spectrum. The initial step bound δ_c is taken to be the lesser of the quantities $a|\mathsf{D} \cdot \mathbf{x}|$ (a is a user-specified scaling factor) and the length of the first step $|\mathbf{x}_+ + \mathbf{x}_0|$ calculated according to Eq. [9]. Subsequently, δ_c may be either increased or decreased depending upon how closely the change of the residual function predicted by Eq. [12], Δf_{pred}, matches the calculated change, $\Delta f_{\mathrm{calc}} = |\mathbf{f}(\mathbf{x}_0)| - |\mathbf{f}(\mathbf{x}_+)|$. If $\Delta f_{\mathrm{calc}}/\Delta f_{\mathrm{pred}} \geqslant 0.75$, then either Eq. [12] is a good approximation to $|\mathbf{f}(\mathbf{x})|$ within the trust region or the function continues to decrease rapidly beyond the step bound, and δ_c may be increased, typically by a factor of 2. If $\Delta f_{\mathrm{calc}}/\Delta f_{\mathrm{pred}} < 0.1$, Eq. [12] is not a good approximation throughout the trust region, and δ_c must be decreased. This modification of the basic Levenberg–Marquardt method is among those most generally recommended for calculation-

intensive problems such as EPR spectral fitting in recent treatises on nonlinear-least-squares methods (*33, 34, 38*).

After the completion of a step, $\mathbf{x}_+$ is taken as the new $\mathbf{x}_0$ and the procedure repeats until it converges to a minimum. Convergence is judged according to one or more of the following criteria: (i) *function convergence,* where the relative change in residuals $(|\mathbf{f}(\mathbf{x}_+)|^2 - |\mathbf{f}(\mathbf{x}_0)|^2)/|\mathbf{f}(\mathbf{x}_0)|^2$ after a step is less than some predefined tolerance, `ftol` (typically 10^{-4}); (ii) *parameter convergence,* for which the scaled norm of the step $|\mathsf{D}(\mathbf{x}_+ - \mathbf{x}_0)|^2$ is less than the tolerance `xtol` (typically 10^{-5}), and (iii) *gradient convergence,* in which the scaled norm of each column of $\mathsf{J}(\mathbf{x}_+)$ is less than some prespecified tolerance `gtol` (typically 10^{-6}), which can occur if $\mathbf{x}_+$ is at an extremum of $\mathbf{f}(\mathbf{x})$. True convergence according to the third of these criteria is relatively rare; most often, termination due to gradient convergence signals an extremely weak dependence of the spectrum on all of the fitting parameters.

Separable Fitting Parameters

In separable least-squares problems, it is possible to partition $\mathbf{x} = (\xi, \eta)$ such that for a given set of nonlinear parameters ξ, $\|\mathbf{f}(\xi, \eta)\|$ is easily minimized with respect to η. Such a separability has the advantages that no initial estimate for η is needed, and fewer nonlinear search parameters are involved, reducing the time needed to compute J. One important special case of a separable problem for the present application is when η contains only linear parameters. Although it is possible to minimize a function of this type using a single-step iterative algorithm (*39*), we utilize an equivalent but somewhat less efficient procedure in which the nonlinear functions depending on ξ are first calculated and the linear parameters in η determined by linear least-squares at each iteration of the nonlinear parameter search. This ''nested'' procedure is necessary to accommodate the possibility of spectral shifting (cf. below), but it does take implicit advantage of the theorems proved by Golub and Pereyra (*39*) that establish the properties of global convergence for this method.

Previous versions of our slow-motion fitting programs for CW EPR spectra (*15–18*) have relied on a different sort of separation-of-parameters technique for optimizing the calculation. In those versions, the fitting parameters were divided into two groups according to whether a change in the parameter required a retridiagonalization of the L matrix. Parameters which enter only into the continued-fraction part of the spectral calculation, such as field range and offset, inhomogeneous linewidth, phase, and scale, were optimized separately by a nested nonlinear-least-squares procedure. This scheme was designed to avoid costly recalculation of the L matrix during the course of the minimization. In the present version, the same objective is achieved by maintaining the list of search parameters in an order such that the parameters that do not require recalculation of L precede all others in the list. The overall efficiency is further increased by carrying out the optimization of the separable parameters without resorting to nonlinear minimization methods, as we now describe.

Scale Factors for Multicomponent Spectra

Separable linear parameters appear in the minimization problem when the experimental data are to be fitted using a sum of spectral components, each with its own scaling factor. If all of the calculated component spectra are normalized to the same integral (or double integral, for first-derivative spectra), this scaling procedure is equivalent to determining the relative population of each species, which is a common objective in the analysis of multicomponent spectra in general (*40–42*). Specifically, at each iteration of the nonlinear parameter minimization, the scaling factors a_i are obtained by solving a linear equation of the form

$$\mathbf{h} = \mathbf{\Phi}\cdot\mathbf{a} = [\phi_1 \oplus \phi_2 \oplus \cdots]\cdot\mathbf{a}, \qquad [13]$$

where $\mathbf{h}$ is the data vector and the design matrix $\mathbf{\Phi}$ is composed of column vectors corresponding to each of the calculated component spectra. In the case that several spectra are being fitted globally, the spectra are packed into the column vectors ϕ_i in the same order as the experimental data. The solution vector $\mathbf{a}$ is obtained from a standard QR factorization of the design matrix $\mathbf{\Phi}$ (*43*),

$$\mathbf{\Phi} = \mathsf{QR}; \quad \mathbf{a} = \mathsf{R}^{-1}\mathsf{Q}^{\mathrm{T}}\mathbf{h}, \qquad [14]$$

where Q is an orthogonal rectangular matrix, and R is an upper triangular square matrix. Then the best-fit spectrum for the given set of nonlinear parameters is given by

$$\phi(\mathbf{x}) = \sum_i \mathbf{a}_i\phi_i(\mathbf{x}). \qquad [15]$$

Because of the very large memory requirements it would impose, this type of multicomponent analysis has not been implemented in the 2D spectral calculations described above.

The use of an orthogonal decomposition method such as the *QR* factorization has the advantage that linear dependencies or near-dependencies among the component spectra can be identified and removed. That is, given two components having practically indistinguishable spectra (aside from a possible linear scaling factor), one of them may be eliminated from the linear-least-squares problem. This is accomplished by utilizing column pivoting during the *QR* procedure and examining the diagonal elements of the resultant R matrix.

The method outlined above bears some resemblance to other methods of multicomponent EPR spectral analysis that have been described in the literature. The iterative

subtraction technique of Evans and Morgan (*41*) is a form of orthogonal decomposition that offers a similar ability to distinguish the number of independent components in a given spectrum. It suffers the drawback that it requires regions in which there is little overlap between the component spectra in order to perform optimally. The common factor analysis employed by Moens *et al.* (*42*) is based on a method that is widely applied in chemistry and related fields (*40*) and also relies on orthogonal decomposition. However, this type of approach requires a set of experimental spectra in which the relative populations of the individual components are varied, which is usually impractical for slow-motion spectra.

Spectral Shifting

The vector of separable parameters η may contain nonlinear parameters as long as η is uniquely determined for each ξ. One nonlinear EPR parameter that is usually separable in this sense is the trace of the g tensor, which determines the center frequency (or field position) of the spectrum. This quantity is particularly important to fit in cases where either the spectrometer frequency or the absolute magnetic field is not known to sufficient accuracy. Most spectral fitting procedures that are used to obtain full g-tensor information from randomly (or partially) oriented spin probes include all three tensor components as nonlinear parameters in the fit. However, one component may be removed from the set of nonlinear parameters by introducing a *spectral shift*, $\Delta\omega$, which is determined at each iteration. In the past (*15–19*), a nested procedure has been used, wherein the optimum spectral shift $\Delta\omega$ is located by a search algorithm such as a binomial or golden-section search, while a linear-least-squares fit is used to optimize the baseline shift and scale factor for the signal amplitude. The use of spline-interpolated data allowed a minimum number of data points to be fitted. In the procedure described below, we utilize a different approach involving rapid calculation of the correlation function between the model spectrum and experimental data, which may be used to determine the spectral shift in one-dimensional spectra, as well as the shift along one dimension of a 2D spectrum. A similar procedure is used for shifting multicomponent spectra, as described below.

The sum of the squared differences between an experimental spectrum $h(\omega)$ and a shifted calculated spectrum $\phi(\omega - \Delta\omega)$ can be written (assuming continuous functions for the moment)

$$\begin{aligned}\|f(\Delta\omega)\|^2 &= \int_{-\infty}^{+\infty} d\omega[h(\omega) - \phi(\omega - \Delta\omega)]^2 \\ &= \int_{-\infty}^{+\infty} h(\omega)^2 d\omega + \int_{-\infty}^{+\infty} \phi(\omega - \Delta\omega)^2 d\omega \\ &\quad - 2\int_{-\infty}^{+\infty} h(\omega)\phi(\omega - \Delta\omega)d\omega. \end{aligned} \quad [16]$$

The first two terms on the right-hand side of this equation are positive constants, whereas the third term is simply the correlation function for $h(\omega)$ and $\phi(\omega)$, which is a function of $\Delta\omega$. Thus, the sum of residuals can be written

$$\|f(\omega)\|^2 = \|h(\omega)\|^2 + \|\phi(\omega)\|^2 - 2\,\mathrm{corr}\{h, \phi\}(\Delta\omega) \quad [17]$$

from which it is evident that the minimum in the residuals $\|f(\Delta\omega)\|^2$ occurs at the maximum of the correlation function $\mathrm{corr}\{h, \phi\}(\Delta\omega)$. The problem is thus reduced to finding the value of $\Delta\omega$ at this maximum. For discrete data, it is straightforward to calculate $\mathrm{corr}\{h, \phi\}(\Delta\omega)$ utilizing the properties of the discrete fast Fourier transform, and the desired value of $\Delta\omega$ may then be obtained by polynomial interpolation of the maximum from the tabulated correlation function. The accuracy of this interpolation is critical to the smooth and reproducible performance of the Levenberg–Marquardt algorithm, since the extrema in the correlation function are quite sharp, i.e., the residuals depend very strongly on $\Delta\omega$. The specific interpolation scheme used is the five-point Lagrange formula for equally spaced abscissas (*44*); the root of this function's derivative within the five-point region is then obtained by a simple polynomial root-finding algorithm (*37*).

Combined Spectral Shifting and Scaling

It frequently occurs in multicomponent fitting problems that both the spectral shift and the individual spectral scaling factors must be determined. In this case, it is possible to combine the methods described above to determine both the linear and nonlinear components of η. For this application, it is most convenient to use the following form of the linear equation for the scale factors:

$$\begin{bmatrix} \mathrm{corr}\{h, \phi_1\}(\Delta\omega) \\ \mathrm{corr}\{h, \phi_2\}(\Delta\omega) \\ \mathrm{corr}\{h, \phi_3\}(\Delta\omega) \\ \vdots \end{bmatrix} = \begin{bmatrix} \|\phi_1\|^2 & \phi_1\cdot\phi_2 & \phi_1\cdot\phi_3 & \cdots \\ \phi_2\cdot\phi_1 & \|\phi_2\|^2 & \phi_2\cdot\phi_3 & \cdots \\ \phi_3\cdot\phi_1 & \phi_3\cdot\phi_2 & \|\phi_3\|^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} a_1(\Delta\omega) \\ a_2(\Delta\omega) \\ a_3(\Delta\omega) \\ \vdots \end{bmatrix}. \quad [18]$$

This equation is equivalent to the normal equation form of a linear least-squares problem, with the exception that $\mathrm{corr}\{h,\phi_i\}(\Delta\omega)$ for a specific value of $\Delta\omega$ replaces $h\cdot\phi_i$ in the left-hand vector elements. The norms and dot products

of the component spectra that appear in the matrix are first calculated and stored as shown. Since this matrix is independent of $\Delta\omega$, it is sufficient to carry out a single *QR* decomposition of it that may subsequently be used for solving Eq. [18] for a series of different $\Delta\omega$ values. Then the correlation functions between the data and each of the component spectra, $\mathrm{corr}\{h, \phi_i\}(\Delta\omega)$, are calculated using the discrete fast Fourier transform. For each $\Delta\omega$ in the domain, the left-hand-vector is substituted with the values of $\mathrm{corr}\{h, \phi_i\}(\Delta\omega)$, and the optimal scaling factors for each component at the given $\Delta\omega$ are found by solving the set of linear normal equations in a manner similar to Eq. [14]. The set of solutions for a range of $\Delta\omega$ values may be calculated at once by forming a matrix on the left-hand side of Eq. [18] with a column for each $\Delta\omega$. The sum function $\Sigma_i\, a_i(\Delta\omega)\mathrm{corr}\{h, \phi_i\}(\Delta\omega)$ is then tabulated as a function of $\Delta\omega$ and the optimal value of $\Delta\omega$ determined by polynomial interpolation as described above for the single-component case.

Statistical Inference from Nonlinear-Least-Squares Minimization

An essential part of experimental measurements using least-squares spectral analysis is the estimation of measurement error, that is, some estimate of how close the parameters at the least-squares minimum are to the ''true'' parameter values. The Levenberg–Marquardt and related quasi-Newton methods have a number of advantages for nonlinear-least-squares analysis that are underutilized in magnetic-resonance applications. These algorithms all share the feature that the Jacobian matrix is calculated during the parameter search. Upon convergence of the search, the Jacobian contains information about the local topology of the χ^2 function that, under certain conditions, can be used to draw statistical inferences about the parameter uncertainties, the relative importance of each parameter, and correlations amongst the fitting parameters.

Basic Assumptions Required for Statistical Analysis

In order to obtain useful expressions for parameter confidence regions in terms of the Jacobian matrix and the experimental noise, it is necessary to make some assumptions about the experimental data. Specifically, the data are assumed to have the form

$$h_i = \phi_i(\mathbf{x}^\dagger) + \varepsilon_i \qquad [19]$$

where the ε_i are assumed to be normally distributed random values with a standard deviation σ_ε. A second, more profound assumption expressed in Eq. [19] is that the experimental spectrum underlying the noise can be exactly reproduced by the model $\phi(\mathbf{x})$ for some specific set of parameters $\mathbf{x}^\dagger$.

Under these assumptions it is possible to express an uncertainty bound for each individual parameter and joint confidence regions for subsets of parameters using information obtained from the covariance matrix c, which is estimated from the Jacobian matrix evaluated at the residual minimum $\mathbf{x}^*$ as

$$\mathsf{c}(\mathbf{x}^*) = \mathsf{C}^{-1}(\mathbf{x}^*) = [\mathsf{J}^{\mathrm{T}}(\mathbf{x}^*)\mathsf{J}(\mathbf{x}^*)]^{-1}. \qquad [20]$$

Equation [20] implies additional assumptions that are used to derive the statistical expressions given below, namely (i) that the residual function $\mathbf{f}(\mathbf{x}^*)$ is well approximated by a multidimensional Taylor expansion around the minimum which includes the curvature matrix $\mathsf{C}(\mathbf{x}^*)$ in its second-order term, and (ii) that $\mathsf{C}(\mathbf{x}^*)$ is in turn well approximated by the product $\mathsf{J}^{\mathrm{T}}(\mathbf{x}^*)\mathsf{J}(\mathbf{x}^*)$. These assumptions are violated when $\mathbf{f}$ exhibits a highly nonlinear dependence upon the parameters in $\mathbf{x}$, i.e., when there is significant local curvature of the surface at $\mathbf{x}^*$, so that higher-order terms of the Taylor expansion are nonnegligible.

Parameter Uncertainties and Confidence Regions

For minimizations of the weighted residuals, the confidence bounds for each parameter at the χ^2 minimum are given by (*37*)

$$\delta x_i = \pm\sqrt{\Delta\chi^2_p c_{ii}} \qquad [21]$$

and the boundary of the joint confidence region for a given subset of parameters in the p-dimensional vector $\mathbf{x}_p$ is given by

$$(\mathbf{x}_p - \mathbf{x}_p^*)\cdot\mathsf{C}_p\cdot(\mathbf{x}_p - \mathbf{x}_p^*) \leqslant \Delta\chi^2_p, \qquad [22]$$

where C_p denotes the curvature matrix corresponding to the subset of p parameters in x_p. In Eqs. [21] and [22], the quantity $\Delta\chi^2_p$ is calculated from the χ^2 probability distribution function for p degrees of freedom and a prespecified confidence interval α such that the probability that the true solution vector lies within the region $\chi^2(\mathbf{x}_p) - \chi^2(\mathbf{x}_p^*) \leqslant \Delta\chi^2$ is given by α. It should be noted that $p = 1$ in Eq. [21], whereas in Eq. [22], p is the number of parameters in the subset for which the confidence region is being calculated. This approach was previously used (*15–18*) in slightly modified form (*35, 45*).

