

Analysis of multicomponent exponential magnetic resonance relaxation data: automatable parameterization and interpretable display protocols

Charles D. Eads

Corporate R&D, The Procter & Gamble Company

8700 S. Mason-Montgomery Road, Mason, OH 45040, USA

ABSTRACT

Multicomponent exponential decay data arise frequently in magnetic resonance measurements and in numerous other circumstances. For any such data set, there are many different distributions of relaxation rates that can fit the data to high precision, so choosing or designing an optimal analysis can be challenging and ambiguous. Nevertheless, multicomponent exponential relaxation data provide valuable information for scientific, engineering, medical, and other applications. This report describes and illustrates a set of automatable relaxation analysis protocols that are model-agnostic and suited to extracting information under circumstances when little prior knowledge about the underlying distribution is used. Methods described include: parameterizations based on small geometrically incremented exponential basis sets and orthogonalized or transformed versions of those basis sets, autoregressive and least-squares determination of rates and amplitudes of small numbers of discrete components that fit the data precisely, calculation and comparison of average relaxation rates and average relaxation times to simply describe rate distributions, and quasi-continuous renderings that are regularized in a Fourier domain and that convey the resolution of the analysis via peak widths. Protocols are generalized to treat two-dimensional relaxation measurements. A second part of this report discusses mathematical and physical underpinnings of the methods. This discussion, organized around the Sloppy Model perspective, justifies the methods advocated here, and may be useful for insight and understanding, uncertainty analysis, and to inform practical implementations of multicomponent exponential decay analysis software.

INTRODUCTION

Numerous protocols (Whittall, 1996) (Istratov and Vyvenko, 1999) are available for fitting multicomponent exponential relaxation measurements. Displays of the results of the different protocols can look quite different from each other. This is illustrated in Figure 1, which compares some of the parameterizations and displays discussed in this report. Even though the plotted results from different protocols look quite different, in all cases, the original data are reproduced to high precision. All protocols that fit data to within its noise capture the same information about the relaxing system even if they differ in how that information is conveyed.

It is also apparent that, without additional knowledge or suppositions, the true underlying model for a system cannot be uniquely determined since different models give experimentally indistinguishable fits. As emphasized in Figure 1, even discrete vs. continuous underlying distributions cannot be distinguished. If prior knowledge about the relaxing system is available, the best course of action is usually to use that knowledge for guidance (Mailhiot et al., 2018) (Van Landeghem et al., 2010), and when possible to directly fit parameters in a model to the data. Textbook approaches to parameter estimation and model selection (Bevington, 1969) (Sivia and Skilling, 2006) can be applied in these cases.

The focus of this report is on analysis methods that can be applied with little or no supervision, and that are suited to situations where little or no prior knowledge is available or used beyond assuming that the curves are exponential decays having one or more components. The methods discussed therefore do not attempt to determine which underlying model is most appropriate for the system measured. Rather, the emphasis is on methods that extract and convey the information objectively provided by the measurement in a manner suited to the purpose of the measurement. Envisioned applications include the use of nuclear magnetic relaxation measurements for automated process control, medical diagnostics, raw material screening, quality assurance, and numerous other situations. The methods are also valuable for fundamental and mechanistic investigations if essential limitations of multicomponent exponential decay analysis are considered. Emphasis is placed on numerically efficient and stable parameterizations that can pass results automatically to computers for further analysis. However, conveying information from the analyses to people expecting certain formats is also considered. Though the focus of this report is on magnetic resonance, multicomponent exponential decay curves arise in numerous other circumstances, and the methods advocated here are likewise broadly applicable.

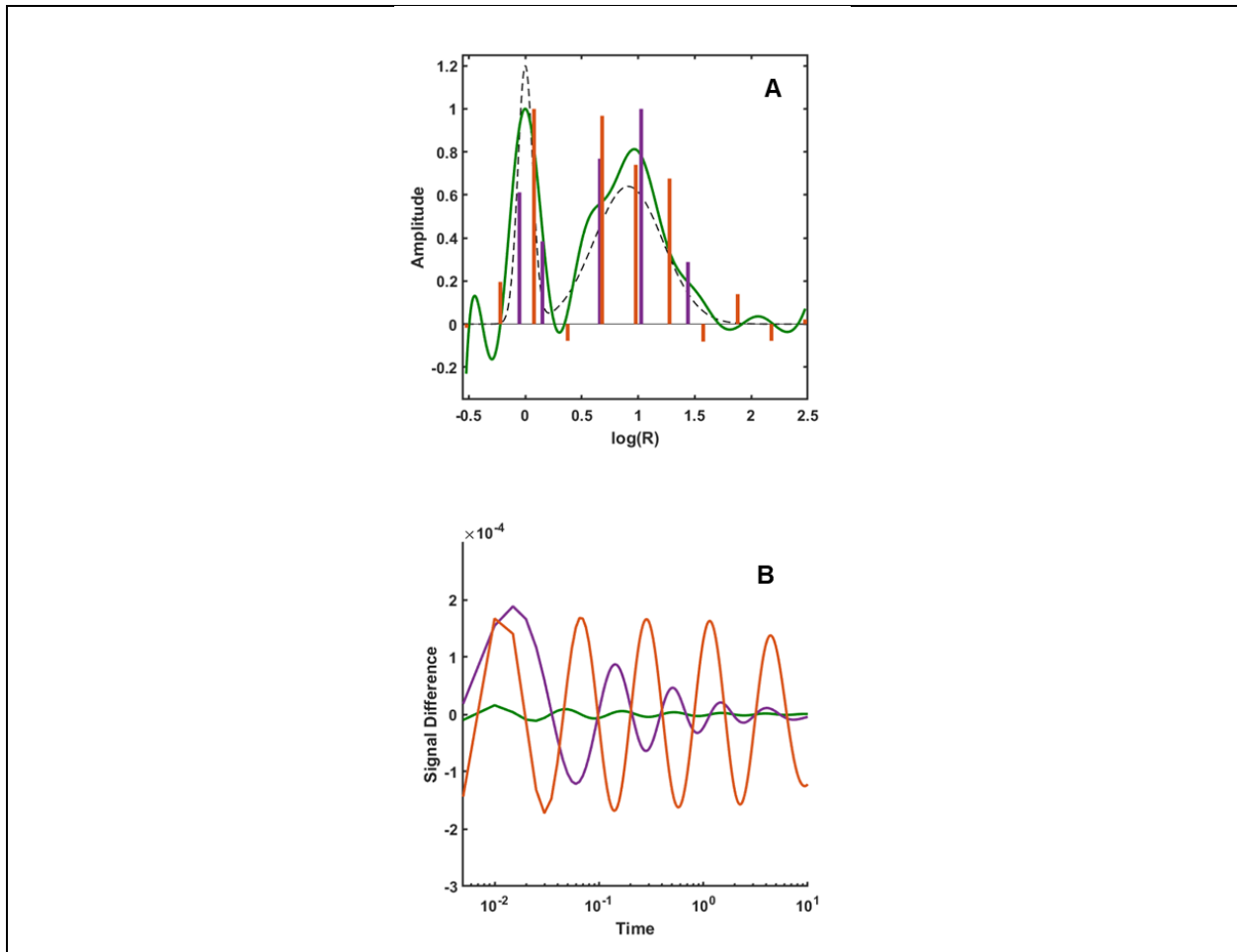


Figure 1. Parameterizations often appear different but fit the data well. (A) dotted black line gives starting distribution of relaxation rates used to generate a multicomponent exponential decay curve; green line shows fitted distribution from SVD truncated to 14 components; orange spikes show amplitudes and rates of a small (11) basis parameter set fitted to the decay curve; purple spikes show discrete fit of five components using ARM+LSQ. (B) differences between decay curves generated by the fits in Panel A compared to the decay curve generated from the starting distribution. All parameterizations reproduce the original decay curve to within 2 parts in 10,000 or better, a higher precision than is often available from real-world measurements. Time and rate units are sec and sec^{-1} , respectively. The decay curve had 2001 evenly spaced points ranging from 0 to 10 seconds. Amplitudes in Panel A are arbitrarily scaled to allow comparison of parameterizations. Differences in Panel B are relative to a decay curve whose initial point is one arbitrary unit.

Overview of the fitting problem

Analysis of multicomponent exponential decay data entails finding a set of spectral amplitudes s_α and relaxation rates r_α to create a modeled data set, $y(t_i)$, that best agrees with a time-dependent signal $d(t_i)$, sampled at time points t_i

$$y(t_i) = \sum_{\alpha=1}^{n_\alpha} s_\alpha e^{-r_\alpha t_i} \quad (1.1)$$

$$d(t_i) = y(t_i) + \varepsilon_i. \quad (1.2)$$

In these expressions, n_α is the number of components, and ε_i is a random noise contribution, assumed uncorrelated in time and normally distributed around zero. For simplicity the independent variable t is referred to as time in this report, but the problem and solutions apply to many different independent variables or functions. As another example from magnetic resonance, t might refer to a function of squared gradient strengths, pulse durations, and delays in magnetic resonance diffusion measurements.

In the limit of continuous distributions of decay rates, the model becomes

$$y(t) = \int_{\alpha} d\alpha s(\alpha) e^{-r(\alpha)t}. \quad (1.3)$$

Being the formal definition of the Laplace transform of the distribution $s(\alpha)$, this expression motivates the prevalent use of the moniker *Inverse Laplace Transform*, *ILT*, for quasi-continuous renderings of $s(\alpha)$ and for algorithms that find such renderings.

This report will frequently describe the multicomponent exponential fitting problem in matrix-vector notation,

$$\mathbf{y} = \mathbf{B} \mathbf{s}, \quad (1.4)$$

$$\mathbf{d} = \mathbf{B} \mathbf{s} + \boldsymbol{\varepsilon} = \mathbf{y} + \boldsymbol{\varepsilon}, \quad (1.5)$$

where \mathbf{y} is the data vector predicted by the model, \mathbf{B} is a basis matrix whose columns contain unit-intensity single component decay curves, \mathbf{s} is a spectral vector containing appropriate amplitudes, \mathbf{d} is a column vector containing the data points, and $\boldsymbol{\varepsilon}$ is the noise vector.

Equation (1.4) can be modified to use different bases by inserting matrices \mathbf{Q} and their inverses \mathbf{Q}^{-1} to transform the basis and spectral vector,

$$\mathbf{y} = (\mathbf{BQ})(\mathbf{Q}^{-1}\mathbf{s}) = \mathbf{B}_Q \mathbf{s}_Q. \quad (1.6)$$

Transformations will include orthonormalization of the basis functions to gain numerous advantageous properties, or conversion to a Fourier basis, which facilitates regularization and its interpretation.

Equations (1.4) and (1.5) (or their transformed analogs) are to be solved for the vector \mathbf{s} and sometimes the matrix \mathbf{B} , to maximize the parameter likelihood function. As usual, it is equivalent and more convenient to minimize the negative of the log of the likelihood function. When the noise is normally distributed around zero, then within an inconsequential additive constant, the negative log likelihood is given by the familiar expression

$$\chi^2 = \sum_{i=1}^{n_d} \frac{(d_i - y_i)^2}{2\sigma^2}. \quad (1.7)$$

Here, the sum is over all n_d data points, and σ is the standard deviation of the noise, assumed here to be equal for all data points.

Protocols for solving Equations (1.1)-(1.5) can mostly be placed into two broad categories, often denoted as discrete and continuous (Whittall, 1996) (Istratov and Vyvenko, 1999). For discrete analysis, a long-standing (Provencher, 1976) favored approach is to estimate or hypothesize initial values for a set of component decay rates and amplitudes, and then to refine those initial values using a least-squares minimization procedure. Given reasonable initial parameter values, computational aspects of the minimization problem are not currently challenging thanks to fast computers and efficient algorithms. However, there is continuing development of principles and methods for discovering the appropriate number of discrete components, for finding initial parameter values, and for alternatives to least-squares parameter estimation. Approaches to finding discrete solutions based on autoregressive methods, and in particular the Matrix Pencil Method (Fricke et al., 2020), appear to be particularly successful, and are adopted here.

Finding continuous distributions $s(\alpha)$, or more precisely quasi-continuous renderings, that fit multicomponent relaxation data is an ill-posed problem—there is no unique solution that is stable to the noise. Regularization, which constrains the fitted distributions in ways intended to suppress unphysical features in the displayed results, must be imposed. Regularization protocols have long been essential to widely used quasi-continuous analysis protocols (Butler et al., 1981) (Provencher, 1982). Many protocols are available, including imposition of non-negativity and the use of Tikhonov regularization. Developing and applying more advanced regularization methods that apply in specific situations, for example when the underlying distributions have both narrow and broad components (Borgia et al., 1998) (Reci et al., 2017), remains an active aspect of research in multi-component exponential relaxation data analysis.

Two-dimensional relaxation measurements have similar advantages to two-dimensional spectroscopy measurements. Advantages include better resolution and more information about interactions and exchange among spin systems having different relaxation rates. Two-dimensional relaxation measurements have been used for decades (English et al., 1991) (Lee et al., 1993), but early applications were hindered by challenges with data analysis. Their use accelerated following the publication of a rapid 2D Inverse Laplace Transform method (Song et al., 2002), and they are now widely applied. An incomplete list of a few of the many current research topics related to two-dimensional relaxation measurements include development of processing algorithms (Su et al., 2019) (Fricke et al., 2020), classification of food products (Greer et al., 2018), porous media characterization (Silletta et al., 2018), and investigations of biomaterials like cartilage (Mailhiot et al., 2018). There does not appear to be widespread agreement on the best methods for processing or for parameterizing the information content of 2D measurements.

Multi-component exponential relaxation analysis is ill-posed because of the high degree of linear dependence among single-component relaxation curves. The origin and extent of linear dependence can be appreciated by inspecting the first terms of the Taylor series for the logarithm of a sum of two decaying exponential functions,

$$\log\left(s_\mu e^{-r_\mu t} + s_\nu e^{-r_\nu t}\right) = \log\left(s_\mu + s_\nu\right) - \frac{s_\mu r_\mu + s_\nu r_\nu}{s_\mu + s_\nu} t + \frac{s_\mu s_\nu}{2(s_\mu + s_\nu)^2} (r_\mu - r_\nu)^2 t^2 + \frac{s_\mu s_\nu (s_\mu - s_\nu)}{6(s_\mu + s_\nu)^3} (r_\mu - r_\nu)^3 t^3 + \dots \quad (1.8)$$

To first order in time, the sum of a pair of components behaves as a single component whose decay rate is the amplitude-weighted average of the components. By similar reasoning, to first order, a single-component relaxation curve can approximate a set of nearby relaxation curves. In this level of approximation, exponential decay curves are linearly dependent. Higher order terms, proportional to powers of $(r_\mu - r_\nu)t$, are responsible for a degree of linear independence among components. Roughly speaking, and depending on the relative amplitudes of the components, these terms become important only at times that are long enough so that the quantity $(r_\mu - r_\nu)t$ becomes significant. If the signal decays into the noise or if the signal is truncated before these terms become apparent, then the data cannot be used to distinguish a distribution of components having relaxation rates in the range between r_μ and r_ν from a single component having the amplitude-weighted average of these relaxation rates.

Equation (1.8) thus gives a simple and useful perspective on why there is no unique quasi-continuous solution and on why multi-exponential decay analysis has low resolution. A single relaxation curve can closely mimic a pair of curves having nearby relaxation rates, and a pair of curves with nearby relaxation rates can closely mimic a single-component curve. The ranges of rates over which these ambiguities apply depend on the noise level and duration of the recorded data. Furthermore, it is apparent that innumerable self-canceling combinations of components can be constructed and included in any model without significantly influencing the decay curve generated by the model.

These properties may be understood as an example of *Sloppy Model* behavior (Transtrum et al., 2015). Sloppy model methods apply when detailed descriptions of a system may involve many underlying details and parameters, but most parameters values are highly or entirely uncertain when fit to data. In these cases, a few combinations of parameters can nonetheless be well-defined. Viewing the multicomponent exponential relaxation problem from this perspective provides insight and informs rigorous and practical approaches to dealing with essential ambiguities. This perspective is described further in Part 2 of this report.

The remainder of this report is organized as follows. Part 1 presents and illustrates a set of analysis tools that meet the goals outlined above. These methods include the use of small exponential basis sets and transformed versions of those basis sets to efficiently capture and record the information available from the data. A practical definition of resolution is offered based on the geometric interval needed to generate basis sets that fit the data adequately. Autoregressive methods, in particular the Matrix Pencil Method, are then used. The resulting parameters are subjected least-squares refinement and to uncertainty analysis using the Fisher Information Matrix and using Monte Carlo explorations of likely parameter values. Average relaxation rates and average relaxation times are then introduced as parameters that partially but usefully describe the effective width of distributions of relaxation rates underlying the measured decay curves. An approach to regularization based on viewing the relaxation rate distributions in a Fourier domain is presented. Unlike other regularization methods, this endows widths of peaks in quasi-continuous renderings with definite and recognizable physical meaning. The last

section of Part 1 generalizes the methods to two-dimensional data sets. Part 2 of the report provides deeper conceptual and mathematical descriptions of some underpinnings of the methods. The discussion in Part 2 is organized around the concept of sloppy models and Fisher Information.

PART 1: DESCRIPTION AND ILLUSTRATION OF METHODS

Small Basis Parameterization Strategies

A pragmatic definition of resolution for multi-exponential fitting

It follows from Equation (1.8) and the accompanying discussion that the amplitudes of only a small set of regularly incremented single-component decay curves can parameterize multicomponent relaxation data, regardless of the richness of the underlying distribution. Quantifying this idea leads to a practical definition of the resolution achievable from a noisy decay curve.