For unweighted residual minimizations, t-distribution statistics are used to estimate confidence bounds for each parameter, and F-distribution statistics are used to estimate joint confidence regions, as (*34*)

$$\delta x_i = \pm t^{\alpha/2}_{n-p} s\sqrt{c'_{ii}} \qquad [23]$$

$$(\mathbf{x}_p - \mathbf{x}_p^*)\mathsf{C}'_p(\mathbf{x}_p - \mathbf{x}_p^*) \leqslant ps^2F^{\alpha}_{p,n-p}, \qquad [24]$$

where n is the number of data points. The primes on $\mathbf{c}$ and C_p are added to distinguish them from the matrices in Eqs. [21] and [22], since the Jacobian matrix from which they are calculated is not scaled by the spectral noise distribution.

Parameters outside the set of nonlinear parameters may be included in this type of analysis as well, so long as it is possible to calculate the partial derivative of the residuals with respect to such parameters. For the scaling factors, the partial derivatives are simply the unscaled individual component spectra; similarly, the partial derivative of the residuals with respect to the shift parameter $\Delta\omega$ is simply the first derivative of the scaled total spectrum with respect to field (or frequency), which may be calculated numerically. Columns containing the appropriate first derivatives for separable parameters are appended to the Jacobian matrix, and their uncertainties and covariances calculated as for the nonlinear parameters.

If the underlying assumptions expressed in Eq. [19] are valid, then s^2/n will be an unbiased estimate of the average σ^2 (*34*). Thus, in the limit of very large sample sets, t and F statistics should give equivalent results to χ^2 statistics for normally distributed errors. For the simulated spectra utilized in this work, which do have normally distributed noise, the two approaches do indeed give nearly identical results. However, for real experimental data, we find that Eqs. [23] and [24] tend to give larger estimates than Eqs. [21] and [22] for both the uncertainty bounds and the confidence regions.

It is worth reemphasizing at this point that *none* of the expressions given above provide a valid estimate of parameter uncertainties in the cases that (a) the errors in the spectrum are not normally distributed, (b) the underlying spectrum cannot be exactly represented by the model function $\mathbf{f}$ at some point in the search parameter domain (i.e., one is fitting with an incorrect model), or (c) the residual function has high curvature or nonlinearity in the neighborhood of the minimum. Although we have found that the third of these difficulties can largely be avoided by proper scaling of the dynamic parameters (in log space), the large majority of experimental spectra fail to meet the first two criteria rigorously. It might therefore appear that Eqs. [21]–[24] are of limited utility for estimating experimental errors; however, it is still possible to obtain reasonable uncertainty estimates from them in many cases, as will be discussed under Examples and Discussion below.

Calculation of Two-Parameter Error Ellipses

Parameter confidence regions are most often used to examine correlation between only two parameters (an example is given under Examples and Discussion). Here, we give the expressions used to calculate two-parameter error ellipses from the output of the fitting program. For the ith parameter, the program reports the uncertainty bound δx_i calculated according to either Eq. [21] or Eq. [24]. The program also reports a matrix of the correlation coefficients r_{ij} between each pair of parameters. The covariance matrix c (or c') for the two-parameter subset is then constructed as

$$\begin{bmatrix} \delta x_i^2 & r_{ij}\delta x_i\delta x_j \\ r_{ij}\delta x_i\delta x_j & \delta x_j^2 \end{bmatrix} \equiv \begin{bmatrix} c_{ii} & c_{ij} \\ c_{ij} & c_{jj} \end{bmatrix}. \qquad [25]$$

Then, defining $\mathbf{dx} = \mathbf{x} - \mathbf{x}^*$, Eq. [22] for two variables can be written as

$$\Delta\chi^2 = \begin{pmatrix} dx_i \\ dx_j \end{pmatrix} \cdot \begin{bmatrix} c_{ii} & c_{ij} \\ c_{ij} & c_{jj} \end{bmatrix}^{-1} \cdot \begin{pmatrix} dx_i \\ dx_j \end{pmatrix} \qquad [26]$$

which may be expanded as

$$\Delta\chi^2 = \frac{1}{D}\,(dx_j^2c_{ii}^2 - 2c_{ij}dx_i\,dx_j + dx_i^2c_{jj}^2) \qquad [27]$$

with $D = c_{ii}c_{jj} - c_{ij}^2$. Solving this equation for dx_2 gives the two boundaries of the error ellipse

$$dx_2 = \frac{1}{c_{11}}\,[c_{12}dx_1 \pm \sqrt{(c_{12}^2 - c_{11}c_{22})dx_1^2 + c_{11}\Delta\chi^2}] \qquad [28]$$

which has real roots for

$$dx_1 \leqslant \left|\frac{c_{11}\Delta\chi^2}{c_{11}c_{22} - c_{12}^2}\right|. \qquad [29]$$

In the case of large correlations ($r_{ij} > 0.9$), it is possible to use the major axis of the ellipse to define a linear relationship between the two parameters in the region of the solution:

$$dx_2 = \tan\zeta\; dx_1; \quad \tan 2\zeta = \left(\frac{2c_{12}}{c_{11} - c_{22}}\right). \qquad [30]$$

One approach to reducing the dimension of the search space in the case of highly correlated parameters is to use Eq. [30] to impose a linear constraint on the fitting parameters (*34*).

Quality of Fit Criteria

A wide variety of statistical tests and empirical criteria have been used to measure the quality-of-fit for various minimization procedures. Here, we briefly review a selection of the criteria that have been applied to EPR spectra in the literature and compare them for application to slow-motion spectra.

Reduced chi-squared test. The χ^2 given by Eq. [8b] at the minimum of the weighted residuals offers a semiquantitative measurement of the goodness-of-fit, since the true χ^2 should be distributed about the value $n - p$ with a standard deviation of $\sqrt{2(n-p)}$, where n is again the number of points and p is the number of fitting parameters. Thus, for an ''ideal fit,'' i.e., one in which the only source of deviation between data and calculation is normally distributed spectral noise, the *reduced* χ^2, given by $\chi^2_{\text{red}} = \chi^2/(n-p)$, should be distributed about unity with a standard deviation of $\sqrt{2/(n-p)}$.

This criterion is fairly strict. In practice, a χ^2_{red} approaching unity can readily be obtained using simulated datasets with normally distributed pseudorandom noise, but it is very difficult to achieve for experimental data, especially when global fits to two or more spectra are involved. Any deviations between the data and the calculated fit that are detectably larger than the spectral noise level will produce a relatively large χ^2_{red}. In practice, fits with χ^2_{red} values as large as 50–100 can still appear reasonable to the eye (cf. Examples and Discussion).

Quality-of-fit parameter. A related quantity that can be used to estimate the quality-of-fit is Q (*37*), the cumulative probability distribution function for χ^2, which assumes values from 0 to 1. Essentially, Q is the probability of finding the experimental χ^2 based on the expected distribution of χ^2 for the available degrees of freedom. If Q is very small, the discrepancies between data and model are unlikely to be random fluctuations.

Q has proven to be of almost no practical utility for slow-motion spectral fitting, mainly because it goes to zero very rapidly as χ^2 deviates from $n - p$. In practice, Q for many of the experimental data sets we have examined is smaller than the floating-point machine precision of our workstation, and it can be very small even for fits that by other criteria are quite acceptable. Fits to simulated data sets typically give Q values of 0.1–0.6; however, any value above 0.001 should be considered an excellent fit to experimental data.

Correlation coefficient, r. A more commonly used criterion is the correlation between the data and the calculated spectrum. Alternative forms of the correlation coefficient r that have been used in EPR analysis include the linear correlation coefficient (*19, 46*)

$$r_{\text{lin}} = \frac{\Sigma_i (h_i - \bar{h})(\phi_i - \bar{\phi})}{[\Sigma_i (h_i - \bar{h})^2 \, \Sigma_i \, (\phi_i - \bar{\phi})^2]^{1/2}} \qquad [31]$$

and the correlation coefficient used by Jackson (*47*)

$$r = \frac{\Sigma_i \, h_i \phi_i}{(\Sigma_i \, \phi_i^2)^{1/2}}, \qquad [32]$$

where $\bar{h}$ and $\bar{\phi}$ respectively are the mean values of the data and the calculated function. The linear correlation coefficient given in Eq. [31] ranges between −1 and +1, with 1 corresponding to a ''perfect'' fit, whereas that of Eq. [32] depends upon the scale of the spectrum.

For slow-motion EPR spectra, the correlation coefficient appears to be a relatively weak criterion for an acceptable fit. Some fits that exhibit obvious flaws and fail by other quantitative criteria have still given linear r_{lin} values greater than 0.95. The fits to simulated spectra with normally distributed noise shown below all gave r_{lin} values above 0.999.

Residual index, R. This type of statistic is similar to one used in X-ray crystallographic analysis and has been applied to magnetic-resonance spectral fitting in a few cases. Alternative forms for R that have been applied in EPR include (*48*)

$$R_{\text{abs}} = \frac{\Sigma_i \, |h_i - \phi_i|}{\Sigma_i \, |h_i|} \qquad [33]$$

and (*42*)

$$R_{\text{sqr}} = \frac{\Sigma_i \, (h_i - \phi_i)^2}{\Sigma_i \, (h_i - \bar{h})^2}, \qquad [34]$$

both of which approach zero for a perfect fit. These expressions essentially give the total deviation between the data and model as a fraction of the integrated signal amplitude in some form. A value of $R_{\text{abs}} < 0.05$ has been suggested as a criterion for a fit that is ''quite good'' (*14*). The value of R_{sqr} should be smaller for an equivalent fit; $R_{\text{sqr}} \approx 0.01$ or less has been suggested as an appropriate criterion for a good fit (*42*). In practice, the actual cutoff values are likely to depend on the experimental signal-to-noise ratio. Fits to simulated slow-motion data with an S/N of 100 give R_{abs} values of 0.04–0.06, whereas the best fits to experimental nitroxide spectra with comparable signal-to-noise typically produce $R_{\text{abs}} \approx 0.1$. The residual index thus appears to be a more useful empirical measure than the linear correlation coefficient r, but much less rigorous than the χ^2 and related tests.

Condition of Fit

The Jacobian matrix utilized in the Levenberg–Marquardt and related quasi-Newton minimization algorithms possesses some useful diagnostic features for identifying ill-conditioned fit, i.e., when one or more parameters do not significantly affect χ^2, or when two parameters are linearly dependent. In the first of these cases, J will have a zero column, whereas in the second case, two columns of J will differ by only a multiplicative factor. Either of these conditions leads to a zero determinant of the full J matrix, so that Eq. [9]

cannot be solved without removing the appropriate columns of J. Although these extreme pathologies are rarely encountered in practice, the full implementation of the *QR* decomposition algorithm with pivoting can be used to identify when J is ''nearly'' rank-deficient according to some predefined tolerance.

EXAMPLES AND DISCUSSION

We have developed two general computer programs, NLSL and NL2DC (cf. end of this section), based on the models and methods outlined above, for least-squares analysis of CW EPR and 2D EPR data sets, respectively. Magnetic parameters as well as dynamic parameters associated with a selection of different models may be optimized. Both programs also allow for global analysis of a series of spectra obtained by varying one of the experimental conditions (e.g., sample orientation or 2D mixing time) and both incorporate the separated-variables approach to reduce the parameter search space. In addition, program NLSL can also carry out fitting of multiple components in either a single or a global minimization.

Among the more significant challenges for ''fully automated'' spectral analysis are the determination of initial search parameter values, and the presence of local minima in the function to be minimized. The tasks of choosing a starting point and avoiding false minima become particularly difficult to automate when a large number of fitting parameters or multiple data sets are involved. The NLSL and NL2DC programs are therefore not intended to carry out fully automated least-squares analysis; rather, they are based on the design philosophy that the optimal fitting algorithm is one which best combines the intuition of the user with the speed and (presumably) the objectivity of a computer.

In order to achieve such a combination, the programs are based on a simple command language interface that enables *real-time, interactive* modification of the parameter values, search parameter set, or fitting procedure itself. The minimization may be interrupted at any time, modified, and restarted. In most operating environments, the NLSL program also displays the current best-fit function graphically, providing a convenient means of choosing starting parameters.

Whereas programs for real-time, interactive fitting of motionally narrowed spectra have been available for some time (*49–52*), such an approach had been prohibitive for slow-motion spectra because of the much larger computational burden involved. Critical steps in the extension of such methods to slow-motion analysis have included (i) the utilization of sparse matrix methods such as the Lanczos algorithm and related methods (*11–13*), particularly for obtaining the eigenvectors required in 2D calculations (*28*); (ii) minimization of the basis sets required (*27*); (iii) the introduction of an efficient modification of the basic Levenberg–Marquardt method; and (iv) efficient optimization of separable parameters in order to reduce the dimension of the parameter search space, as described under Methods. These enhancements have improved the efficiency of the calculation to the point where it is possible to carry out routine interactive analysis on an IBM RS-6000 RISC laboratory workstation operating under the AIX (UNIX) system, even for relatively large-scale problems. For example, a series of global minimizations for liquid crystal spectra obtained at two director orientations for eight different temperatures takes about 20 minutes to complete, operating interactively. For four fitting parameters, this corresponds to approximately 1500 individual spectral calculations, or about 750 ms per calculation. The program may also be run in production mode by specifying the fitting commands in a script file once an appropriate set of conditions for the minimizations has been identified.

In the examples that follow, we illustrate some of the more significant applications of slow-motion EPR spectral fitting by the Levenberg–Marquardt method. In many cases, the data sets employed are simulated by adding normally distributed noise to a theoretical spectral lineshape. In addition to their illustrative purpose, such spectra provide a useful gauge of how accurately the least-squares procedure can recover the true model parameters. The idealization of the data also allows an assessment of the intrinsic resolution of different types of spectra and fitting procedures with respect to the model parameters. Where appropriate, examples using real experimental data will be given, particularly with reference to procedures for statistical analysis of the fitting parameters under nonideal conditions.

Parameter Correlation: Resolution of Rotational Anisotropy

Although correlations amongst the search parameters are widely neglected in EPR spectral fitting, they occur quite frequently and provide an important criterion for assessing the reliability of the information derived from the fit. This is particularly true for slow-motion spectra, where the large number of available model parameters frequently leads to ''overfitting'' of the data; that is, more quantities may be sought in the minimization procedure than can accurately be determined from the available data. Thus, the appearance of large correlations amongst fitting parameters can be a useful diagnostic for identifying such cases and determining the intrinsic resolution of the data with respect to the model parameters, as we now illustrate.

A major goal in the analysis of slow-motion EPR spectra is to determine reliably the anisotropy of the spin probe rotational motion. This may be particularly important for nitroxide-labeled macromolecules which can exhibit highly anisotropic reorientational behavior. The sensitivity of the ESR spectrum to rotational anisotropy may be expected to

vary as one proceeds from the limit of motional narrowing through the slow-motion region. For example, in the case of an axially symmetric rotational diffusion tensor, the average rotational rate constant $\bar{R} \equiv \sqrt[3]{R_\perp^2 R_\parallel}$ and the rotational anisotropy $N \equiv R_\parallel / R_\perp$ can uniquely be determined from the homogeneous linewidths of the three ^{14}N hyperfine lines in the motionally narrowed spectrum, assuming that the magnetic (g and A) tensors of a nitroxide are known (*53–55*). As the motion slows, the degree to which $\bar{R}$ and N may be independently determined decreases until the rigid limit is reached, at which point they cannot be measured from the spectrum at all. In the incipient rigid limit, only the faster components of anisotropic motion are detectable; for example, with $N > 1$, the spectrum will become more sensitive to $R_\parallel$ and less sensitive to $R_\perp$ as the motion slows. At the slowest motions, there will be a range of combinations of $\bar{R}$ and N that give the correct value of $R_\parallel$ and therefore fit the data reasonably well, but produce an incorrect value for $R_\perp$. Such situations are detected by a high correlation coefficient between $\bar{R}$ and N when these components are simultaneously varied in a minimization.

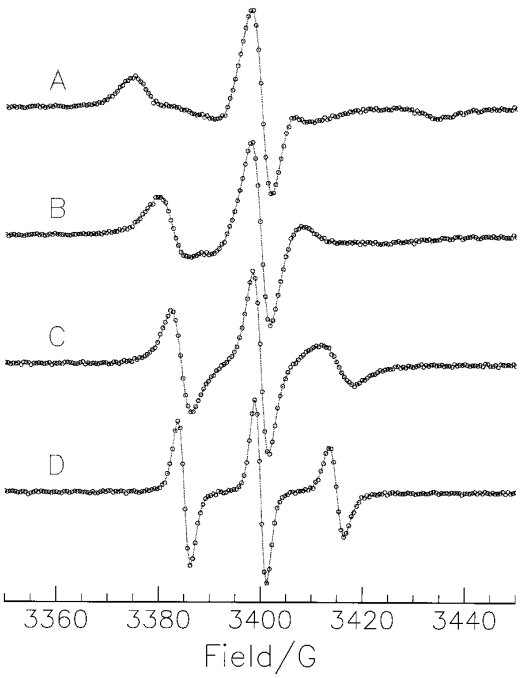

FIG. 2. Series of simulated X-band slow-motion CW EPR spectra (open circles) calculated using magnetic parameters of CSL (cf. Table 2) and assuming anisotropic Brownian rotational diffusion in an isotropic solvent for a range of rotational diffusion rates. The value of $R_\perp$ used in each calculation was (A) 3.0×10^6, (B) 1.0×10^7, (C) 3.0×10^7, and (D) 1.0×10^8 s^{-1}. The rotational anisotropy $N = R_\parallel / R_\perp$ was 10 for all the spectra. Normally distributed random noise was added to each spectrum to give a signal-to-noise ratio of 100. Solid lines show nonlinear-least-squares fits to the spectra as described in the text; parameters obtained from the fits are given in Table 3.

TABLE 2
Summary of Magnetic Tensor Parameters Used for the Simulated Spectra Used in This Work

	Data set				
Parameter	1	2	3[a]	4	5
gxx	2.0089	2.0096	2.0089	2.0089	2.0081
gyy	2.0021	2.0063	2.0058	2.0063	2.0061
gzz	2.0058	2.0021	2.0021	2.0021	2.0024
axx	5.6	5.4	4.9	5.0	5.6
ayy	33.8	5.1	4.9	5.0	5.3
azz	5.3	33.6	33.0	33.0	33.8
rx	[b]	7.326	7.699	[c]	7.930
ry	[b]	7.628	7.699	[c]	7.930
rz	[b]	8.477	8.398	[c]	9.362
betad	15		30		20
b0	3400	[d]	3405	3400	3262
gib0	2.0	[e]	2.0	2.0	1.125
gib2			0.5		
c20			2.50		4.520
c22			−1.25		1.432

Note. Explanation of parameter names and units for each parameter are given in Table 1.

[a] Calculated for psi = 0, 30, 60, and 90, and for a MOMD spectrum with nort = 15.

[b] A set of rprp and rpll values were used as indicated in Table 3.

[c] A set of rbar values was used for each of two components as specified in Table 7.

[d] Calculated for 89000 and 3400.

[e] 3.0 G at b0 = 89000; 2.0 G at b0 = 3400.

We illustrate this effect for simple Brownian rotational diffusion in an isotropic solvent with an axially symmetric anisotropic diffusion tensor. Figure 2 shows a series of simulated X-band slow-motion CW EPR spectra which were calculated using magnetic parameters intended to represent the cholestane spin label (CSL) (set 1 in Table 2), assuming Brownian rotational diffusion in an isotropic solvent over a range of rotational diffusion rates ($R_\perp = 3.0 \times 10^6$, 1.0×10^7, 3.0×10^7, and 1.0×10^8 s^{-1}) with a constant anisotropy of $N = 10$. Normally distributed random noise was added to each spectrum to produce a signal-to-noise ratio of 100.

For each of the spectra shown, least-squares minimizations were carried out with respect to the inhomogeneous Gaussian broadening parameter and either the average and the anisotropy of the diffusion tensor components ($\log_{10}\bar{R}$/s^{-1} and $\log_{10}N$) or the axial Cartesian components ($\log_{10}R_\perp$/s^{-1} and $\log_{10}R_\parallel$/s^{-1}). The χ^2_{red} of each fit (shown by solid lines in Fig. 2) was close to unity, and the starting parameter values were recovered within the reported uncertainties (cf. Table 3).

The most significant aspect of the fits shown in Fig. 2 is the increasing degree of correlation that is observed between

TABLE 3

Best-Fit Parameters and Correlation Coefficients for Each of the Simulated Spectra Shown in Fig. 2, Found by Varying the Axial Components (rpll, rprp) or the Spherical Components (rbar, n) of the Rotational Diffusion Tensor

Run	χ^2_{red}	Parameter	Starting value	Fit value	r^a	Parameter	Starting value	Fit value	r^b
1	1.089	rprp	6.477	6.500 ± 0.083	−0.362	rbar	6.810	6.825 ± 0.054	−.961
		rpll	7.477	7.474 ± 0.017		n	1.000	0.974 ± 0.091	
2	1.142	rprp	7.000	6.997 ± 0.029	−0.290	rbar	7.333	7.330 ± 0.018	−.691
		rpll	8.000	7.995 ± 0.018		n	1.000	0.998 ± 0.038	
3	0.967	rprp	7.477	7.478 ± 0.013	−0.062	rbar	7.810	7.810 ± 0.014	0.503
		rpll	8.477	8.473 ± 0.034		n	1.000	0.995 ± 0.038	
4	1.070	rprp	8.000	7.975 ± 0.030	−0.445	rbar	8.333	8.314 ± 0.021	0.250
		rpll	9.000	8.991 ± 0.062		n	1.000	1.016 ± 0.080	

[a] Correlation coefficient between parameters rprp and rpll.
[b] Correlation coefficient between parameters rbar and n.

log $\bar{R}/s^{-1}$ and log N toward the slower motions. For the fastest motion shown, the correlation is only 0.25, whereas a very strong negative correlation of −0.96 is observed at the slowest motion. In contrast, no significant correlation appears throughout the range of motional rates when the minimization is carried out with respect to log $R_\perp/s^{-1}$ and log $R_\parallel/s^{-1}$: the correlation between these two parameters remains below about 0.45 in magnitude. The differences between the two cases for the slowest motion are illustrated more graphically in Fig. 3, which plots the χ^2 contours for the spectrum corresponding to the slowest motion (top of Fig. 2) vs parameter values for both the (log $\bar{R}/s^{-1}$, log N) and the (log $R_\perp/s^{-1}$, log $R_\parallel/s^{-1}$) pairs. The dotted ellipses show joint confidence regions for the fitting parameters calculated from Eq. [28] using $\Delta\chi^2 = 5.99$ (95% significance level for two parameters). From this figure, it can be seen that the strong correlation between log $\bar{R}/s^{-1}$ and log N produces an elongated, tilted error ellipse that conforms reasonably well to the local topology indicated by the χ^2 contours. In contrast, the long axis of the error ellipse for log $R_\perp/s^{-1}$ vs log $R_\parallel/s^{-1}$ is not tilted, but lies approximately parallel to the $R_\perp$ axis, again conforming to the χ^2 contours and indicating that there is little correlation between the $R_\perp$ and $R_\parallel$ parameters.

In general, the appearance of large (>0.9) correlations between fitting parameters indicates that the experimental data are insensitive to one of the correlated parameters, or some combination of them. In the example shown here, the poorly determined quantity is readily identified by utilizing a different linear combination of tensor quantities to reduce or eliminate the correlation. In the absence of strong correlation, the relative uncertainties for log $R_\perp/s^{-1}$ vs log $R_\parallel/s^{-1}$ are easily interpreted and clearly show that the poorly determined quantity is $R_\perp$, consistent with the expectation that the spectrum should be less sensitive to the slower components of R as the rigid limit is approached.

Another use of the correlation coefficients as a measure of spectral resolution at different EPR frequencies is shown in Fig. 4. Simulated spectra were calculated for a nitroxide exhibiting a fully anisotropic rotational diffusion tensor (set 2 in Table 2) at 250 and at 9.5 GHz with added Gaussian noise to give $S/N = 100$ at both frequencies. The results of the least-squares analyses are given in Table 4. The minimization using the high-frequency data yields the starting parameters to within the reported uncertainties for all three elements of the rotational diffusion tensor. At the lower frequency, however, R_x and R_y are incorrectly estimated to be nearly equal, with a very substantial and unacceptable correlation coefficient of −0.98. The linear relationship derived from this correlation according to Eq. [30] gives a slope of −1, indicating that $\log_{10}R_x/s^{-1}$ and $\log_{10}R_y/s^{-1}$ can only vary so as to maintain a constant sum (i.e., the product R_xR_y is constant). Thus, the lower-frequency data are insensitive to the ratio R_x/R_y that measures nonaxial anisotropy in R. This result is consistent with our previous observations that high-frequency ESR spectra are more sensitive to the rotational anisotropy than are X-band spectra (*55*).

The appearance of correlations amongst the elements of a given tensor generally indicate a limitation in the resolution of the experimental data with respect to that tensor. Such correlations can frequently be avoided by (i) fixing the value of one of the quantities involved, (ii) restricting the tensor to a higher symmetry, or (iii) using a different representation for the tensor as illustrated above. By analogy, correlations between coefficients of the ordering potential usually indicate that the higher-order term is not uniquely resolvable and should not be varied in the minimization.

Although the strong correlations observed in slow-motion spectral fitting occur most frequently between similar parameters, correlations between different types of parameters that do not lend themselves so easily to a physical interpretation can also arise. The most common such correlations occur

between rotational diffusion tensor elements and the inhomogeneous linewidth, diffusion tilt angles, or higher-order (nonaxial) terms in the orienting potential. In such cases, it is generally necessary to remove one of the correlated variables from the list of search parameters. Another possible solution is to impose a linear constraint such as Eq. [30] upon the correlated variables, thus effectively reducing the dimension of the search space.

Global Spectral Analysis: Aligned vs Unaligned Samples

An important aspect of the nonlinear-least-squares program described here is its ability to perform global analysis for a series of spectra obtained by varying one of the parameters in a given EPR experiment. Global analysis of multiple datasets has found widespread use in the study of time-resolved optical anisotropy decays (*56, 57*), and has recently been applied to the simultaneous analysis of fluorescence anisotropy decay and saturation-transfer EPR spectra using the basic Levenberg–Marquardt algorithm (*20*). Such an approach has also been utilized in the analysis of rigid-limit EPR spectra obtained at different frequencies (*48*). It has previously been applied by us to fit slow-motion EPR spectra taken at different temperatures to obtain the Arrhenius parameters of the rotational reorientation directly (*18*), and to fit liquid crystal spectra at different sample orientations simultaneously (*15, 58*). Additionally, a series of 2D EPR spectra for different mixing times have been fitted simultaneously (*21, 22*).

The main advantage of global analysis is that it places additional experimental constraints upon the residual function in the parameter space, thus better defining the region of

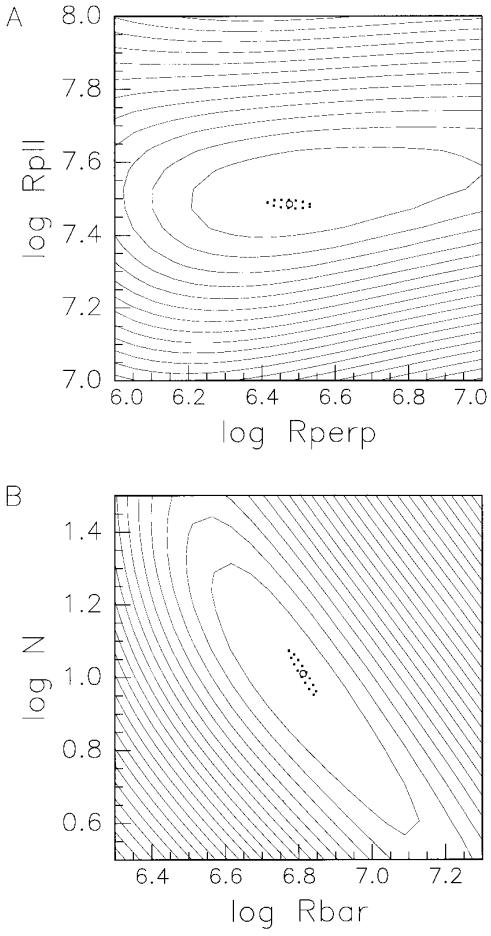

FIG. 3. Chi-squared contour maps plotted as a function of parameter values for least-squares fits to the spectrum at the top of Fig. 2. The variable parameters were (A) $\log_{10}$ of the rotational diffusion rates about the axis parallel (`rpll`) and perpendicular (`rprp`) to the long axis of the diffusing probe, or (B) $\log_{10}$ of the average rotational diffusion rate (`rbar`) and rotational anisotropy (n). The correlation coefficients were -0.36 and -0.96 for (A) and (B), respectively. Dotted ellipses show joint 95% confidence regions for the fitting parameters calculated for $\Delta\chi^2 = 6.30$ as described in the text. Note the agreement between the local contour predicted by the error ellipse and χ^2 contours.

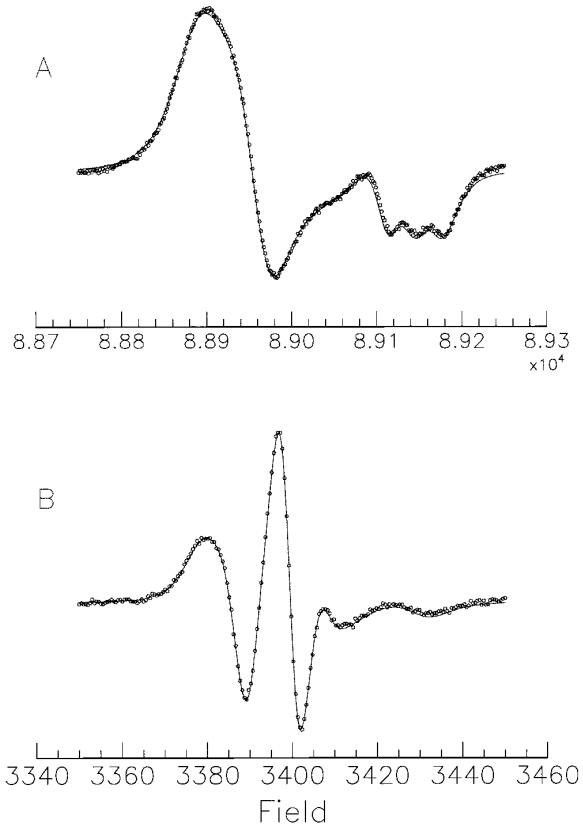

FIG. 4. Least-squares fits to simulated spectra (open circles) of a nitroxide at (A) 250 GHz and at (B) 9.5 GHz for a probe with fully anisotropic rotational diffusion (set 2, Table 2). Note the large difference in field scales. The rotational anisotropy was $R_x/R_y = 2.0$ and $\sqrt{R_xR_y}/R_z = 10.0$. Least-squares spectra corresponding to the parameters given in Table 4 are shown by the solid lines.

TABLE 4
Rotational Diffusion Components Measured from Least-Squares Analysis of Simulated 250 and 9.5 GHz Spectra Shown in Fig. 4 for a Nitroxide with a Fully Anisotropic Rotational Diffusion Tensor

	Parameter	Starting value	Least-squares value	Correlation matrix		
				rx	ry	rz
250 GHz	rx	7.326	7.314 ± 0.012	1	−0.353	0.800
	ry	7.628	7.667 ± 0.039		1	−0.475
	rz	8.477	8.470 ± 0.015			1
9 GHz	rx	7.326	7.448 ± 0.018	1	−0.984	−0.382
	ry	7.628	7.540 ± 0.017		1	0.212
	rz	8.477	8.088 ± 0.033			1

allowed solutions and reducing otherwise high correlations amongst the parameters. The principles of global analysis may be applied to many different types of series in which a single EPR parameter is varied. We have implemented a very general approach to this strategy: in principle, it is possible to carry out a global fit with respect to a series in any parameter that can also be varied as a search parameter. In practice, not every EPR parameter is practical as a series variable; the more useful series variables include (i) the director tilt angle for studying an ordered system at a series of orientations; (ii) the spectrometer field, for fitting spectra measured at a series of different EPR frequencies; (iii) the mixing time in a 2D exchange (2D ELDOR) experiment; and (iv) different coherence pathways for a given 2D spectrum.