Resolution is expressed as the geometric interval between relaxation rates needed to construct an adequate set of fixed-rate basis curves for use in Equation (1.4). To make this more precise, a definition of the term *adequate* is required. As one definition of the term, consider a geometrically incremented basis set \mathbf{B} and a single component, unit intensity test decay curve \mathbf{y}_{test} , chosen from within the range of relaxation rates spanned by the basis. Denote the spectral vector giving the least-squares fit to that curve as \mathbf{s}_{test} . Somewhere in this fitted curve there is a point of maximum deviation having the value $\delta_{max,test} = \|\mathbf{B}\mathbf{s}_{test} - \mathbf{y}_{test}\|_{max}$. Among all possible test curves \mathbf{y}_{test} in this range, there is a worst case, \mathbf{y}_{wc} , giving the largest maximum deviation $\delta_{max,wc}$. The term *adequate* means that for a basis set built using a given geometric interval, the value of $\delta_{max,wc}$ is less than or equal to some maximum tolerable deviation. This suits present purposes, though other definitions of *adequate* can be devised, depending on the approach to data acquisition, processing algorithm, and error analysis.

Figure 2 illustrates some features of an example of a curve used to determine resolution using this protocol. The horizontal axis shows $\delta_{max,wc}$. The vertical axis is the increment between the basis curves that gives that maximum deviation. If the horizontal axis is interpreted as the maximum tolerable error for a fit to a noisy decay curve, the vertical axis can be interpreted as the resolution.

Compared to the number of points usually displayed when most Inverse Laplace Transform algorithms are used, the resolution achievable for realistic signal-to-noise ratios is quite coarse. For the case considered in Figure 2, a maximum deviation of ~0.3% corresponds to a geometric increment of ~2, and coefficients of only 11 fixed-rate decay curves suffice to parameterize data for a range of relaxation rates spanning three orders of magnitude. As data quality improve, the resolution becomes less sensitive to the maximum tolerable error, slowly and asymptotically approaching a limiting value of 1. According to Figure 2, a range of maximum tolerable errors from 10^{-6} to 10^{-3} corresponds to a range of resolutions between approximately 1.3 and 1.7

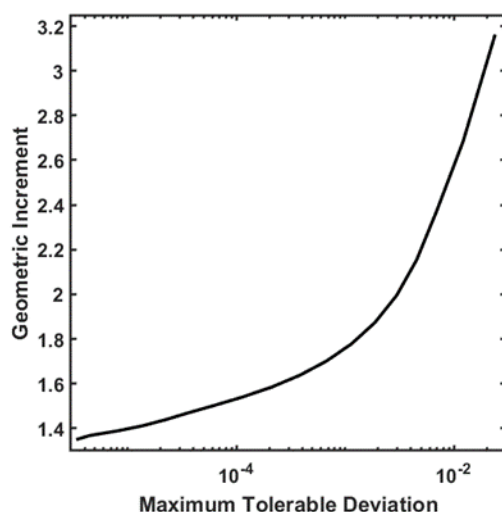


Figure 2. Resolution as a function of maximum tolerable deviation between fitted and true decay curves. Resolution is defined as a geometric increment between relaxation rates used to construct a set of decay curves that can fit data to within a maximum tolerable deviation. Different resolution curves will be obtained depending on the time increment and the range of decay rates included in construction of the basis curves. This curve was constructed using a range of decay rates between 0.3 and 300 sec^{-1} . Basis and fitted curves consisted of 2001 points incremented linearly over 10 seconds. 1000 curves having relaxation rates spaced geometrically throughout the range were tested to identify the worst case.

Implementation of Small and Orthonormal Basis Parameterizations

Since a small number of basis curves with coarsely incremented relaxation rates suffices to precisely fit multicomponent exponential decay curves, a very simple protocol can be used to parsimoniously capture all the information in a relaxation measurement without making any assumptions about the underlying distribution of relaxation rates. This amplitude-only protocol, here denoted SBP (Small Basis Parameterization), starts by picking a series of single-exponential basis components whose fixed rates are geometrically incremented within the range of expected relaxation rates using the resolution factor as the increment. These basis curves are assembled into a basis matrix \mathbf{B} . Then, Equation (1.5) is solved for the spectral vector \mathbf{s} giving the least-squares fit to the data. Because the number of basis curves is low, the spectral vector can be found directly without needing to consider ill-conditioning of the basis matrix or regularization of the spectral vector. An analysis of this type is shown in Figure 1.

Among the notable shortcomings of this small basis parameterization, it is difficult to ascribe clear meaning to the values in the resulting spectral vector. Least-squares fitting using a small basis can exploit small differences of large amplitudes to accommodate physically inconsequential differences between the original and the modeled decay curves. This can lead to coefficients that oscillate, thereby obscuring the physical meaning of the parameters. This shortcoming can be alleviated, and many advantages accrued, if the basis set is appropriately orthonormalized before fitting to the data. For orthonormal basis parameterization (OBP), the fitting function given by Equation (1.4) is modified according to Equation (1.6) to give

$$\mathbf{y} = \mathbf{B}_o \mathbf{s}_o = (\mathbf{B}\mathbf{O})(\mathbf{O}^{-1}\mathbf{s}). \quad (1.9)$$

Here, \mathbf{O} orthonormalizes the small basis matrix \mathbf{B} based on its Singular Value Decomposition (SVD) to give the orthonormal basis \mathbf{B}_o . This choice of SVD for orthonormalization vs. other possible methods such as QR decomposition, and the advantages claimed of OBP in the following paragraphs, are justified and discussed in Part 2 of this report.

One advantage of OBP is that the resulting parameter values are uncorrelated. A second advantage is that, in the Bayesian perspective, the probability distributions for all the orthogonal basis parameter values have the same variance, equal to the noise variance of the data. Another advantage is that each parameter is the scalar product of the orthogonalized basis function with the decay curve and is in that sense a functional of the time-domain data. As will be illustrated, these parameters are simultaneously functionals of the underlying distribution of relaxation rates, and in that sense have discernable physical meaning. The use of spectral-domain and time-domain functionals to parameterize and exploit multi-component relaxation analysis is well established. Often, the functionals are calculated based on ILT-like interpretations of the data (Parker and Song, 2005). The recent use of logarithmic moments calculates useful functionals directly from the data (Petrov and Stapf, 2017).

Figure 3 illustrates some consequential properties of orthogonal basis sets and related Fourier basis sets, discussed in more detail in Part 2 of this report. Panel A shows the orthonormalized basis functions. Higher index curves have increasing numbers of zero crossings. The weightings of intensities from the underlying distribution for these curves are shown in Panel B of Figure 3. Being functionals of the data with the indicated mappings to the rate distributions, orthogonal basis parameters give information about the positions of rates in the underlying distribution. The precision with which the underlying distribution can be specified is low because increasingly rapid oscillations in the rate domain basis functions create increasing amounts of intensity cancellation in the time domain basis functions. Panel C shows the singular values for these basis functions, which reflect the decreasing importance of the orthogonal basis functions in fitting to data. The solid line shows the singular values for a large (256) finely incremented basis. These values fall off very rapidly and approximately exponentially. By about 15 singular values, the importance drops below one part in 10,000. Small basis sets follow nearly the same curve but have somewhat less importance for the higher indices. This provides another indication of the low resolution available from multicomponent exponential analysis.

Panel D of Figure 3 shows consequences along the time axis of using Fourier weightings of the rate distribution (not shown). These time domain functions are not orthogonal though they have been normalized. The advantage of the Fourier basis, as described below, is that it facilitates regularization when pseudo-continuous representation of the information in the decay curve is desired. The Fourier basis is discussed further in Part 2 of this report.