In the following example, global fitting to spectra from

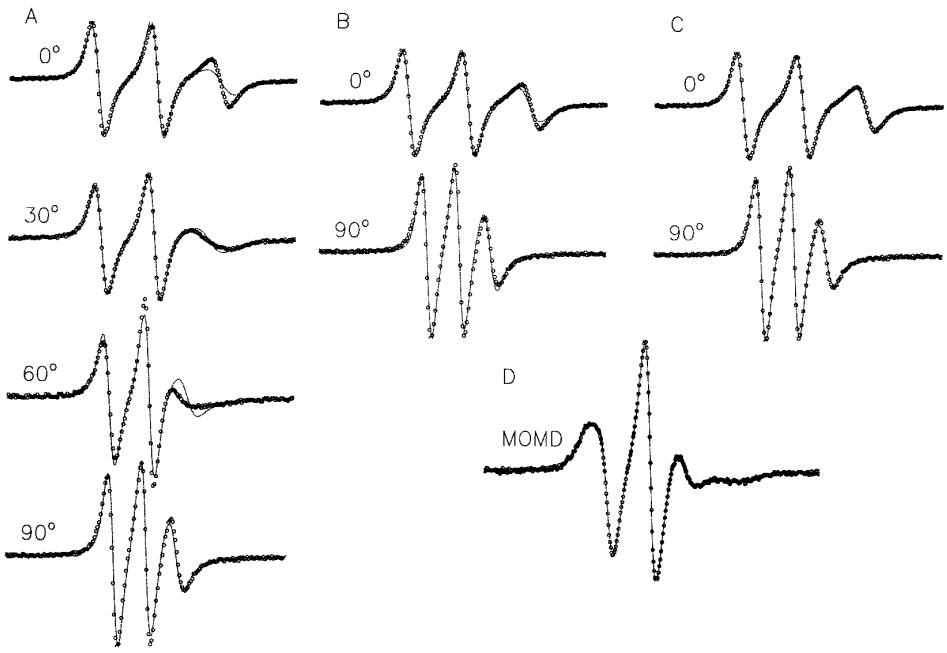

FIG. 5. Least-squares fits to simulated X-band EPR slow-motion spectra (open circles) for the 5-PC spin probe (set 3, Table 2) in an aligned membrane at director tilt angles of $\Psi = 0°$, 30°, 60°, and 90°, and for macroscopically disordered membrane domains (MOMD) model. Minimizations were carried out using a mismatched motional model with $\beta_d = 0$; shown are (A) global fit to all orientations of the aligned membranes; (B) global fit for $\Psi = 0°$ and 90° only; (C) individual fits for $\Psi = 0$ and 90°; and (D) individual fit to the MOMD spectrum. Fits using the correct model and least-squares parameters given in Table 5 did not deviate above the noise level. Solid lines show nonlinear-least-squares fits of a mismatched model (see text) using least-squares parameters given in Table 6. Note that the mismatched model leads to significantly greater discrepancies in the global fit to four orientations.

aligned samples obtained at a series of orientations will be illustrated. This approach, in combination with methods that provide a high degree of membrane alignment such as pressure-annealing (*59*) and isopotential spin-dry ultracentrifugation (*58, 60*), has been essential for enhancing the structural and dynamic information that can be obtained from spin-labeled membrane systems by analysis of slow-motion EPR spectra (*16, 59, 61*).

Global fits to spectra from aligned membranes will also be compared to fits to spectra from macroscopically disordered samples, i.e., using the MOMD model. Many membrane systems are not amenable to alignment by the methods just mentioned, and in such cases one is restricted to randomly oriented dispersion samples with an accompanying loss of resolution. It is therefore important to be able to utilize the fitting procedure to assess the quality of the information obtained from unaligned vs aligned membranes systems.

Figure 5 presents simulated data for a lipid spin probe in macroscopically aligned membranes at director tilt angles of 0°, 30°, 60°, and 90° as well as in macroscopically disordered

TABLE 5

Comparison of Parameter Correlation Matrices for Single-Orientation and Global Least-Squares Fits to Simulated Spectra from Aligned Membranes at Various Director Tilt Angles, and to a MOMD Spectrum (See Text for Definition)

Type of fit	Parameter	Least-squares value	Correlation coefficients						
			gib0	gib2	rbar	n	betad	c20	c22
0° only	gib0	2.020 ± 0.061	1		0.6037	0.5534	−0.1963	−0.3904	0.5554
	gib2								
	rbar	7.936 ± 0.005			1	0.0334	−0.7258	0.1053	−0.0015
	n	0.677 ± 0.034				1	0.1695	−0.955	0.991
	betad	29.931 ± 0.151					1	−0.1261	0.2725
	c20	2.512 ± 0.022						1	−0.921
	c22	−1.272 ± 0.051							1
90° only	gib0	2.508 ± 0.029	1		0.666	−0.142	−0.395	−0.403	0.325
	gib2								
	rbar	7.931 ± 0.007			1	−0.119	−0.518	−0.375	0.187
	n	0.701 ± 0.069				1	0.798	0.592	−0.732
	betad	30.000 ± 0.407					1	0.819	−0.767
	c20	2.488 ± 0.084						1	−0.945
	c22	−1.230 ± 0.116							1
0° and 90°	gib0	2.007 ± 0.031	1	−0.622	0.597	0.114	−0.255	0.342	0.318
	gib2	0.507 ± 0.021		1	−0.236	l−0.516	−0.133	−0.743	0.674
	rbar	7.935 ± 0.004			1	−0.404	−0.758	0.290	−0.484
	n	0.693 ± 0.015				1	0.504	−0.043	0.181
	betad	29.948 ± 0.136					1	0.120	0.302
	c20	2.507 ± 0.017						1	−0.852
	c22	−1.261 ± 0.030							1
0°, 30°, 60°, 90°	gib0	1.999 ± 0.033	1	−0.836	0.514	−0.138	−0.154	−0.136	0.284
	gib2	0.517 ± 0.036		1	−0.287	−0.009	0.063	0.047	−0.123
	rbar	7.933 ± 0.002			1	−0.022	−0.343	0.156	−0.095
	n	0.688 ± 0.001				1	0.024	0.119	−0.192
	betad	29.990 ± 0.145					1	0.504	−0.127
	c20	2.504 ± 0.025						1	−0.839
	c22	−1.256 ± 0.035							1
MOMD	gib0	1.772 ± 0.344	1	−0.964	0.270	0.213	0.340	−0.040	0.323
	gib2	0.847 ± 0.332		1	−0.198	−0.123	−0.333	−0.068	−0.205
	rbar	7.937 ± 0.021			1	−0.222	−0.676	0.153	−0.312
	n	0.556 ± 0.020				1	0.339	−0.887	0.926
	betad	30.333 ± 1.918					1	−0.103	0.544
	c20	2.612 ± 0.183						1	−0.858
	c22	−1.335 ± 0.349							1

Note. The simulated spectra are shown in Fig. 5. (Least-squares spectra for the parameters given below are not visible in the figure.)

vesicles, but using the same set of magnetic and dynamic parameters. The simulation parameters were chosen to typify the 5-PC probe (set 3 in Table 2) including a substantial diffusion tilt angle, $\beta_d \equiv 30°$. The MOMD spectra were obtained using the sinψ-weighted average of the spectra calculated for 15 different values of ψ. Normally distributed noise was added to give a signal-to-noise ratio of 100 for all of the spectra.

The spectra corresponding to aligned membranes were fitted in several different ways, including (i) separate fits for $\psi = 0°$ and $\psi = 90°$, (ii) global fits for $\psi = 0°$ and $\psi = 90°$, and (iii) global fits for $\psi = 0°$, 30°, 60°, and 90°. The parameters varied in each minimization were `gib0`, `gib2`, `rbar`, `n`, `c20`, `c22`, and `betad` (cf. Table 1). For the individual fits to the $\psi = 0°$ and $\psi = 90°$ spectra, $\Delta^{(2)}$ was not varied since the orientation dependence of the inhomogeneous linewidth cannot be determined from a single-orientation spectrum. The same parameters were varied in the fit to the MOMD spectrum. For all of the minimizations, including the MOMD fit, the reduced χ^2 was near unity.

The correlation matrices of the variables for each of the fits listed in Table 5 clearly indicate that, even for aligned samples, the single-orientation spectra contain insufficient information to permit unambiguous determination of all of the model parameters. This is especially apparent in the $\psi = 0°$ fit, which exhibits correlations exceeding 0.9 between the rotational anisotropy N and both orienting potential coefficients, as well as between the potential coefficients themselves. The correlations observed in both the $\psi = 90°$ fit and the MOMD fit are somewhat less severe: both exhibit high correlation between the potential coefficients, and the fit to the MOMD spectrum also exhibits correlation between the $\Delta^{(0)}$ and $\Delta^{(2)}$ Gaussian inhomogeneous broadening parameters, which are incorrectly predicted by the minimization.

The very strong correlations disappear when global analysis is performed; for example, in the 0°, 90° global fit, there are no correlations greater than 0.90, and the overall correlation is almost entirely removed when spectra from four angles are included. Thus, if a sufficient number of different orientations are employed, it becomes possible to support the independent determination of as many as seven parameters, permitting a more detailed description of the molecular dynamics. In this context, it is not surprising that the fits to the MOMD spectrum exhibit somewhat less parameter correlation than that of the $\psi = 0°$ spectrum, since the spin probe is sampled over a wider range of molecular orientations in both cases. Also, it has been found in the past that the 90° orientation is the more sensitive for labeled lipids (*59, 61*).

An even greater contrast between the resolving power of multiorientation vs single orientation and MOMD data is apparent in the estimated parameter uncertainties from the different types of minimization. For most of the dynamic parameters of interest, the relative uncertainty of the parameter from each fit follows the order MOMD $>$ 0° $>$ (0°, 90°) $\approx$ (0°, 30°, 60°, 90°), consistent with the relative reliability of the fits indicated by the correlation matrices. Significantly, the uncertainties from the MOMD fit are 5–10 times larger than those from the global fits, owing to the loss of resolution in the macroscopically disordered sample.

Another revealing test of the relative information content of MOMD vs single-orientation spectra is to compare the results obtained for an incorrect or approximate theoretical model. This situation more closely approaches the conditions under which real experimental data are usually first analyzed, since at least the initial fitting of experimental data should be undertaken using the simplest possible model. To mimic

TABLE 6

Parameters Derived from Individual and Global Least-Squares Fits to Simulated Spectra from Aligned Membranes at a Selection of Director Tilt Angles, and to a Spectrum Exhibiting the Same Dynamic and Ordering Parameters, but with Macroscopic Disorder (MOMD)

	Starting value	0° only	90° only	0° and 90°	0°, 30°, 60°, 90°	MOMD
χ^2_{red}		13.569	6.231	33.011	32.852	1.831
r_{lin}		0.9958	0.9977	0.9877	0.9862	0.9982
R_{sqr}		0.106	0.082	0.168	0.174	0.083
`gib0`	2.000	4.249 ± 0.034	2.412 ± 0.038	3.999 ± 0.025	3.279 ± 0.024	0.645 ± 0.149
`gib2`	0.500			−0.628 ± 0.028	−0.311 ± 0.025	1.952 ± 0.143
`rbar`	7.932	9.092 ± 0.006	8.497 ± 0.024	8.271 ± 0.004	8.130 ± 0.003	8.051 ± 0.016
`n`	0.699	2.754 ± 0.002	1.542 ± 0.074	0.702 ± 0.001	0.719 ± 0.001	0.736 ± 0.036
`c20`	2.5	1.653 ± 0.002	1.633 ± 0.030	1.409 ± 0.004	1.308 ± 0.002	1.824 ± 0.014
`c22`	−1.25	0.760 ± 0.002	−1.642 ± 0.041	0.073 ± 0.014	0.062 ± 0.011	−1.243 ± 0.002

Note. A mismatch was introduced into the fitting model by constraining the diffusion tilt angle β_d to be 0 when the simulated spectra had $\beta_d = 30°$. The least-squares fits to each spectrum are shown by the solid lines in Fig. 5.

a defect in the motional model, the least-squares analysis was carried out as described above, but with the diffusion tilt angle β_d fixed at zero. The best fits from this experiment are shown in Fig. 5 and the least-squares parameters given in Table 6.

As the results in Table 6 indicate, the mismatch in motional model produces a very dramatic difference in the quality of the fits and the least-squares parameters obtained for the different types of minimization. The individual fits to the 0° and 90° spectra appear reasonably close as judged by their appearance and χ^2_{red}, but they lead to significant overestimates of the anisotropy and the average rotational diffusion rate. The global analyses are much more sensitive to the mismatch, as can be seen by the significant deviations between data and calculation on the left-hand side of Fig. 5. Surprisingly, the fit to the MOMD spectrum is rather good, with a χ^2_{red} approaching unity, even for the mismatched model. Even more significant is that the MOMD fit provides the best approximation to the starting parameters. Thus, in the case of a mismatched model, the MOMD spectrum appear to reflect the average behavior of the spin probe better, at the expense of resolution to certain details of the motional model.

Some other aspects of the comparison between aligned and unaligned membranes deserve comment. In the test case presented here, the magnetic, dynamic, and ordering parameters were assumed to be the same for both the MOMD spectrum and the spectra from oriented samples. For real membrane samples, local structure and dynamics may differ substantially between the two cases due to differences in membrane curvature or different phase behavior in aligned multilayers vs nonaligned vesicles (*58*). The results presented here show how information from the curvature matrix used in the Marquardt–Levenberg method can be used define the resolution of MOMD spectra and also indicate what may reliably be inferred about the differences between well-aligned multilayers and randomly oriented dispersions in comparing fits from these two types of samples.

Finally, it should be noted that there are some potential technical pitfalls peculiar to fitting MOMD spectra using the Levenberg–Marquardt or related quasi-Newton minimization methods. The problem lies in the calculation of the Jacobian, i.e. the partial derivatives of the spectra, which may not be fully averaged for a given number of orientations even though the spectrum itself is sufficiently averaged. This is because the partial derivative spectra at individual orientations exhibit features that are in general more numerous and sharper than the features in the original spectra, qualitatively resembling higher-derivative spectra with respect to the magnetic field. Because of this effect, significantly more orientations must be averaged to produce an acceptably smooth partial derivative than are needed to obtain an acceptable first-derivative spectrum. Figure 6 illustrates this effect: at the top of the figure is the first-derivative MOMD spectrum corresponding to the best-fit spectrum shown in Fig. 5, calculated using 10 orientations. The bottom two plots are the partial derivative of the same spectrum with respect to the c_{20} potential coefficient, calculated using 10 orientations (middle curve) and 20 orientations (bottom curve, scaled by a factor of 5). When only 10 orientations are used, there remain relatively large-amplitude oscillations in the middle of the partial derivative spectrum, which are nearly completely averaged away out when 20 orientations are used. Significant differences between the partial derivatives for 10 and 20 orientations were found for all of the dynamic and ordering parameters.

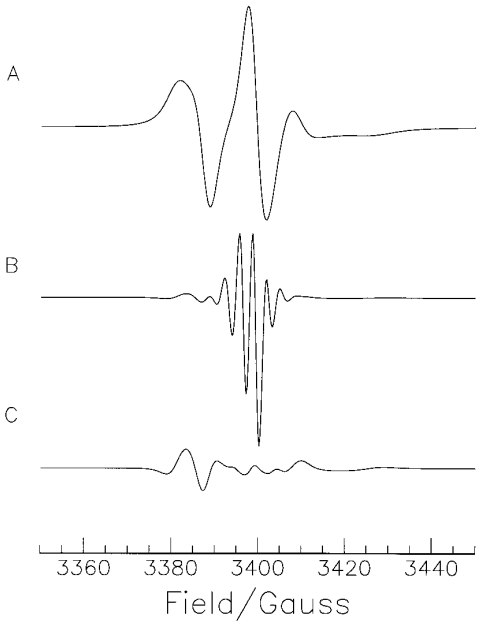

FIG. 6. (A) Calculated best-fit spectrum to the MOMD data shown in Fig. 5d with $\beta_d = 30°$, obtained by averaging 10 orientations. (B) Partial derivative of the spectrum in (A) with respect to the c_0^2 ordering potential coefficient, also calculated using 10 orientations. (C) Same partial derivative spectrum as in (B), but calculated using 30 orientations. The scale in (C) is expanded by a factor of five.

The practical consequence of an incompletely averaged partial derivative MOMD spectrum is that the resulting oscillations artificially increase the norm of the column of the Jacobian matrix corresponding to the given parameter. Since it follows from Eq. [9] that the step size for each parameter is inversely related to the norm of the corresponding column in $\mathbf{J}$, the incomplete averaging has the effect of slowing down the convergence along the parameter dimension. We have found that acceptable averaging of the partial derivative spectra generally requires about 2.0–2.5 times as many orientations as are necessary to obtain a smooth and completely averaged MOMD spectrum. The extra computation time required for the additional orientations is partially compensated by the improved convergence properties of the minimi-

zation with the correct partial derivative information. Correct partial derivatives are also essential for accurate statistical analysis; in this case, an artificially large gradient of the residual function with respect to a given parameter will result in an underestimate of the uncertainty in that parameter.

Multicomponent Fitting

A frequently occurring problem in EPR analysis is the separation and identification of individual components in a spectrum that contains more than one EPR species. Significant effort has been directed toward identifying components of EPR spectra that differ in their magnetic properties or chemical identities (*41, 42*).

Multicomponent spectra can also arise in the case of chemically identical species that exhibit different mobilities. One common approach to separating species based on their mobilities is the spectral subtraction technique, which has been extensively applied to characterize lipids associated with integral proteins in membrane systems (*62*). Spectral subtraction is most effective when the mobilities of the two species differ significantly and one of the components is in the motionally narrowed regime. It also requires that one obtain the spectrum for one of the components by itself, which is assumed not to change in the presence of a second component. Such methods become increasingly unreliable when there is significant overlap between the component spectra, or when more than two species are present. In this case, more sophisticated approaches are required, such as iterative spectral subtraction (*41*). These methods rely on the assumption that the individual component spectra are unchanged by the presence of other components in the system, an approximation that may be violated in cases where species with different mobilities are undergoing physical or chemical exchange. In the 2D format, one approach to resolving species with different mobilities has been the field swept spin-echo experiment (*63, 64*), which permits resolution of the species based on their different T_2 relaxation times.

The use of an orthogonal decomposition such as the *QR* factorization to quantify individual spectral components has the very useful feature that linear dependencies amongst the component spectra may be detected, and the number of components reduced accordingly. Most important, it affords a more rigorous and accurate means of quantifying the properties of the individual species than, e.g., spectral subtraction or trial-and-error. Another very useful feature is that error analysis may be performed on the scale factors just as for the other fitting parameters, providing estimates of uncertainty and correlation with other parameters. We illustrate these features in the following examples.

Figure 7 shows a series of simulated two-component X-band CW EPR spectra calculated for two isotropically diffusing nitroxides with identical magnetic parameters (cf. set 4 in Table 2) but different rotational diffusion constants, $\bar{R}_1$ and $\bar{R}_2$. The geometric mean of $\bar{R}_1$ and $\bar{R}_2$ was kept fixed at 3.00×10^7 and the ratio $\bar{R}_1/\bar{R}_2$ was 50, 20, 10, 5, 2, and 1.1, from top to bottom. The relative population of the more mobile component in the simulated spectra was maintained at 30% in all of the spectra, and random noise was added to produce a signal-to-noise ratio of 100. The values of the two diffusion rates and Gaussian inhomogeneous broadening parameters were varied in the least-squares minimizations, which utilized the nested linear-least-squares procedure described under Methods to determine scaling factors for each component. Two components could be fitted to every spectrum but that at the bottom ($\bar{R}_1/\bar{R}_2 = 1.1$), and the χ^2_{red} was close to unity in all cases. Relevant least-squares parameters from the fits are given in Table 7. The scaling factors were

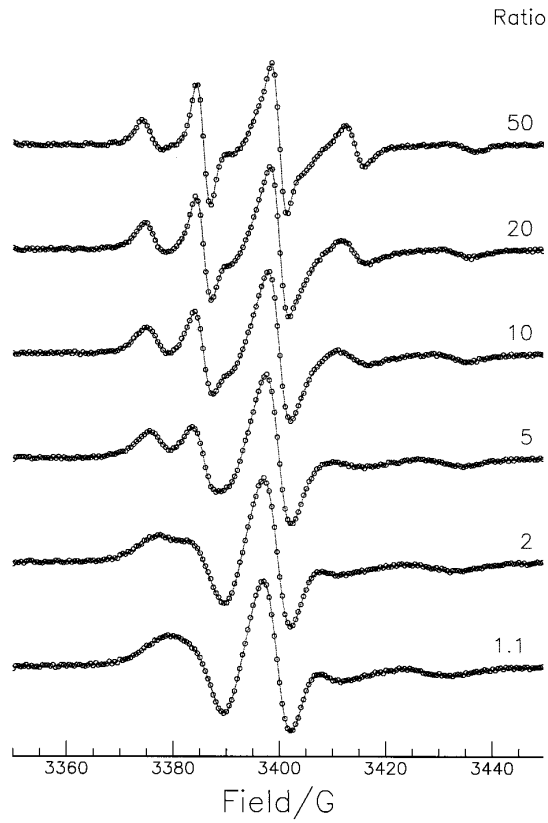

FIG. 7. Series of simulated two-component X-band CW EPR spectra calculated for two isotropically diffusing nitroxides with identical magnetic parameters (set 4, Table 2) but different rotational diffusion constants, $\bar{R}_1$ and $\bar{R}_2$. The geometric mean of $\bar{R}_1$ and $\bar{R}_2$ was kept fixed at $3 \times 10^7\ s^{-1}$ and the ratio $\bar{R}_1/\bar{R}_2$ was 50, 20, 10, 5, 2, and 1.1, from top to bottom. The relative population of the more mobile component in the simulated spectra was 30% in all of the spectra, and normally distributed random noise was added to produce a signal-to-noise ratio of 100. Solid lines show two-component nonlinear-least-squares fits to the spectra which could be obtained for all but the bottom spectrum. Parameters obtained from the fits are given in Table 7.

TABLE 7

Least-Squares Rotational Diffusion Rates and Spectral Scaling Factors for Nonlinear-Least-Squares Fits to Simulated Spectra with Two Components Having Different Rotational Rates in the Ratios Indicated

	Starting values		Least-squares values		Scale factors		Correlation (site 1 vs 2)	
$\bar{R}_2/\bar{R}_1$	rbar (1)	rbar (2)	rbar (1)	rbar (2)	a_1	a_2	rbar	a
50	6.628	8.327	6.683 ± .045	8.329 ± .038	0.706 ± .035	0.294 ± .016	0.073	−0.523
20	6.827	8.128	6.857 ± .040	8.137 ± .038	0.714 ± .032	0.286 ± .018	0.141	−0.636
10	6.997	7.977	6.993 ± .030	8.025 ± .038	0.731 ± .032	0.269 ± .019	0.351	−0.735
5	7.128	7.827	7.146 ± .034	7.850 ± .052	0.723 ± .055	0.277 ± .041	0.672	−0.892
2	7.327	7.628	7.371 ± .045	7.729 ± .190	0.859 ± .169	0.141 ± 1.60	0.913	−0.992
1.1	7.457	7.498		7.509 ± 3.126	0.000	1.000	0.999	—

Note. The spectra and accompanying least-squares fits are shown in Fig. 7. Reported uncertainties correspond to the 68% confidence level.

included in the error analysis as described under Methods, and their uncertainties as calculated from Eq. [21], together with the correlation coefficient between the two scaling factors, are also given in Table 7.

At the very highest ratios, the individual components are quite well resolved by the fitting procedure, both in terms of the accuracy of the relative populations and in the accuracy of the motional parameters derived for each component. Although starting parameter values of the calculation are recovered within the reported uncertainties even down to a ratio of 2, the uncertainties become quite large, thus limiting the utility of the information derived from the spectrum.

Again, we find that the correlation coefficients amongst the parameters offer a very sensitive measure of the resolving power of the spectrum. In this case, resolution is lost with respect to discriminating spectral components rather than the individual characteristics of a single component. As the ratio of the mobilities of the two species gets smaller, the spectra become progressively less distinguishable, and the correlation between the rotational rates, as well as that between the scaling factors (cf. Table 7), increases dramatically. At the lowest ratio of 1.1, the constituent spectra become virtually indistinguishable, and the procedure eliminates the unnecessary component by returning one of the scale factors as zero.

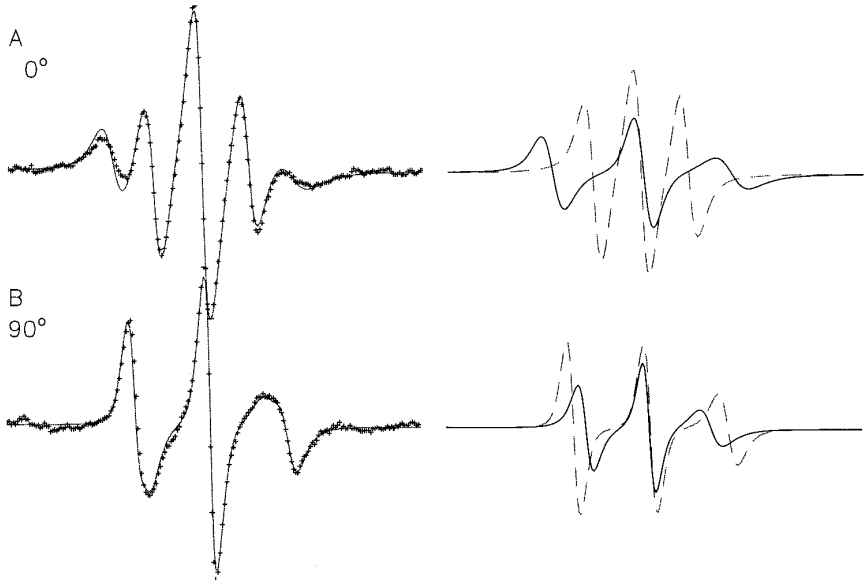

FIG. 8. Experimental X-band CW EPR spectra of the CSL spin probe diffusing in well-aligned multilayers containing a mixture of the lipid POPC, cholesterol, and dioctanoylglycerol at (A) 0° and (B) 90° director tilt. Solid lines show the least-squares global fit of two spectral components to the data. The component spectra calculated in the fit are shown on the right-hand side for both tilt angles; they consist of a well-aligned phase (dashed lines) and a disordered, MOMD-like phase (solid lines).

Figure 8 illustrates a recent practical application of multicomponent fitting in our laboratory, which has been utilized in conjunction with some of the other features of program NLSL described above, including global analysis and the fitting of MOMD spectra. Shown are experimental X-band CW EPR spectra of the CSL spin probe diffusing in well-aligned multilayers containing a mixture of the lipid POPC, with 10 mol% cholesterol, and 10 mol% dioctanoylglycerol (DOG) at 36°C, obtained at director tilt angles of 0° and 90°, together with the least-squares fits. Two components are well resolved from these spectra. They are plotted on the right-hand side for both tilt angles, utilizing the least-squares fits. Two components consist of a well-aligned phase (light lines) and a disordered, MOMD-like phase (darker lines), which is independent of orientation. Significantly, the appearance and amount of the disordered phase appears to vary almost reversibly with temperature and membrane composition. Using multicomponent analysis, it has been possible to characterize accurately the molecular dynamics and ordering in each of the two phases, as well as their relative amounts, over the entire temperature range for which a second phase is observed.

Statistical Analysis of Experimental Slow-Motion Spectra

When the approximations expressed in Eq. [19] are not valid, reliable statistical analysis requires a more robust estimate of the parameter confidence regions than the linear-limit expressions given in Eqs. [21]–[24]. The best method for this application is Monte Carlo estimation of the parameter probability distribution (*37*, *65*). The basic approach, once a minimum $\mathbf{x}^*$ has been located, is to generate a set of synthetic spectra (typically several hundred) starting from the least-squares spectrum $\phi(\mathbf{x}^*)$ and adding noise having the same probability distribution as the residuals $\mathbf{f}(\mathbf{x}^*)$. The noise may be derived either by generating a pseudorandom number with a known probability distribution function (*37*) or by randomly permuting the residuals $\mathbf{f}(\mathbf{x}^*)$ and adding them back to $\phi(\mathbf{x}^*)$ (*66*). For each spectrum synthesized in this manner, a minimization is performed, and the distribution of the resulting set of solutions $\{\mathbf{x}_j^*\}$ is then used to approximate the probability distributions of the fitting parameters.

In practice, experimental slow-motion EPR spectra almost never rigorously satisfy the assumptions that are required in order to interpret the linear-limit expressions given in Eqs.

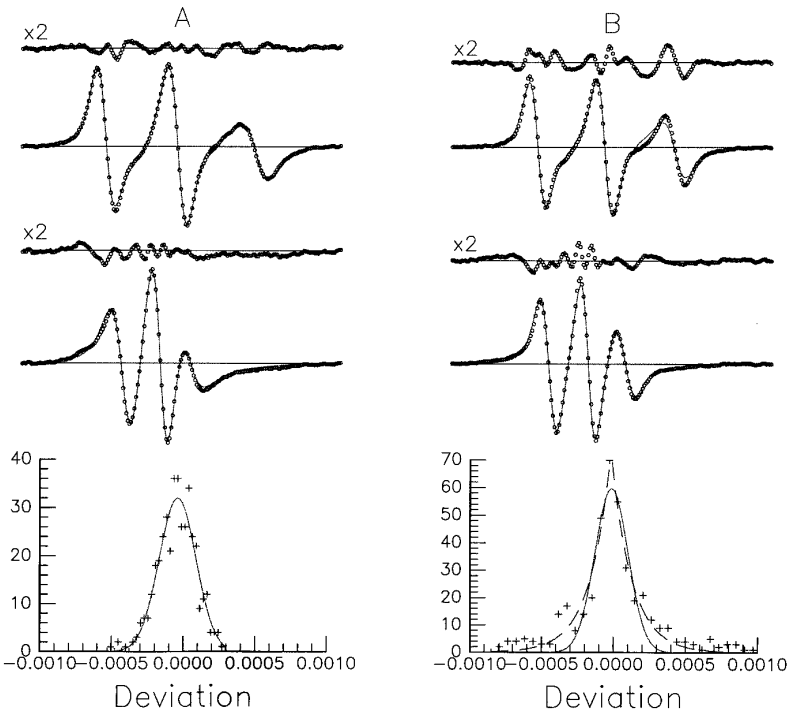

FIG. 9. Experimental X-band CW EPR spectra of the 5-PC spin probe diffusing in well-aligned multilayers of the lipid POPC at (A) 25°C and (B) 65°C for director tilt angles of 0° (upper spectra) and 90° (lower spectra). Solid lines show the global least-squares fits. Parameter uncertainties and correlations for the fits are given in Table 8. Above each spectrum are plotted the residuals for the fit, scaled by a factor of two. The bar graphs at the bottom show a binned distribution of the combined residuals for 0° and 90°, superimposed with a Gaussian curve fit by least squares to the distribution data at both temperatures. The dashed line on the plot at bottom left shows the best approximation of a double-exponential function to the residual distribution.

[21]–[24] as valid estimates of the parameter uncertainties. At reasonable signal-to-noise ratios, small nonrandom deviations between calculated and experimental spectra almost inevitably appear, even in the best fits. (At low signal-to-noise ratios, when the residuals are dominated by the spectral noise, the residuals do exhibit more nearly random behavior, which simplifies the error analysis.) Systematic deviations may occur because an incorrect model is being employed, or because the model is itself approximate at some level, and cannot reproduce the underlying spectrum exactly. They may also result from small but systematic experimental errors such as nonlinearities in the field sweep or signal amplification.