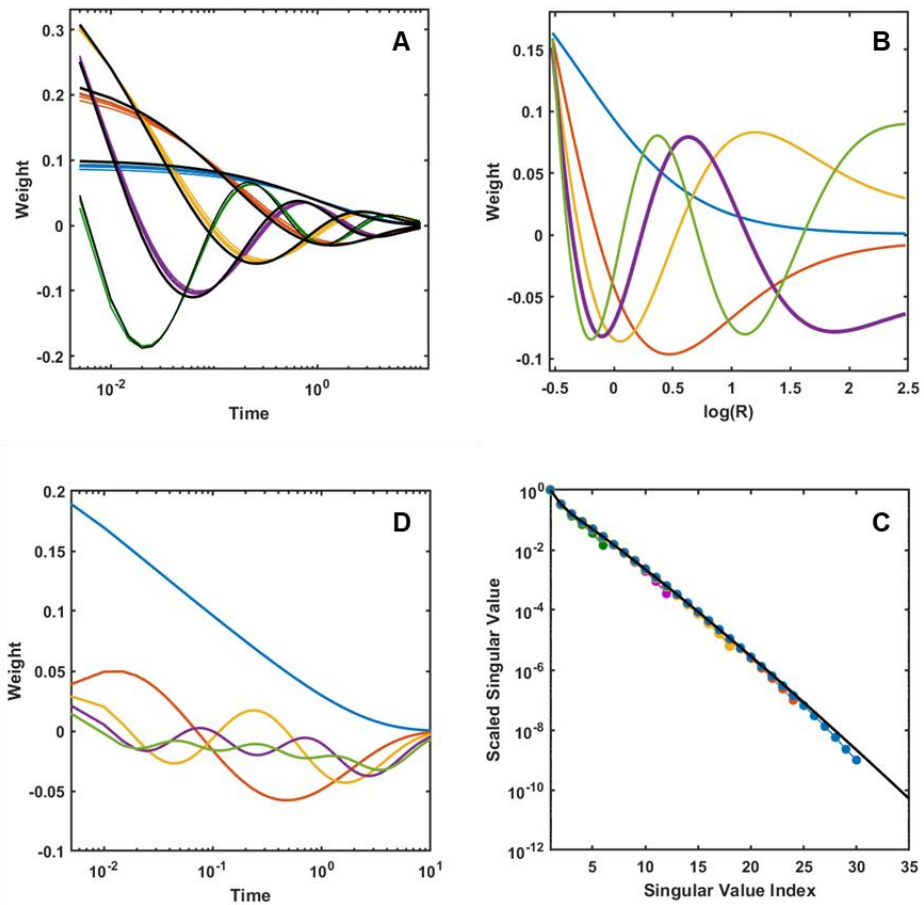


Figure 3. Basis vectors for orthonormal and Fourier basis parameterization. (A) Shapes of the first five orthonormalized basis vectors, plotted on a logarithmic scale to reveal behavior at short times. These were generated using 6, 12, 18, and 25 basis functions geometrically incremented within the same range. Shapes are quite similar and approach the limiting shape from SVD of a large (256 functions) basis set (black curve). (B) Weightings in the spectral dimension corresponding to the orthonormalized basis vectors determined from the large basis set. (C) Comparison of the relative singular values as a function of basis set size. For all basis set sizes, these remain very close to the singular values of a large basis set. The first several singular values and vectors are *stiff* and well defined by the data. The rapid decrease of these values shows that high-index coefficients have low importance for fitting regardless of the size of the basis set. They correspond to *sloppy* parameters. For orthonormalized functions determined by SVD, the number of zero crossings increases with the index into the SVD for both the time-domain and rate-domain weightings. (D) Normalized basis functions generated from Fourier weighting of components of a large (256 function) basis set. Data functionals calculated from curves in D lead to a Fourier series of coefficients that may be easily regularized. To facilitate visualization, signs of all basis functions are adjusted to have their first point non-negative in this figure. Though plotted using logarithmic time axes, the curves were computed and used with linear time spacing.

OBP is especially useful when used for comparisons among different samples thanks to the commensurate parameters it generates. To illustrate the need for commensurate parameters, consider the examples of on-flow measurements or large numbers of related measurements on patients. Results may be compared or used to train or exploit machine learning models. Assume that each measurement is performed using the same experimental protocol with comparable instruments, and that the same orthonormal basis set is used for analysis of each one. Under these conditions, each parameter is a functional that weighs, integrates, and combines precisely the same decay rate ranges of the underlying distribution. Each parameter also reveals unique (orthogonal) information about that underlying distribution. In contrast, a common approach of finding an appropriate number of components, and fitting rates and intensities for those components, may generate incommensurate parameterizations. For example, the approach might find different numbers of components for different samples, in which case it will not be clear how to handle missing components, and it will not be possible to know which components should be compared from sample-to-sample. Orthonormal basis parameterization has none of these ambiguities.

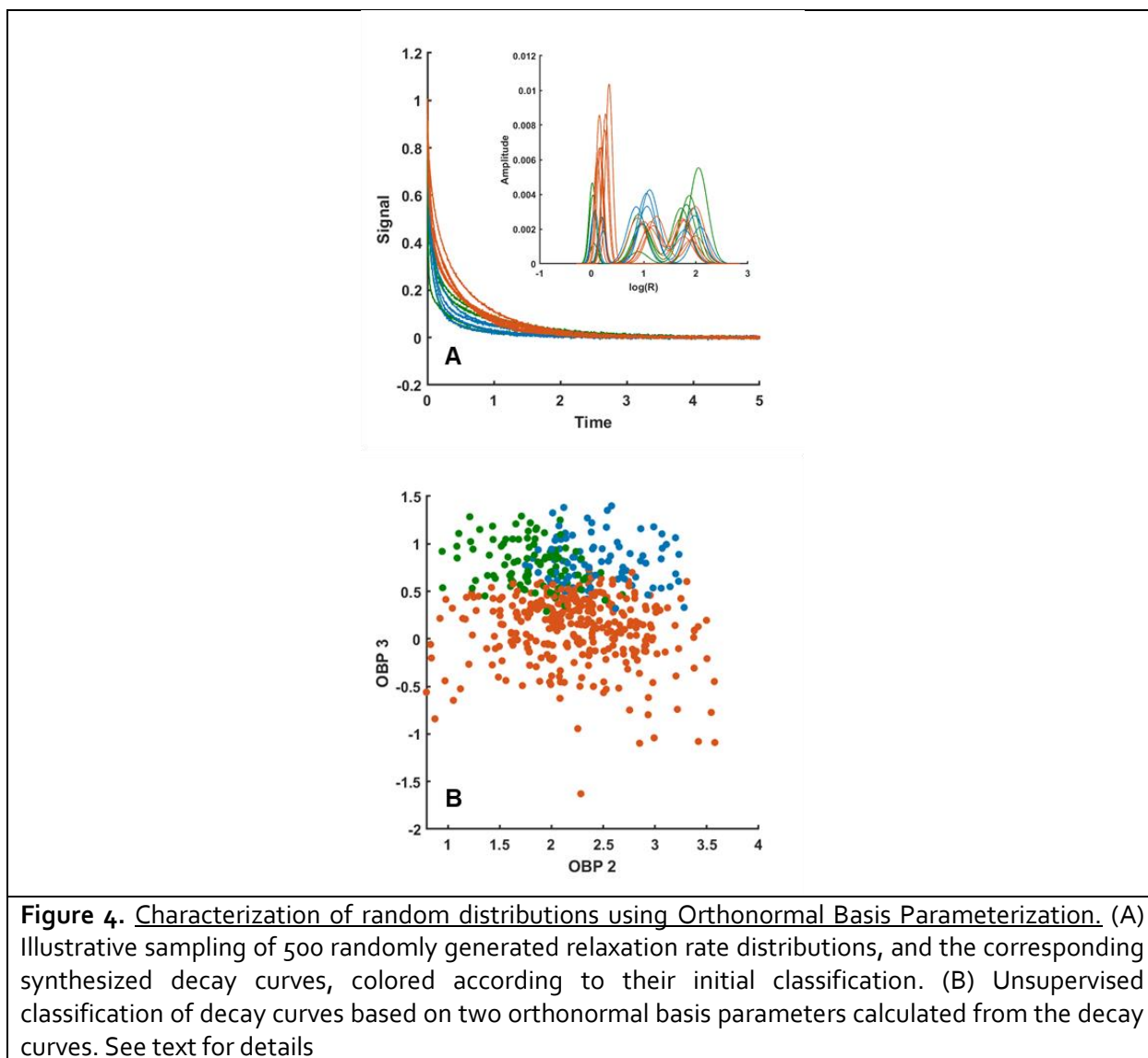


Figure 4 provides an example of OBP-based classification using simulated data. This exercise was intended to mimic the permeation of water into a porous material. For these situations, free water typically has slow transverse relaxation rates, while absorbed water may sample different environments where it has faster relaxation rates. It could be important to distinguish among materials that differ in their pore properties.

To roughly mimic this and to construct Figure 4, 500 relaxation rate distributions consisting of three components each were randomly synthesized to create a test data set. To generate each distribution, intensities of three components were randomly chosen and scaled to a total intensity of 1. To each of the three intensities, a relaxation rate was assigned. These relaxation rates were chosen randomly from within three pre-specified ranges (1.0-2.65, 7.0-18.5, 49.0-129.6) with components in each range having pre-specified widths (log widths = 0.15, 0.4, 0.4). Gaussian (on the log axis) distributions for each component centered on the randomly chosen relaxation rate, with the specified widths and intensities, were summed to create an underlying distribution for each simulated sample. These distributions of relaxation rates were used to generate a decay curve for each simulated sample. The decay curves had 2001 points evenly spaced over 5 seconds, with the slowest and fastest relaxation rates residing in the range 0.5 and 700 per second. Random Gaussian noise with a standard deviation of 0.003 was added to each curve.

Prior to the analysis, samples were arbitrarily classified as "orange" if the slowest rate contained > 25% of the total intensity. The remaining samples were classified as "blue" if the remaining intensity was >50% in the middle rate component, and "green" if the remaining intensity was > 50% in the fastest component. Example relaxation rate distributions and decay curves are shown in Figure 4. The decay curves were processed using OBP, and dot plots of pairs of parameter values were constructed and colored according to the classification. Even without employing sophisticated classification methods, it is evident from Figure 4 that the unsupervised OBP treatment can successfully classify most samples.

Discrete Adjustable Rate Component Methods

Autoregressive approaches and the Matrix Pencil method

As described in references (Fricke et al., 2020) (Hua and Sarkar, 1990) (Sarkar and Pereira, 1995), multicomponent exponential time series can be efficiently analyzed using any of a number of linear algebraic methods. Though several names are used for this class of methods, in this report the term Autoregressive Modeling (ARM) is preferred. Methods in this class rely on model-independent matrix factorizations or polynomial root-finding exercises to extract parameters. The results do have physical meaning if interpreted as exponential models, but they can be applied to non-exponential data. Autoregressive methods used in NMR include various implementations of Linear Prediction (Koehl, 1999), the Filter Diagonal Method (Mandelshtam, 2001), and the Matrix Pencil Method. The Matrix Pencil method is favored in this report.

Autoregressive methods applied to real-valued but noisy multicomponent exponential data can generate large numbers of parameters. Complex-valued components modeling oscillatory behavior often appear among these parameters. These appear as complex-conjugate pairs to ensure a net real-valued contribution, and they typically model noise or artifacts of imperfect measurements, such as oscillations in the first several spin-echoes sometimes observed in transverse relaxation measurements. Component selection heuristics are therefore needed to determine the number of components to extract, and to evaluate if the components are physically meaningful. For example, oscillatory components can be

excluded because they model noise or non-exponential components. Components having decay rates that are fast relative to the decay sampling rate can often be excluded because they model noise in the first few points. Depending on circumstances, components decaying (or growing) much more slowly than the total acquisition time may be rejected as unphysical, but they also can model baseline offsets or equilibrium levels towards which magnetization relaxes.

Least squares refinement

Least-squares optimization of parameters generated by the Matrix Pencil Method usually improves fit quality. A speculative reason for this is that Matrix Pencil implementations involve Singular Value Decomposition, which minimizes squares of deviations in directions perpendicular to the solution vectors. Least-squares fitting, as usually implemented and as used here, minimizes squares of deviations perpendicular to the time axis. These differences in the measure of distance could lead to different optima. In any case, parameter value adjustments resulting from least-squares refinement are usually small. It is nevertheless worthwhile when precise characterization or comparison is needed. In this report, the autoregressive Matrix Pencil Method and component selection heuristics are used to determine the number of justifiable components, and to provide high-quality initial values for the parameters. Least-squares refinement is routinely and reliably used to quickly find nearby local optima. The combined method is herein denoted ARM+LSQ.

Uncertainty analysis of ARM+LSQ parameters

Uncertainty analysis frequently involves the matrix of mixed second derivatives of χ^2 with respect to the parameters, called the Fisher Information Matrix. The inverse of this matrix gives the parameter variance-covariance matrix. From the Bayesian perspective, the diagonal elements of a parameter variance-covariance matrix give lower bounds for variances for the parameter values, assuming the distribution of likely parameter values is normal. These variances (or the corresponding standard deviations) are easily calculated and reported with the parameters themselves. Monte Carlo methods can also be used to explore ranges of parameter values. For data sets where more comprehensive uncertainty analysis is justified, or when formulas for the Fisher Information Matrix elements are not available, it can be advantageous and informative to carry out such explorations. A more detailed treatment of parameter uncertainty analysis is provided in Part 2 of this report.

Figure 5 shows the results of using ARM+LSQ to analyze a transverse relaxation measurement of a sample of ground roasted coffee beans mixed with an equal weight of water. Results are presented as a spike plot, where the positions of the vertical spikes indicate the relaxation rates of components, and the heights of the spikes indicates the amplitudes of the contributions of those components. Component selection was based on finding the maximum number of components that yielded exclusively real-valued relaxation rates and amplitudes. A maximum of 6 components could be justified. Attempting to extract a smaller number of parameters gives components that are spaced more widely. The uncertainties in these parameters were determined using the Fisher Information Matrix as described above and in more detail in Part 2 of this report. The pseudo-continuous distributions reflect the resolution of the measurement and were found using the convolution method described in detail below.

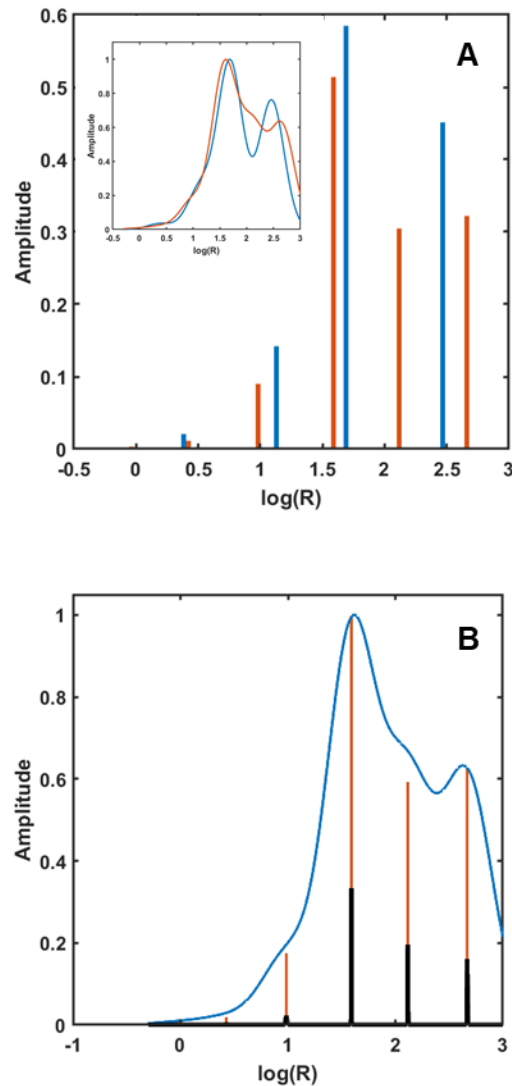


Figure 5. Discrete fits using ARM+LSO, related pseudo-continuous renderings, and uncertainty analysis. Data are from a transverse relaxation measurement of equal weights of water and roasted ground coffee beans. (A) Spike plot showing rates and their amplitudes for 4-component (blue) and 6-component (orange) fits. Inset is the pseudo-continuous rendering of these fits obtained by convolving the spike plot with a Gaussian function to show the resolution of the data. (B) Preferred form for displaying results. The spike plot showing the discrete fit of six components (orange) conveys the information available from the data. The superimposed pseudo-continuous distribution (blue) conveys the resolution available from the measurement. The parameter uncertainties (heavy black line) are encoded as the widths of Gaussian functions. On this scale, the uncertainties are small relative to the line thickness. Figure 7 shows a case where noisier data leads to greater, more easily visualized, uncertainties.

Average Relaxation Rates and Average Relaxation Times

It is frequently useful to summarize properties of a distribution using a few parameters, for example, using moments or other data functionals. In this work, compact and simple summary statistics are provided for the multicomponent exponential relaxation problem by the average relaxation rate, $\langle R \rangle$, and the reciprocal of the average relaxation time, $1/\langle T \rangle$. These are denoted herein as the *rate-average relaxation rate*, and the *time-average relaxation rate*, respectively. Rate-average relaxation rates weigh rapidly relaxing components more heavily, whereas time-average relaxation rates weigh slowly relaxing components more heavily. In general $\langle R \rangle \geq 1/\langle T \rangle$, and the values are the same only when there is a single sharp component. It is also noteworthy that for a data trace $d(t)$, $\langle R \rangle$ is the initial slope of a plot of $\log(d(t))$ vs. t , while $\langle T \rangle$ is the integral of the trace if the first point is normalized to 1 and if the trace is measured for a time long compared to its slowest relaxation rate. $\langle R \rangle$ is subject to increased uncertainty if the initial part of the decay curve is not adequately sampled, and $1/\langle T \rangle$ is subject to increased uncertainty if the decay curve is not sampled long enough.

Underlying distributions can be partially characterized using these parameters. For example, if the underlying rate distribution is assumed log-normal, and if μ and σ are the location and width parameters in the log-normal distribution, then

$$\log(\langle R \rangle) = \mu + \frac{\sigma^2}{2} \quad (1.10)$$

$$\log\left(\frac{1}{\langle T \rangle}\right) = \mu - \frac{\sigma^2}{2}. \quad (1.11)$$

Whether the underlying distribution is log-normal or not, the size of the difference between these averages (or their logarithms) gives a measure of the width of that distribution. The idea and intent are reminiscent of the weight-average and number-average molecular weights that are widely used in polymer chemistry as simple measures that characterize polymer polydispersity. In practice, calculating rate-average relaxation rates and time-average relaxation rates is conveniently achieved using amplitudes and rates from either the SBP analysis or from ARM+LSQ.

Quasi-Continuous Renderings

Widely used Inverse Laplace Transformation algorithms applied to magnetic resonance relaxation data do not produce mathematically unique and reversible transforms, but instead produce interpretations of the data. In this sense, the ILT moniker can be somewhat misleading, and in this report the phrase *quasi-continuous rendering* is preferred. As usually understood, these renderings of amplitude as a function of relaxation rate must be constructed using many more than the 10~15 basis amplitudes that can be determined uniquely from the data alone. Many different amplitude combinations can be found that are mutually compensatory and that can assume wide ranges of values while remaining fully compatible with the data, as discussed in the Introduction and described in more detail in Part 2 of this report.