These considerations suggest that Monte Carlo estimation is unavoidable if one requires accurate error analysis of EPR slow-motion parameters, however impractical this may be for routine use, even on modern laboratory workstations or personal computers. However, it is possible to avoid such a time-consuming procedure in many cases of practical significance, as will be illustrated below.

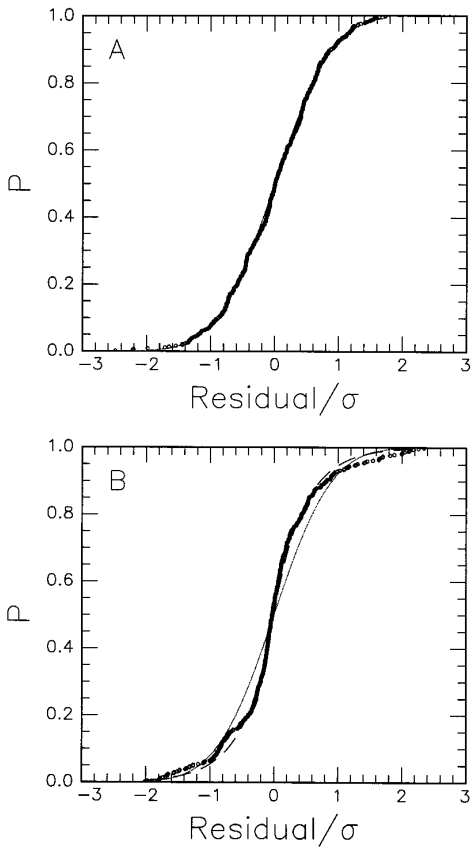

FIG. 10. Cumulative probability distribution plots for the residuals shown in Fig. 9, in comparison with a normal distribution function. (A) Residuals from 25°C data, exhibiting nearly normal distribution; (B) residuals from 65°C data, showing significant deviation from a normal distribution (solid line) but closer correspondence with a double-exponential distribution (dashed line).

Fouse and Bernhard (*67*) have demonstrated the use of the Monte Carlo method in EPR spectral analysis. They carried out a Gauss–Newton least-squares fit of the hyperfine splitting for an organic radical trapped in a single crystal as a function of crystal orientation. Using the procedure described above with normally distributed errors, they obtained the uncertainties in the hyperfine tensor principal values and principal axis direction cosines and showed that these were consistent with the uncertainties predicted from the curvature matrix according to Eqs. [21] and [22].

The results of Fouse and Bernhard should be generally applicable in EPR spectral fitting. For most EPR spectra, including rigid-limit, slow-motion, and multidimensional spectra, the dependence of the residuals on the fitting parameters will be sufficiently well behaved that the elliptical uncertainty region given by Eq. [22] is a good approximation to the local topology of the function. Consequently, when normally distributed noise is added to the calculated function $\mathbf{g}(\mathbf{x}^*)$ in the Monte Carlo procedure, all of the criteria that make the linear-limit confidence region a reliable estimate of the parameter probability distribution are satisfied. Thus, it should generally be possible to utilize the curvature matrix instead of Monte Carlo estimation in the case of normally distributed residuals, even though the residuals may contain systematic deviations exceeding the ambient noise level.

In the presence of nonrandom deviations, the error estimates based on unweighted residuals will be significantly larger than those from weighted residuals. This is because correlated deviations above the noise level contribute to the unweighted residual norm s^2 (cf. Eqs. [23] and [24]) but by definition do not contribute to the average spectral noise $\langle\sigma\rangle$. Thus, when $s^2 \gg \langle\sigma\rangle^2$, the uncertainty estimates from the *unweighted* residuals using t and F statistics according to Eqs. [23] and [24] should correspond to the Monte Carlo estimates for normally distributed residuals.

It might therefore appear that one should always minimize with respect to unweighted residuals, especially since χ^2 minimization suffers the additional drawback that some knowledge of the spectral noise is required, which may admit errors due to incorrect noise estimation. However, the two types of minimization can sometimes lead to appreciable discrepancies in the ''best-fit'' parameters when the noise levels differ significantly amongst the experimental spectra in a global fit. It is generally desirable to retain the weighting of a χ^2 minimization, so that the features of noisy spectra are less influential than those of ''clean'' spectra. In such cases, the confidence regions obtained from the unweighted residuals may be estimated by scaling those obtained from the weighted residuals as follows. The elements of the c and C matrices in Eqs. [21]–[24] are approximately related by

$C_{ij} = C'_{ij}/\langle\sigma\rangle^2$ and $c_{ij} = \langle\sigma\rangle^2 c'_{ij}$ where $\langle\sigma\rangle$ is estimated for the n spectra using

$$\langle\sigma\rangle = \left[\sum_{i=1}^{n} \frac{\sigma_i^2 m_i}{m_{\text{tot}}}\right]^{1/2}, \qquad [35]$$

in which σ_i and m_i are the standard deviation of the noise and the number of points in the ith spectrum and $m_{\text{tot}} = \Sigma_i m_i$. In the limit of large m_{tot}, the χ^2, t, and F distributions become equivalent; that is, $\Delta\chi_1^2(\alpha) \approx (t^{\alpha/2}_{m_{\text{tot}}-1})^2$ and $\Delta\chi_p^2(\alpha) \approx pF^{\alpha}_{m_{\text{tot}}-p,p}$. Thus, the value of δx_i given by Eq. [21] may simply be multiplied by $s/\langle\sigma\rangle$ to estimate that given by Eq. [23]; similarly, the boundary for the joint confidence region ellipsoid on the right-hand side of Eq. [22] may be scaled by $s^2/\langle\sigma\rangle^2$, to obtain the uncertainty region predicted by Eq. [24].

To illustrate this scaling procedure and when it may be applied, we show in Fig. 9 two typical examples of fits to experimental slow-motion spectra from the 5-PC spin probe diffusing in an aligned POPC membrane at director tilt angles of 0° and 90°, obtained at temperatures of 25 and 65°C, respectively. Least-squares minimizations were carried out with respect to the weighted residuals using fixed values of the diffusion tilt angle β_d, rotational anisotropy N, and inhomogeneous broadening parameters. Error analysis was restricted to the variable parameters $\bar{R}$, c_0^2, and c_2^2.

The calculated spectra shown by the solid lines in Fig. 9 seem qualitatively to satisfy the prevailing standards for acceptable fits to experimental slow-motion EPR data, especially considering that they are based on global minimizations for two director tilt angles. The agreement is quite good to the eye for both $\psi = 0°$ and $\psi = 90°$, although there remain some perceptible deviations, especially around the high-field peak of the $\psi = 0°$ spectrum at the higher temperature (right-hand side).

By any quantitative criterion, however, the fits shown in Fig. 9 are unacceptable for the purposes of error analysis. The reduced χ^2 values are respectively 115 and 270 for the fits shown on the left- and right-hand sides. More importantly, the residual functions plotted above each spectrum in an expanded scale clearly reveal the presence of systematic errors above the level of the ambient noise. This qualitative observation is supported by more quantitative tests for systematic behavior in residuals, such as the runs test (*68*) and the Durbin-Watson test based on serial lag plots of the residuals (*69*). These tests indicate that the residuals at both temperatures exhibit very high correlation, which is typical for the vast majority of fits to experimental slow-motion spectra.

However, the residuals at the two temperatures do have significantly different distributions as can be seen from the binned distribution plots at the bottom of Fig. 9 and also the cumulative probability distributions shown in Fig. 10. The plots in Fig. 10 were obtained by sorting the m_{tot} residuals r_i in increasing order, and then calculating the cumulative probability associated with each residual, P_i, as the fraction of data points having a smaller residual, $P_j = (i - \frac{1}{2})/m_{\text{tot}}$. The solid line in each plot represents a normal distribution having the same mean $\bar{r}$ and standard deviation σ_r as $\{r_i\}$, given by

TABLE 8
Comparison of Parameter Uncertainties and Correlation Coefficients Estimated by Different Methods for Two Fits to Experimental Spectra from the 5PC Spin Probe in Well-Aligned POPC Multilayers

		Parameter uncertainty			Correlation coefficients: From covariance matrix		Correlation coefficients: From multiple minimizations	
	Parameter	Eq. [22]	Scaled estimate	Multiple minimization	c20	c22	c20	c22
Normally	rbar	0.0011	0.0118	0.0120	0.329	−0.248	0.413	−0.306
Distributed	c20	0.0085	0.0904	0.0596		−0.898		−0.995
Residuals	c22	0.0140	0.149	0.110				
Nonnormally	rbar	0.0008	0.0136	0.0393	0.568	−0.505	0.063	−0.097
Distributed	c20	0.0076	0.125	0.397		−0.952		−0.994
Residuals	c22	0.0103	0.1711	0.570				

Note. Parameter uncertainties were obtained using Eq. [22], the scaling procedure described in the text, and estimation based on 100 minimizations started from randomly chosen initial points. Correlation coefficients were calculated directly from the covariance matrix or estimated as the linear correlation coefficients for the distribution of the 100 parameter sets.

$$P_{\text{normal}}(x) = \frac{1}{2}\left(\text{erf}\left[\frac{x - \bar{r}}{\sigma_r}\right] + 1\right). \qquad [36]$$

From Fig. 10, it can be seen that the residuals at the lower temperature (Fig. 10A) conform well to a normal distribution, whereas those from the higher temperature (Fig. 10B) deviate substantially, particularly in the ''wings'' of the distribution. A more quantitative test for determining whether a set of residuals obeys a given probability distribution is the Kolmogorov-Smirnov (K-S) test (*37, 70*). For comparison with a normal distribution, the K-S statistic D is defined as

$$D_{\text{normal}} = \max_{1 \leq i \leq m} |P_i - P_{\text{normal}}(r_i)|. \qquad [37]$$

The D statistic itself has a cumulative distribution function that gives the significance level at which the hypothesis that two distributions are the same may be rejected. The significance level may be obtained by comparing a given value of D to tabulations of the K-S cumulative distribution function (*70*) or by direct numerical calculation (*37*). For the 400 residuals plotted in Fig. 10A, the value of D_{normal} is 0.033, corresponding to a significance of 0.78, whereas those in Fig. 10B give $D_{\text{normal}} = 0.141$ and a significance level of 0.0031, clearly indicating a strong departure from a normal distribution. Thus, according to the arguments given above, one may expect the scaling procedure to predict the results of Monte Carlo estimation reasonably well for the first spectrum, but not for the second.

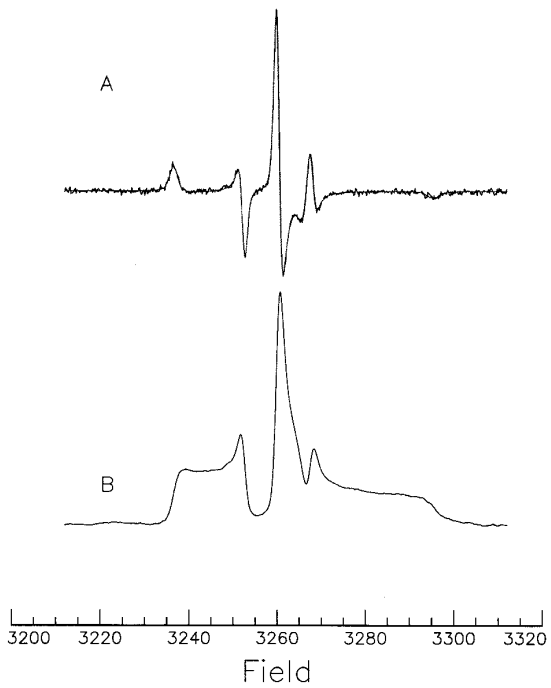

FIG. 11. Simulated CW EPR MOMD spectrum for a CSL probe exhibiting fast motion and high ordering (set 5 in Table 2). (A) First-derivative CW lineshape with added noise ($S/N = 100$), and least-squares fitted spectrum (Table 9). (B) Integral of simulated data.

This prediction is borne out by a comparison of the parameter uncertainties that were calculated for the fits shown in Fig. 9 in three different ways: (a) using the estimates from the weighted residuals given by Eq. [21]; (b) scaling the estimate in (a) by $s/\langle\sigma\rangle \cong \sqrt{\chi^2_{\text{red}}}$; and (c) estimation based on 100 minimizations with randomly selected initial parameter values. The results are compared in Table 8. As can be seen from this tabulation, the uncertainty bounds calculated from the weighted residuals are consistently much smaller than the estimates by multiple minimization for all parameters, regardless of the distribution of the residuals. The scaled estimates are quite comparable in the case of normally distributed residuals; however, they still severely underestimate the uncertainty for the nonnormally distributed residuals.

In cases of extreme departure from a normal residual distribution, several alternative approaches should be considered. The observation of such distributions call into question both the appropriateness of the fitting model and the use of mean-squared deviation as a best-fit criterion. If the least-squares fit cannot be substantially improved by utilizing alternative models, it may be necessary to refine the ''optimal'' parameter values by minimizing with respect to a maximum-likelihood estimator for a nonnormal residual distribution (*37*). Nonnormal residual distributions from slow-motion EPR spectral fits may be well approximated instead by the double-exponential distribution. Figure 9 shows one such case: the nonnormal residuals in Fig. 9B also show a double-exponential distribution (dotted line) which conforms quite well to the observed distribution and gives a significance factor 0.28 when compared to the experimental distribution by the K-S test. For this type of residual distribution, the maximum-likelihood estimator ξ is the mean *absolute* deviation of the fitting function (*37*).

$$\xi = \frac{1}{M}\sum_{i=1}^{M} |h_i(\mathbf{x}) - \phi_i(\mathbf{x})|, \qquad [38]$$

which is relatively easy to program, and should be considered as an alternative minimization method in the case of nonnormal residuals.