Regularization can be understood from this perspective as choosing values of these ill-defined parameter combinations to suppress undesirable features appearing in the spectrum of relaxation rates. In this view,

the highly parsimonious spike representations resulting from ARM+LSQ are objectively regularized because all undefinable parameter combinations are set to zero. However, by tradition, and in the sense used in this report, regularization generates smooth, sometimes broad distributions of relaxation rates, even if the data come from discrete underlying source distributions. The widths of peaks in the distributions depend not only on the source distributions and the noise, but also on the regularization algorithm and the degree of regularization applied.

It is reasonable to ask, and hard to answer: When traditional regularization methods are used, what is the physical meaning of the intensity at a given relaxation rate? Non-expert users of quasi-continuous renderings hold tacit and reasonable expectations concerning the physical meaning of these renderings. Reasonable expectations include having a spectrum-like or chromatogram-like appearance, a correspondence of high intensities with the presence of material in the sample having relaxation rates in the vicinity of the intensity maxima, and a lack of visually prominent features that may inadvertently suggest the presence of information that is not justified by the data. It is fair to expect that relatively broad peaks should correspond to sets or distributions of components that span a wide range of relaxation rates.

Part 2 of this report motivates and describes an approach to performing regularization in a controlled way that leads to quasi-continuous renderings that meet these reasonable expectations. Since there is no established meaning for the widths of peaks in these quasi-continuous renderings, in this work a meaning is chosen and imposed to add to the information conveyed by the plots. In particular, the peak widths are regularized to convey the resolution of the data. The approach tames excursions of the ill-defined parameter combinations by regularization in a Fourier domain.

A brief and approximate justification of the method, different from that presented in Part 2, begins by examining the coefficients for constructing ill-defined parameter combinations. These coefficients have fluctuating signs when arrayed along the relaxation rate axis, leading to near cancellation of intensities in the resulting curves along the time axis. The fluctuations have frequency components above that set by the resolution of the measurements. Dealing with rapidly fluctuating ill-defined components is straightforward in the Fourier domain because such high frequency components can be trivially identified and damped or zeroed. Transforming back to the relaxation rate domain produces quasi-continuous renderings having the reasonably expected properties listed above. The resolution of the rendering in the rate domain can be controlled through the number of Fourier components used and the way they are damped.

A very simple application of this idea leads to the following straightforward algorithm for quasi-continuous renderings. The method does not even require overt transformation to and from the Fourier domain. Begin with an initially empty spectral vector with a fine spacing between rates. Parameterize the data using the ARM-LSQ method. For each coefficient in the discrete parsimonious representation of \mathbf{s} , enter the amplitude of the component at the position in the spectral vector that is closest to the rate associated with the component. This is called a spike representation in this report. Convolve the resulting spike representation with a broadening function whose width corresponds to the resolution of the data. A reasonable choice for broadening is a Gaussian shape, though more carefully chosen shapes may be appropriate. Behind the scenes, convolution algorithms may work in a Fourier domain, but that detail is unimportant for practical implementation using high-level programming tools.

As shown in Figure 5, simultaneous display of both the spike plot and the quasi-continuous convolved representation is very useful for visually communicating results to people. Parameter uncertainty can

also be simultaneously shown in such plots. The spike representation of discrete components conveys a parsimonious encapsulation of nearly all the information available from the measurement. If the underlying distribution were discrete, this would be the best representation of that distribution achievable with the number of parameters allowed by the heuristics. The width of signals in the quasi-continuous rendering conveys the resolution of the measurement. Parameter uncertainties from the Fisher Information Matrix can be revealed by superimposing corresponding widths on the spike plot. Investigators who prefer the continuous style of display find all the information extracted from the data in plots such as Figure 5-B. This includes parsimonious discrete component amplitudes and rates, their uncertainties, and the resolution of the data.

Applications to Two-Dimensional Measurement

All the protocols listed above are adaptable to processing and display of two-dimensional measurements. A few aspects of this will be illustrated in this section. Equations (1.12) and (1.13) give expressions for a predicted two-dimensional relaxation spectrum \mathbf{Y} , that agrees with the data matrix \mathbf{D} .

$$\mathbf{Y} = \mathbf{B}_c \mathbf{S} \mathbf{B}_r \quad (1.12)$$

$$\mathbf{D} = \mathbf{Y} + \mathbf{N} = \mathbf{B}_c \mathbf{S} \mathbf{B}_r + \mathbf{N} \quad (1.13)$$

Here, \mathbf{S} is the spectral matrix, \mathbf{B}_c is the column-spanning exponential basis matrix, \mathbf{B}_r is the row-spanning exponential basis matrix, and \mathbf{N} contains the noise.

Two-dimensional Small Basis Parameterization is almost identical to the one-dimensional case. Fixed-rate basis matrices, \mathbf{B}_c and \mathbf{B}_r , are constructed using unit intensity relaxation curves having geometrically incremented relaxation rates. The increments are chosen according to the resolution, and ranges are chosen to span the column and row spaces of the data matrix, just as with the one-dimensional case. Their inverses are used to find a least-squares solution to Equation (1.13) using standard methods. For Orthonormal Basis Parameterization, the procedure is the same except the basis matrices are orthonormalized first using SVD, just as done in the one-dimensional case. For both small basis and orthonormal basis cases, the result is a matrix of coefficients stored in the spectral matrix.

For autoregressive methods in general, Equation (1.13) can be solved uniquely for a chosen rank because \mathbf{B}_c and \mathbf{B}_r are required to have a very limiting and specific (Vandermonde) form, or a least-squares approximation of that form. Recent work (Fricke et al., 2020) describes successful use of the matrix pencil method for analyzing two-dimensional relaxometry data sets. In this report, output parameters are refined using least-squares optimization, as was done for the one-dimensional case. Complications arise for certain two-dimensional protocols including the longitudinal-transverse (T1-T2) measurement, where relaxation towards non-zero magnetization occurs in one or both dimensions. In these cases, rows (or columns) of the spectral matrix are overtly linearly dependent. Equation (1.13) must be appropriately modified in these cases.

For two-dimensional data, uncertainty analysis of ARM+LSQ results using the Fisher Information Matrix approach is possible in principle. However, Monte Carlo methods are used in this report for parameter uncertainty analysis in two-dimensional measurements.

Figure 4 shows analysis of a T₁-T₂ measurement applied to a mixture of ground roasted coffee beans and water (1:1 by weight). Panels A and B show the four components recommended by component selection heuristics for each dimension. The dotted lines show the decay curves calculated from parameters generated by the Matrix Pencil method. The colored lines show the corresponding matrix rows and columns that the algorithm generates during its factorization of the data matrix. The very slowly decaying components correspond to the expected behavior of the offset and equilibrium magnetization for the measurement. To perform the analysis, spacing between increments in the indirect dimension must be even, but in the interest of experimental efficiency, the spacing between time increments was spaced geometrically. Interpolation (cubic) of the values in the indirect dimension was performed to create regularly incremented time values in the indirect dimension. The interpolated curves are shown in Panel C, along with the residuals.

Following two-dimensional ARM, least-squares refinement and MC explorations was performed to generate Panel D. Intensities for each step were summed in a matrix of bins, each bin encompassing a small range of relaxation rate values around those labeling the bin. A contour plot of the results is given in Panel D. Note that one of the cross peaks is negative, a not uncommon finding in this laboratory. A more careful study involving models and simulations is necessary to fully interpret these results. The MC calculations did include variation of the offset and the equilibrium magnetization, as well as the flip angle (see discussion in Part 2). These parameters can contribute to marginalized uncertainties since they may vary from sample-to-sample due to experimental circumstances. They are not included in such plots because they offer no insight into the sample itself.

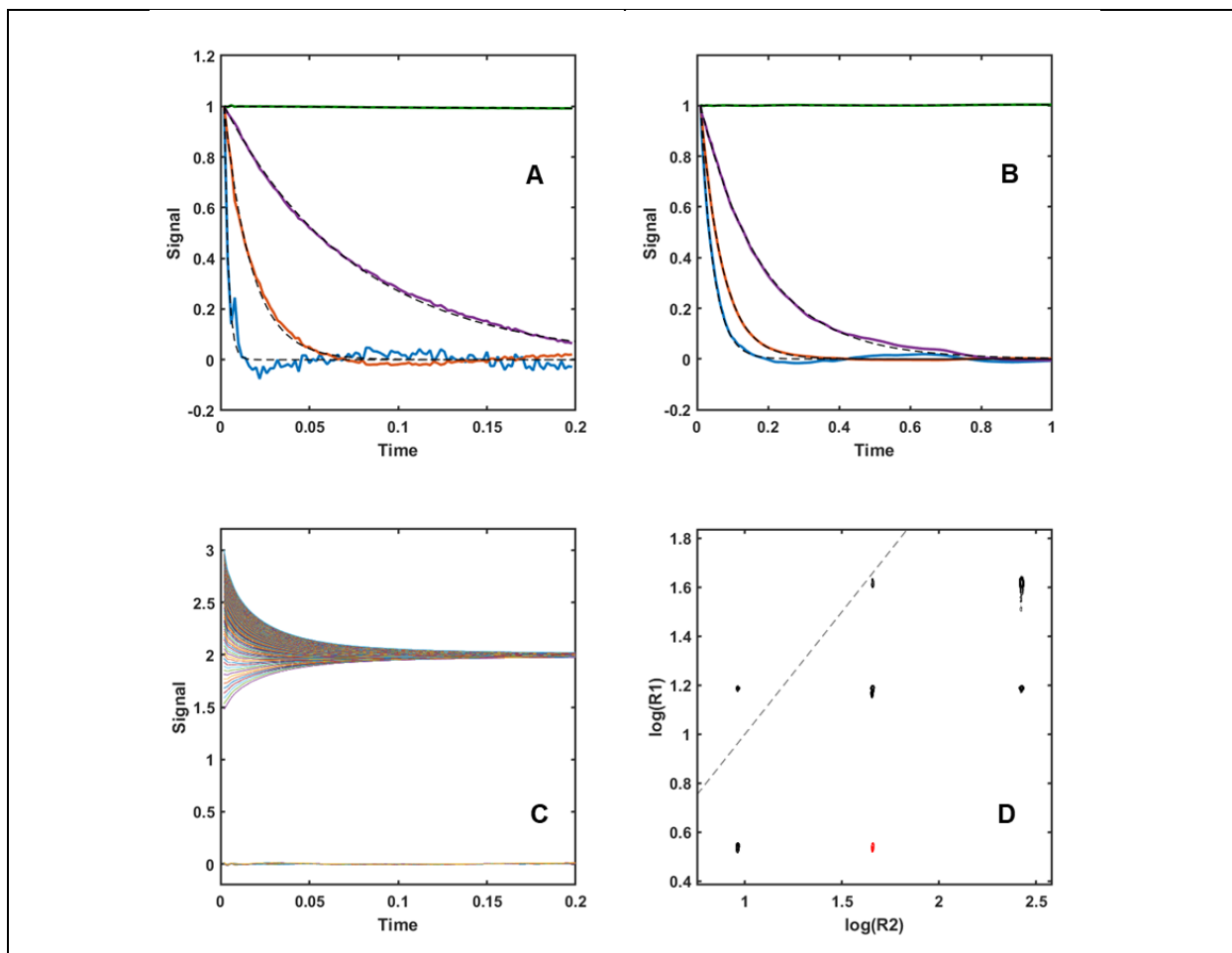


Figure 6. Analysis of longitudinal-transverse two-dimensional relaxation data using ARM+LSQ. Transverse (A) and longitudinal (B) decay components extracted from the Vandermonde matrices determined using the Matrix Pencil method. Based on the component selection heuristics, four components including offsets were computed for each dimension. Dotted lines are based on the ARM-derived relaxation rates, noisy lines are corresponding components that emerge from data matrix factorization using that algorithm. (C) Interpolated original data with offset applied, and residuals after ARM+LSQ analysis. (D) Distributions of rates and intensities displayed by a contour plot calculated using Monte Carlo explorations of probable parameter values. The dotted line shows $R_1=R_2$, and signals above this line must arise from exchange or cross relaxation. Positive and negative intensities are shown in black and red, respectively. For the two-dimensional data set, in contrast to the one-dimensional measurement in Figure 5, fewer points were obtained, and a different sample was used under less optimal conditions, so only 3 non-zero rate components could be obtained.

PART 2: DETAILS AND JUSTIFICATIONS OF THE PROTOCOLS

Fisher Information and Sloppy Models in relation to multicomponent exponential relaxation data

Definitions of Sloppy Models and Fisher Information

Multi-exponential fitting exemplifies a common occurrence in models of complex systems. In detailed descriptions of complex systems, numerous parameters are often used to fully describe the underlying behavior. When examined by a non-specific measurement, many combinations of the detailed parameters become mutually compensatory, and only a few parameter combinations remain relevant to describing the measurement results. Models exhibiting this type of behavior have been denoted *sloppy* (Transtrum et al., 2015). In the context of sloppy models, parameter combinations that make little or no difference to predicting or fitting the data are also called *sloppy*. The few remaining *stiff* parameter combinations capture the experimentally available information and serve as useful descriptors of the measurement. Sometimes, stiff parameter combinations correspond to familiar macroscopic descriptors and properties of systems (Machta et al., 2013). For example, a diffusion coefficient is a stiff parameter that can result from underlying molecular dynamics whose detailed description may be enormously complicated. As one of these sloppy models, multicomponent exponential relaxation can be approached through identification of stiff parameter combinations, determination of stiff parameter values that capture the information available from the data, and determination of the physical meaning of those stiff parameter combinations. Recognition of the sloppy parameter combinations is also valuable, allowing them to be appropriately handled.

Sloppy model publications offer a useful approach for finding the stiff and sloppy parameter combinations based on Fisher Information. Fisher information is well known and used in several fields but, to the knowledge of the author, is not currently used in magnetic resonance relaxation analysis. Useful properties of Fisher Information have been concisely summarized (Coe, 2009) and didactic introductions are found via the sloppy model citations given above, and elsewhere (Caticha, 2015). Since Fisher Information can be used as a unique metric defining distances between probability distributions, it is a foundational concept in the field of Information Geometry (Amari, 2016). Here, only a simplified view of Fisher Information can be provided, and a few of the key results used in this report are listed.

Optimized parameters lie in minima of plots of χ^2 vs. parameter values. For a single parameter problem, if a minimum is steep, characterized by a relatively high second derivative, then the parameter can be considered well-defined, and the information content of data concerning that parameter is large. When a minimum is shallow, characterized by a relatively low second derivative, then the parameter can be considered poorly defined, and the information content of the data concerning that parameter is small. For a single parameter, the Fisher Information I is defined as the second derivative of χ^2 with respect to that parameter. Since the variance in the distribution of a parameter value will be low if it lies in a steep minimum, it is reasonable that the reciprocal of the Fisher information gives the parameter variance.

For multiparameter problems, the matrix of mixed second partial derivatives of χ^2 with respect to the parameters (the Hessian of χ^2) characterizes the information content of the data and is called the Fisher Information Matrix, \mathbf{I} , with elements

$$I_{\alpha,\beta} = \frac{\partial^2 \chi^2}{\partial s_\alpha \partial s_\beta}. \quad (2.1)$$

When the parameter values covary, then the mixed partial derivatives are non-zero, and \mathbf{I} has off-diagonal elements. Another, usually equivalent, expression for Fisher information matrix elements is

$$I_{\alpha,\beta} = \frac{1}{\sigma^2} \sum_{i=1}^{n_d} \frac{\partial y_i}{\partial \alpha} \frac{\partial y_i}{\partial \beta} \quad (2.2)$$

This form follows from Equation (2.1) if it is assumed that the expressions for y_i can be approximated linearly over relevant parameter ranges.

As noted, the inverse of \mathbf{I} predicts the parameter covariance matrix assuming that χ^2 depends linearly on parameter values. However, if the dependence is not linear over a sufficiently large span of parameter values, or if other conditions hold, then the inverse of \mathbf{I} gives only a lower bound on the marginalized variances and covariances, called the Cramér-Rao lower bound. The qualifier *marginalized* is important—it implies that the method gives variances that are larger than they would be if the parameters were varied independently because couplings and mutual compensatory behaviors among parameters are considered.

Eigenvalue decomposition of \mathbf{I} identifies the stiff and sloppy parameter combinations. Eigenvectors corresponding to the large eigenvalues define stiff parameter combinations, while those with small eigenvalues define sloppy parameter combinations.

$$\mathbf{I} = \mathbf{V} \mathbf{\Gamma} \mathbf{V}^T \quad (2.3)$$

Here, $\mathbf{\Gamma}$ is a diagonal matrix containing the eigenvalues of \mathbf{I} , and \mathbf{V} is an orthonormal matrix of eigenvectors. The coefficients of the raw parameters that combine into orthogonal parameter combinations are provided by elements of these eigenvectors. It is assumed throughout this report that the eigenvalues, and the corresponding eigenvectors, are sorted from largest to smallest. It will be seen below that the eigenvalue decomposition of \mathbf{I} also leads to the Orthogonal Basis Parameterization strategy advocated in this report.