Multidimensional Spectroscopy

Nonlinear-least-squares analysis is essential for the quantitative interpretation of 2D FT spectra because of the complicated lineshapes that arise in the slow-motion region, and the subtle interplay of inhomogeneous linewidths and spin relaxation processes in determining the shapes and rela-

tive intensities of auto and cross peaks in 2D spectra (*21, 22, 28*). The analysis is aided by the presence of the second frequency in the spectrum, which offers many of the same advantages as the global fitting of multiple CW spectra discussed above. In addition, the 2D spectrum usually permits the direct determination of quantities that are not accessible by CW EPR on a single sample. These include homogeneous linewidths as well as nuclear relaxation rates and Heisenberg spin exchange. It also allows one to study motions that are too slow to be discerned from the CW EPR spectrum, and it enables one to distinguish important but subtle details of the motional models (*28*). These advantages are enhanced by simultaneously analyzing different linear combinations that provide complementary information from a given measurement, such as the S_{c+} and S_{c-} signals (*22, 22, 28*). The ultimate extension of this strategy is the 2D ELDOR experiment, for which the 2D S_{c+} and S_{c-} signals may be obtained for a series of mixing times (including the

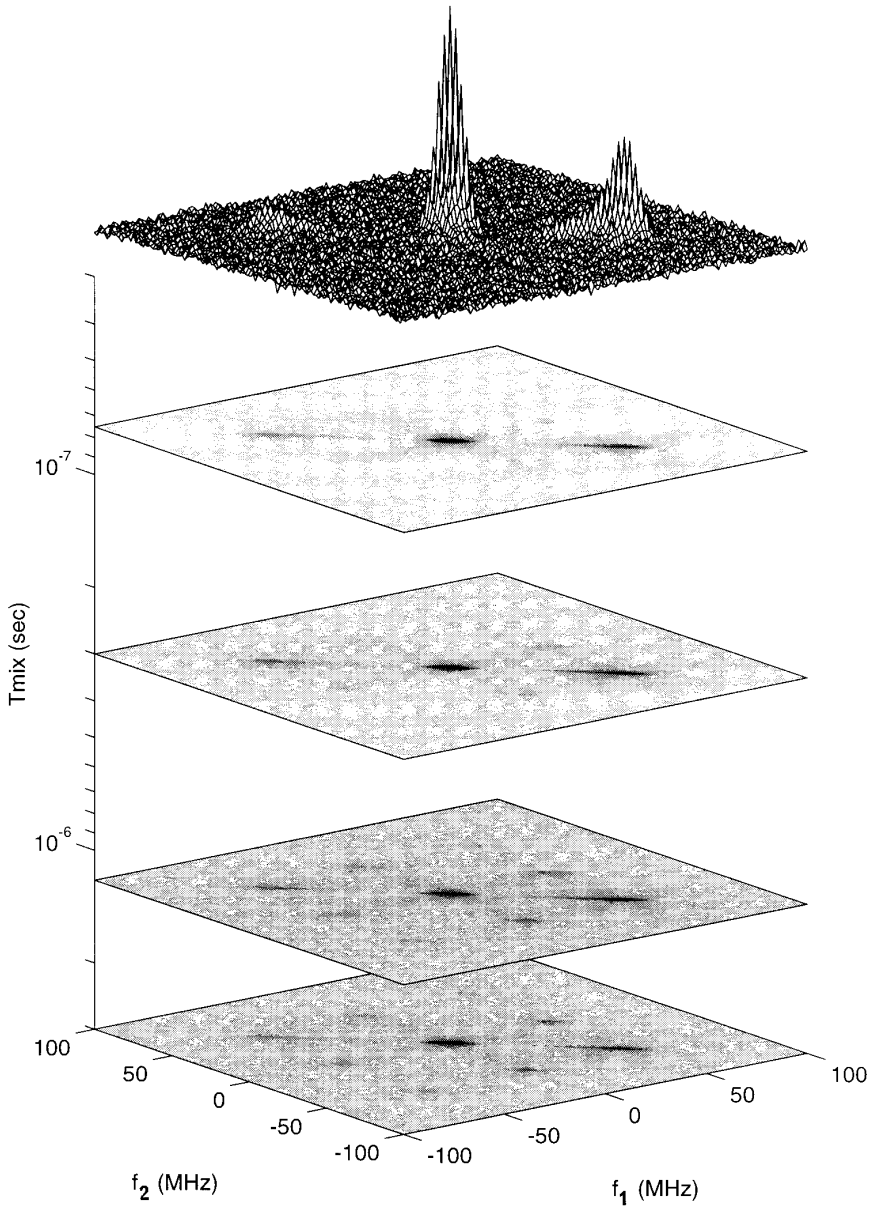

FIG. 12. Simulated 2D ELDOR MOMD spectrum (S_c signal) for a CSL probe exhibiting fast motion and high ordering (set 5 in Table 2). Shown is the COSY spectrum (T = 0 ns, raised surface at top) and 2D ELDOR spectra for mixing times 75, 300, 1200, and 3000 ns (gray-scale images projected onto planes at different levels on the z axis). The intensity scale is separately normalized for each 2D slice. Note the development of slightly elongated cross peaks with increasing T.

COSY spectrum at zero mixing time), effectively yielding a 3D data set which can be analyzed globally by least squares (*21, 22, 28*). Such a series of 2D ELDOR spectra are fitted by the same dynamic and ordering parameters and are characterized by the growth of the cross peaks relative to the auto peaks as a function of mixing time. It affords a large number of experimental constraints that can be used to improve the accuracy of the determination and resolve or reduce ambiguities amongst the fitting parameters.

We now compare fits to simulated data sets representing the CSL in randomly oriented lipid vesicles for CW EPR, 2D ELDOR at a single mixing time, and 2D ELDOR at a series of eight mixing times including the COSY spectrum at zero mixing time. Figure 11 shows two representations of the simulated CW EPR MOMD spectrum for a CSL probe exhibiting relatively fast motion and high ordering (the corresponding parameters are given as set 5 in Table 2). At the top of the figure is the standard first-derivative CW line-

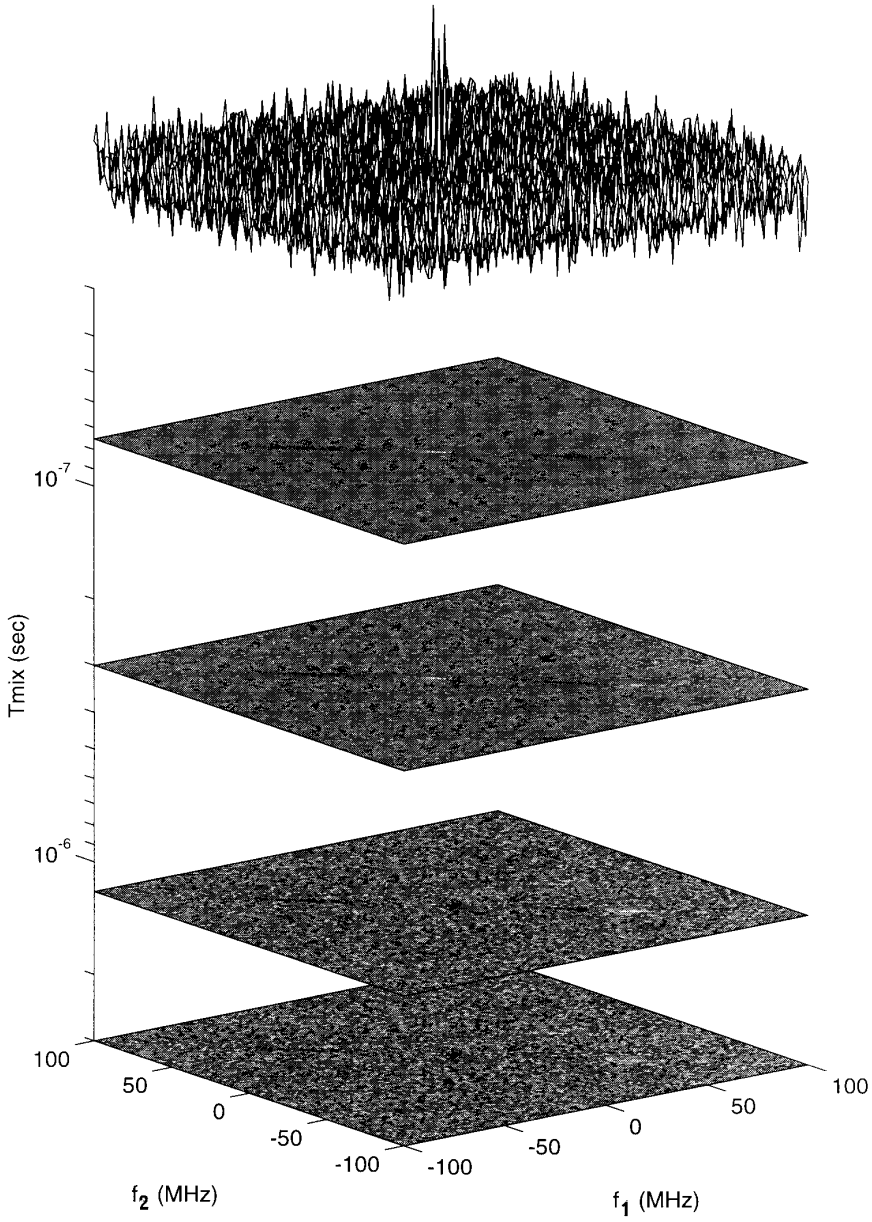

FIG. 13. Residuals from global nonlinear-least-squares fit to the simulated 2D ELDOR MOMD spectrum shown in Fig. 12 (spectra for $T = 150$ and 300 ns included in the fit are not shown). Raised surface at top is the residual plot for the COSY spectrum at $T = 0$. Deviations between data and fit appear as light and dark lines along the negative diagonal of the gray-scale images, respectively indicating under- and overestimation of the spectral intensity. The intensity scale is separately normalized for each 2D slice.

TABLE 9

Least-Squares Rotational Diffusion Rates and Ordering Potential Coefficients for Nonlinear Least-Squares Fits to Simulated CW EPR and 2D ELDOR Spectra of a MOMD Sample Using Parameter Set 5 from Table 2

Type of fit	Parameter	Least-squares value	Correlation matrix			
			gib	rbar	c20	c22
CW EPR	gib	1.108 ± 0.128	1	0.731	0.458	0.101
$\chi^2_{red} = 3.124$	rbar	8.354 ± 0.024		1	0.320	0.045
	c20	4.550 ± 0.034			1	−0.529
	c22	0.1628 ± 0.040				1
2D ELDOR	gib	1.208 ± 0.084	1	0.398	−0.367	0.111
(single T_m)	rbar	8.262 ± 0.014		1	−0.700	0.759
$\chi^2_{red} = 1.085$	c20	4.199 ± 0.019			1	0.110
	c22	0.404 ± 0.005				1
2D ELDOR[a]	gib	1.192 ± 0.058	1	0.427	−0.130	−0.380
(global)	rbar	8.310 ± 0.006		1	−0.515	−0.417
$\chi^2_{red} = 1.050$	c20	4.427 ± 0.006			1	0.596
	c22	0.410 ± 0.001				1

Note. The spectra and accompanying least-squares fits are shown in Fig. 11 for the CW EPR fit and Figs. 12 and 13 for the 2D ELDOR fit. Reported uncertainties correspond to the 95% confidence level.

[a] Errors and correlations were estimated directly from the Jacobian matrix calculated at the starting parameter set.

shape with added noise ($S/N = 100$), plotted together with the least-squares fitted spectrum. At the bottom is the integrated signal, i.e., the CW absorption spectrum, which is included for comparison with the 2D spectrum. The isotropic distribution of microscopic membrane directors produces a sharp central peak and two peaks exhibiting ''powder-like'' patterns on either side.

A 2D ELDOR spectrum (S_{c-} signal) was simulated using the same parameter set for mixing times of 0, 75, 150, 300, 600, 1200, 2000, and 3000 ns. Figure 12 shows the calculated spectrum for selected mixing times (spectra for $T =$ 150, 600, and 2000 ns have been omitted for compactness). The raised surface plotted at the top of the figure shows the appearance of the COSY ($T = 0$) spectrum. A somewhat different representation is used for the nonzero mixing times so that the spectral intensity may be represented as a function of the two spectral frequencies f_1 and f_2 as well as T. The intensity of each 2D spectrum is represented as a gray-scale (or color) intensity map on a flat plane that is projected into a perspective view, and located along the z axis at the corresponding T value. The intensity scale is separately normalized for each 2D slice. From this type of plot, the development of slightly elongated cross peaks can clearly be observed with increasing T.

The data shown in Fig. 12 were fitted first by minimizing with respect to the single 2D spectrum calculated at $T =$ 300 ns and, second, by globally fitting the entire dataset. The χ^2_{red} of both 2D fits were close to unity, and their residuals were very close to normally distributed, with KS significance factors higher than 0.9 in both cases. Some small deviations between the simulated data and the fit are still observable in the plot of residuals in Fig. 13, which is formatted in a similar way to Fig. 12 for comparison. The deviations appear as light and dark lines along the negative diagonal of the 2D slices, respectively indicating under- and overestimation of the spectral intensity by the calculated fit. However, the mesh plot at the top of Fig. 13 shows that the deviations were very close to the ambient noise level.

The results from the CW and 2D fits are given in Table 9. As one might expect for simulated data sets, all three minimizations recovered the starting lineshape and parameters reasonably well. However, the CW EPR results exhibit significantly higher parameter uncertainties, and somewhat higher correlations amongst the fitting parameters. The fit

TABLE 10

Least-Squares Rotational Diffusion Rates and Ordering Potential Coefficients for Nonlinear Least-Squares Fits to Experimental 2D ELDOR Spectra of CSL in a Vesicle Dispersion at 70°C.

		gib	rbar	c20	c22
gib	1.243 ± 0.080	1	0.244	−0.075	−0.235
rbar	8.373 ± 0.022		1	−0.801	−0.226
c20	2.848 ± 0.020			1	−0.402
c22	0.484 ± 0.012				1

Note. The experimental spectrum and residuals are shown in Figs. 13 and 14. Reported uncertainties correspond to the 95% confidence level.

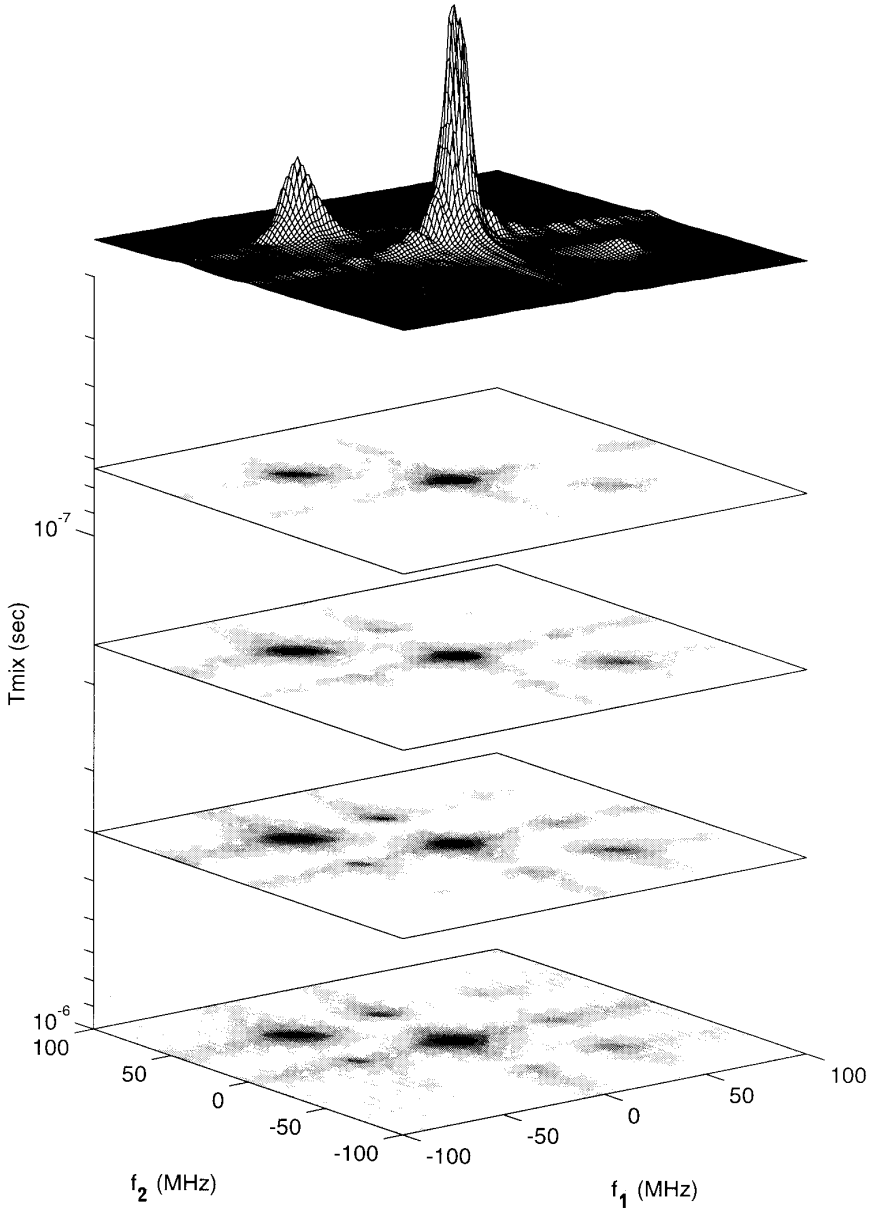

FIG. 14. Experimental 2D ELDOR spectrum (S_c-signal) for the CSL spin probe diffusing in randomly oriented POPC vesicles with approximately 20 wt% water at 70°C. Shown is the COSY spectrum ($T = 0$ ns, raised surface at top) and 2D ELDOR spectra for mixing times of 73, 106, 166, 400, and 1000 ns (gray-scale images projected onto planes at different levels on the z axis). The intensity scale is separately normalized for each 2D slice.

to a single mixing time appreciably reduces the parameter uncertainties, although the correlations remain comparable to those observed in the CW fit. For the global fit, however, both the uncertainties and the correlations are significantly reduced relative to the CW case.

Comparable reductions in parameter uncertainties and correlations are achieved in the global analysis of actual experimental data. Figure 14 shows a 2D ELDOR spectrum (S_c–signal) of the CSL spin probe diffusing in randomly oriented membranes of the lipid POPC with approximately 20 wt% water at 70°C. Parameters obtained by a global fit to spectra obtained at $T = 0$, 73, 106, 166, 400, and 1000 ns are given in Table 10, and the residuals from the fit are shown in Fig. 15.

Program Distribution

The source code for the programs described in this paper may be obtained via anonymous ftp at the Cornell Materials Science Center. The routines are mainly coded in standard

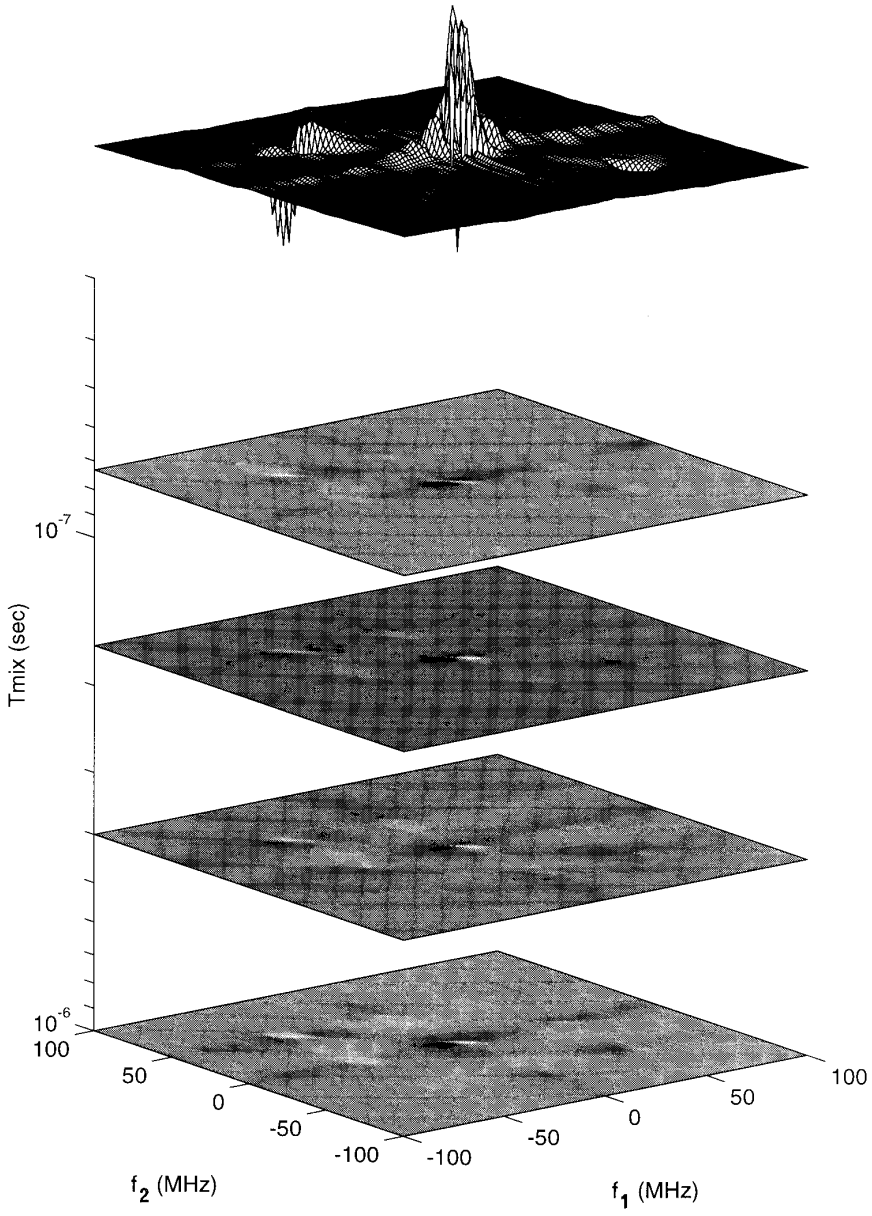

FIG. 15. Residuals from global nonlinear-least-squares fit to the experimental 2D ELDOR spectrum shown in Fig. 14. Raised surface at top is the residual plot for the COSY spectrum at $T = 0$. Deviations between data and fit in the 2D ELDOR spectra appear as light and dark areas in the gray-scale images, respectively indicating under- and overestimation of the spectral intensity. The intensity scale is separately normalized for each 2D slice.

Fortran 77 and compiled using the IBM xlf compiler (comparable to f77 on most Unix systems). There are a few routines written in C that perform low-level functions such as intercepting the Unix ''kill'' signal so that the user may halt the fitting procedure without exiting the program. In program NLSL, additional C subroutines are used to interface with low-level X-windows library functions for graphical monitoring of the fit. Some modification of the C subroutine names may be required in order to run on systems other than the IBM RS6000.

The following directories are available:

EPRLL, version 1.6 (family of single-spectrum calculation programs, including the program (EPRBL) used to determine truncated basis sets).

EPRLF (family of single-spectrum calculation programs similar to EPRLL but including nonsecular terms).

NLSL (CW version of least squares)

NL2DC (2D FT version of least squares using the conjugate gradients matrix solution)

NL2DR (2D FT version using the Rutishauser matrix solution)

To obtain the software, use the following steps:

(1) **ftp ftp.msc.cornell.edu**
(2) At the login prompt, type **anonymous**
(3) Enter your e-mail address as the password.
(4) Type **cd pub/freed** to access the home directory
(5) There an updated list of the available programs and the directories in which they are located may be found in the file ''README.1st.'' Type **get README.1st** to consult this list and determine the directory for the desired program.
(6) Change to the desired directory with the command **cd directory_name**.
(7) Type **get *filename*** to obtain the desired file. Although the large number of subroutine files that compose each program are available individually, several standard formats are provided that allow transfer of all files at once, depending on how the routines are to be unpacked after downloading. Details regarding the available formats may be found in file README.1st.

ACKNOWLEDGMENTS

We gratefully acknowledge extensive contributions of Drs. Richard H. Crepeau and Mingtao Ge and David J. Schneider to the development of earlier versions of the CW EPR and nonlinear-least-squares programs described here and their many useful discussions during the course of this work. Computations and program development were performed at the Cornell Theory Center and the Cornell Materials Science Center. This work was supported by NIH Grants RR07126 and GM25862 and by NSF Grants CHE9313167 and DMR9210638.

REFERENCES

1. J. Gorcester and J. H. Freed, *J. Chem. Phys.* 85, 5375 (1986).
2. J. Gorcester and J. H. Freed, *J. Chem. Phys.* 88, 4678 (1988).
3. J. Gorcester, S. B. Rananavare, and J. H. Freed, *J. Chem. Phys.* 90, 5764 (1989).
4. B. R. Patyal, R. H. Crepeau, D. Gamliel, and J. H. Freed, *Chem. Phys. Lett.* 175, 445 (1990); 175, 453 (1990).
5. W. B. Lynch, K. A. Earle, and J. H. Freed, *Rev. Sci. Instrum.* 59, 1345 (1988).
6. D. E. Budil, K. A. Earle, W. B. Lynch, and J. H. Freed, *in* ''Advanced EPR: Applications in Biochemistry and Biology,'' (A. J. Hoff, Ed.), Elsevier, Amsterdam, 1989.
7. O. Ya. Grinberg, A. A. Dubinskii, and Ya. S. Lebedev, *Russ. Chem. Rev.* 52, 850 (1983).
8. R. T. Weber, J. A. J. M. Disselhorst, L. J. Prevo, J. Schmidt, and W. Th. Wenckebach, *J. Magn. Reson.* 80, 383 (1988).
9. T. F. Prisner, S. Un, and R. G. Griffin, *Isr. J. Chem.* 32, 357 (1992).
10. J. H. Freed, *in* ''Spin Labeling: Theory and Application'' (L. J. Berliner, Ed.), Academic Press, New York, 1976.
11. E. Meirovitch, E. Igner, D. Igner, G. Moro, and J. H. Freed, *J. Chem. Phys.* 77, 3915 (1982).
12. D. J. Schneider and J. H. Freed, *in* ''Spin Labeling: Theory and Application, Biological Magnetic Resonance'' (L. J. Berliner and J. Reuben, Eds.), Vol. 8, Plenum, New York, 1989.
13. G. Moro and J. H. Freed, *J. Chem. Phys.* 74, 3757 (1981).
14. B. Kirste, *in* ''Handbook of EPR Spectroscopy'' (C. P. Poole, Jr., and H. A. Farach, Eds.), AIP Press, New York, 1994.
15. R. H. Crepeau, S. B. Rananavare, and J. H. Freed, *in* ''10th International EPR Symposium, Rocky Mountain Conferency on Analytical Chemistry, Denver, 1987.''
16. Y. K. Shin and J. H. Freed, *Biophys. J.* 55, 537 (1989); 56, 1093 (1989).
17. M. Ge, S. B. Rananavare, and J. H. Freed, *Biochim. Biophys. Acta* 1036, 228 (1990).
18. J. H. Freed, *J. Chem. Soc. Faraday Trans.* 86, 3173 (1990).
19. E. J. Hustedt, C. E. Cobb, A. H. Beth, and J. M. Beechem, *Biophys. J.* 64, 614 (1993).
20. E. J. Hustedt, A. Spaltenstein, J. J. Kirchner, P. B. Hopkins, and B. H. Robinson, *Biochemistry* 32, 1774 (1993).
21. R. H. Crepeau, S. Saxena, S. Lee, B. R. Patyal, and J. H. Freed, *Biophys. J.* 66, 1489 (1994).
22. S. Lee, B. R. Patyal, S. Saxena, R. H. Crepeau, and J. H. Freed, *Chem. Phys. Lett.* 211, 397 (1994).
23. Program POWFIT for simulating EPR powder spectra. Author: David R. Duling, National Institute of Environmental Health Science, 1994.
24. P. G. Fajer, R. L. H. Bennett, C. F. Polnaszek, E. A. Fajer, and D. D. Thomas, *J. Magn. Reson.* 88, 111 (1990).
25. R. F. Campbell, E. Meirovitch, and J. H. Freed, *J. Phys. Chem.* 83, 525 (1979).
26. E. Meirovitch, A. Nayeem, and J. H. Freed, *J. Phys. Chem.* 88, 3454 (1984).
27. K. V. Vasavada, D. J. Schneider, and J. H. Freed, *J. Chem. Phys.* 86, 647 (1987).
28. S. Lee, D. E. Budil, and J. H. Freed, *J. Chem. Phys.* 101, 5529 (1994).
29. L. J. Schwartz, E. Meirovitch, J. A. Ripmeester, and J. H. Freed, *J. Phys. Chem.* 87, 4453 (1983).
30. R. R. Ernst, G. Bodenhausen, and A. Wokaun, ''Principles of Nuclear Magnetic Resonance in One and Two Dimensions,'' Oxford, New York, 1987.
31. D. Gamliel and J. H. Freed, *J. Magn. Reson.* 89, 60 (1990).
32. J. J. Moré, B. S. Garbow, K. E. Hillstrom, ''Users' Guide for MINPACK-1,'' Publication ANL 80-74, National Technical Information Service, Springfield, Virginia 1980.
33. J. E. Dennis, Jr., and R. B. Schnabel, ''Numerical Methods for Unconstrained Optimization and Nonlinear Equations,'' Prentice-Hall, Englewood Cliffs, New Jersey, 1983.
34. G. A. F. Seber and C. J. Wild, ''Nonlinear Regression,'' Chap. 15, Wiley, New York, 1989.
35. IMSL Inc., 2500 City West Blvd., Houston, Texas 75502.
36. Z. L. Mádi and R. R. Ernst, *J. Magn. Reson.* 79, 513 (1988).
37. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, ''Numerical Recipes,'' Cambridge Univ. Press, Cambridge, UK, 1986.
38. P. E. Gill, W. Murray, and M. H. Wright, ''Practical Optimization,'' Academic Press, New York, 1981.
39. G. H. Golub and V. Pereyra, *SIAM J. Numer. Anal.* 10, 413 (1973).
40. E. R. Malinkowski and D. G. Hovely, ''Factor Analysis in Chemistry,'' Wiley, New York, 1991.

41. J. C. Evans and P. H. Morgan, *Anal. Chim. Acta* 133, 329 (1981).

42. P. Moens, P. DeVolder, R. Hoogewijs, F. Callens, and R. Verbeeck, *J. Magn. Reson. A* 101, 1 (1993).

43. J. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart, ''LINPack Users' Guide,'' Society for Industrial and Applied Mathematics, 1979.

44. Abramowitz and I. Stegun, ''Handbook of Mathematical Functions,'' Chap. 25, Dover, New York, 1971.

45. P. R. Bevington, ''Data Reduction and Error Analysis for the Physical Sciences,'' McGraw–Hill, New York, 1969.

46. J. W. Chen, F. P. Auteri, D. E. Budil, R. L. Belford, and R. B. Clarkson, *J. Phys. Chem.* 98, 13452 (1994).

47. R. A. Jackson, *J. Chem. Soc. Perkins Trans.* 2, 523 (1983).

48. W. Bernhard and G. W. Fouse, *J. Magn. Reson.* 82, 156 (1989).

49. M. Barzaghi and M. Simonetta, *J. Magn. Reson.* 51, 175 (1983).

50. C. Chachaty, *J. Chim. Phys.* 82, 629 (1985).

51. D. R. Duling, A. G. Motten, and R. P. Mason, *J. Magn. Reson.* 77, 504 (1988).

52. B. Kirste, *Anal. Chim. Acta* 265, 191 (1992).

53. J. S. Hwang, R. P. Mason, L. P. Hwang, and J. H. Freed, *J. Phys. Chem.* 79, 489 (1975).

54. B. A. Kowert, *J. Phys. Chem.* 85, 229 (1981).

55. D. E. Budil, K. A. Earle, and J. H. Freed, *J. Phys. Chem.* 97, 1294 (1993); K. A. Earle, D. E. Budil, and J. H. Freed, *J. Phys. Chem.* 97, 13,289 (1993).

56. J. M. Beecham, E. Gratton, M. Ameloot, J. R. Knotson, and L. Brand, *in* ''Fluorescence Spectroscopy, Vol. 2: Principles'' (J. R. Lakowicz, Ed.), Plenum, New York, 1991.

57. J. M. Beecham, *in* ''Methods in Enzymology'' (L. Brand and M. L. Johnson, Eds.), Vol. 210, p. 37, Academic Press, San Diego, 1992.

58. M. Ge, D. E. Budil, and J. H. Freed, *Biophys. J.* 67, 2326 (1994).

59. H. Tanaka and J. H. Freed, *J. Phys. Chem.* 88, 6633 (1984); *J. Phys. Chem.* 89, 350 (1985).

60. N. A. Clark and K. J. Rothschild, *in* ''Methods in Enzymology'' (L. Packer, Ed.), Vol. 88, p. 326, Academic Press, New York, 1982.

61. L. Kar, E. Ney-Igner, and J. H. Freed, *Biophys. J.* 48, 569 (1985).

62. D. Marsh, *in* ''Biological Magnetic Resonance'' (L. J. Berliner and J. Reuben, Eds.), Vol. 8, Chap. 5, Plenum Press, New York, 1989.

63. L. Schwartz, G. L. Millhauser, and J. H. Freed, *Chem. Phys. Lett.* 127, 60 (1986).

64. G. L. Millhauser, J. Gorcester, and J. H. Freed, *in* ''Electron Magnetic Resonance in the Solid State'' (J. A. Weil, Ed.), Can. Chem. Soc., 1987, p. 571,

65. M. Straume and M. L. Johnson, *in* ''Methods in Enzymology'' (L. Brand and M. L. Johnson, Eds.), Vol. 210, p. 117, Academic Press, San Diego, 1992.

66. M. Straume and M. L. Johnson, *in* ''Methods in Enzymology'' (L. Brand and M. L. Johnson, Eds.), Vol. 210, p. 87, Academic Press, San Diego, 1992.

67. G. W. Fouse and W. A. Bernhard, *J. Magn. Reson.* 32, 191 (1978).

68. Y. Bard, ''Nonlinear Parameter Estimation,'' p. 201. Academic Press, New York, 1974.

69. N. R. Draper and R. Smith, ''Applied Regression Analysis,'' 2nd ed., p. 153. Wiley–Interscience, New York, 1981.

70. W. W. Daniel, ''Biostatistics: A Foundation for Analysis in the Health Sciences,'' 2nd ed., Wiley–Interscience, New York, 1978.