Fisher Information expressions for multi-component relaxation analysis

To explore the use of \mathbf{I} for the multi-component relaxation analysis problem, first consider a basis matrix \mathbf{B} for the amplitude-only case. This consists of mono-exponential decay curves whose relaxation rates are fixed and spaced geometrically between a lower and upper limit. Using this basis matrix and solving Equation (1.5) gives an estimate of the spectral vector. We can evaluate \mathbf{I} using either Equation (2.1) or (2.2), which both give

$$I_{\alpha,\beta} = \frac{1}{\sigma^2} \sum_{i=1}^{N_d} \exp(-r_\alpha t_i) \exp(-r_\beta t_i) \quad (2.4)$$

Note that when applied to the amplitude-only problem, this matrix is independent of the amplitudes assigned to the components. \mathbf{I} can therefore be written for the amplitude-only case in matrix notation as

$$\mathbf{I} = \frac{1}{\sigma^2} \mathbf{B}^T \mathbf{B} \quad (2.5)$$

A continuous limit expression for the elements of \mathbf{I} can be obtained by replacing the discrete sums in (2.4) with continuous functions that extend to infinity, and by integrating symbolically, giving

$$\hat{I}_{\alpha,\beta} = \frac{1}{\sigma^2} \left(\frac{1}{r_\alpha + r_\beta} \right) \quad (2.6)$$

The carat emphasizes that matrix elements are determined symbolically rather than from discrete sampling. $\hat{\mathbf{I}}$ is very similar to the notorious Hilbert matrix, whose elements are given by $H_{\mu,\nu} = 1/(\mu + \nu - 1)$, where μ and ν are row and column indices. The Hilbert matrix is a prototypical example of a severely ill-conditioned matrix, so $\hat{\mathbf{I}}$ is expected and observed to be similarly ill-conditioned when large numbers of basis curves are used.

Next consider the case wherein both the amplitudes and the logarithms of rates are adjusted. (Different expressions are obtained when the rates are varied linearly.) For this case it is convenient to arrange the Fisher matrix block-wise. The two diagonal blocks give derivatives involving amplitude-amplitude cross terms, $I_{\alpha,\beta}^{s,s}$, and rate-rate cross terms, $I_{\alpha,\beta}^{r,r}$, while the non-diagonal blocks mix derivatives with respect to amplitudes and rates, $I_{\alpha,\beta}^{s,r}$.

The resulting expressions for Fisher matrix elements are

$$I_{\alpha,\beta}^{s,s} = \frac{1}{\sigma^2} \sum_{i=1}^{n_d} \exp(-(r_\alpha + r_\beta)t_i) \quad (2.7)$$

$$I_{\alpha,\beta}^{r,r} = \frac{s_\alpha s_\beta r_\alpha r_\beta}{\sigma^2} \sum_{i=1}^{n_d} t_i^2 \exp(-(r_\alpha + r_\beta)t_i) \quad (2.8)$$

$$I_{\alpha,\beta}^{s,r} = \frac{-s_\beta r_\beta}{\sigma^2} \sum_{i=1}^{n_d} t_i \exp(-(r_\alpha + r_\beta)t_i) \quad (2.9)$$

Truncated Singular Value Decomposition of the basis matrix in relation to Fisher Information
Singular Value Decomposition (SVD) is a widely used matrix factorization algorithm that frequently plays a critical role in solving ill-posed problems like Equation (1.5). SVD is also theoretically related to eigenvalue decomposition of the Fisher information matrix. This enables practical algorithms for finding stiff parameter combinations in the context of Sloppy Models and underlies the orthonormalization strategy suggested here for OBP.

SVD factorizes the basis matrix into a product of three other matrices, traditionally written as

$$\mathbf{B} \xrightarrow{SVD} \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (2.10)$$

The columns of \mathbf{U} form an orthonormal basis set that spans the column space of \mathbf{B} , while the rows of \mathbf{V}^T form an orthonormal basis set that spans the row space of \mathbf{B} . $\mathbf{\Sigma}$ is a diagonal matrix whose elements are the singular values. The singular values weigh the importance of the singular vectors in reproducing \mathbf{B} . The singular values are sorted from largest to smallest, and the corresponding vectors in \mathbf{U} and \mathbf{V} are sorted accordingly.

Orthogonality properties of SVD matrix factors allow construction of a pseudo-inverse which can be used to solve (1.5)

$$\mathbf{B}^{-1} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T \quad (2.11)$$

In the present case, since the columns of \mathbf{B} have a low degree of linear independence, only a limited number of the singular values are numerically significant. This is illustrated in Figure 3, where the singular values (which correspond to square roots of the eigenvalues of the Fisher Information Matrix as demonstrated below) are shown to decay approximately geometrically. For singular values that are very small, elements of the matrix $\mathbf{\Sigma}^{-1}$ become unmanageably large, and their use in data fitting becomes overly sensitive to noise.

A simple and powerful way to deal with reciprocals of vanishingly small singular values recognizes that smaller singular values and the corresponding singular vectors, corresponding to sloppy parameter combinations, refer to the (effective) null space of the basis matrix. Sloppy, null space components make no discernable contribution to the quality of the fit. Their fitted magnitudes give no useful information about the relaxing system. SVD allows these components to be handled appropriately. The simplest way to handle them is to discard them. If one keeps only the first k numerically relevant columns of \mathbf{W} and \mathbf{V} , and keeps the upper left k -by- k sub-matrix of $\mathbf{\Sigma}$ to create the truncated matrices $\mathbf{U}_{(k)}$, $\mathbf{V}_{(k)}$, and $\mathbf{\Sigma}_{(k)}$, then the product

$$\mathbf{B}_{(k)} = \mathbf{U}_{(k)}\mathbf{\Sigma}_{(k)}\mathbf{V}_{(k)}^T \quad (2.12)$$

is the best least-squares rank- k representation of the original matrix \mathbf{B} . Based on this, a naive method for solving Equation (1.5) for a large, finely incremented basis set involves the inverse of this truncated singular value decomposition. The result is

$$\mathbf{s}_{(k)} = \mathbf{B}_{(k)}^{-1}\mathbf{d} = \mathbf{V}_{(k)}\mathbf{\Sigma}_{(k)}^{-1}\mathbf{U}_{(k)}^T\mathbf{d}. \quad (2.13)$$

The subscript k emphasizes that only k independent parameters make their way into the analysis, even if the results are conveyed using a quasi-continuous spectrum- or chromatogram-like display with $\mathbf{s}_{(k)}$ having hundreds or thousands of points. This expression was used to generate the green curve in Figure 1A. In this report, the amount of truncation applied to SVD is always stated in context. Therefore, the subscript k is dropped throughout.

There is a useful connection between the Fisher Information Matrix \mathbf{I} and SVD of the basis matrix \mathbf{B} . Theory shows that, upon singular value decomposition of \mathbf{B} , the matrix \mathbf{V} contains the eigenvectors of $\mathbf{B}^T\mathbf{B}$, while the diagonal elements of $\mathbf{\Sigma}$ are the square roots of the corresponding eigenvalues, so

$\Gamma = \Sigma^2$. According to Equation (2.3), \mathbf{I} is proportional to the same product $\mathbf{B}^T \mathbf{B}$. For the amplitude-only fitting problem, the eigenvalue decomposition of \mathbf{I} is therefore related to the SVD of \mathbf{B} as follows

$$\mathbf{I} = \frac{1}{\sigma^2} \mathbf{V} \Gamma \mathbf{V}^T = \frac{1}{\sigma^2} \mathbf{V} \Sigma^2 \mathbf{V}^T. \quad (2.14)$$

The \mathbf{V} matrices are identical in eigenvalue decompositions and in singular value decompositions of Fisher information matrices in cases where those matrices are sufficiently well-conditioned. This observation is useful because SVD can be used in practice for orthonormalizing basis sets, even when they are ill-conditioned.

Orthonormal Basis Parameterization interpretation and relation to Fisher Information

For the amplitude-only case, the Fisher Information Matrix is given by Equation (2.5). Orthonormalization of the basis set \mathbf{B} to get \mathbf{B}_o gives another Fisher information matrix in that new basis

$$\mathbf{I}_o = \frac{1}{\sigma^2} \mathbf{B}_o^T \mathbf{B}_o = \frac{1}{\sigma^2} \mathbf{E} \quad (2.15)$$

where, since \mathbf{B}_o is orthonormal, \mathbf{E} is the identity matrix. Thus, for an orthonormal basis, the inverse of the Fisher Information Matrix gives the same Cramér-Rao lower bound, σ^2 , for all orthonormalized parameters.

Considering Equations (2.5) and (2.15) for the Fisher Information Matrix in the two basis sets, and finding a transformation that converts \mathbf{I} to \mathbf{I}_o , shows that \mathbf{B}_o can be obtained from \mathbf{B}

$$\mathbf{B}_o = \mathbf{B} \mathbf{O} \quad (2.16)$$

$$\mathbf{O} = \mathbf{V} \Gamma^{-\frac{1}{2}} = \mathbf{V} \Sigma^{-1} \quad (2.17)$$

Equivalently, when \mathbf{B} is subjected to SVD according to Equation (2.12), the orthonormalized version is

$$\mathbf{B}_o = \mathbf{U}. \quad (2.18)$$

Though other orthonormalization strategies are possible, such as QR decomposition, using Equation (2.17) seems most natural given its relation to the Fisher Information Matrix and the Sloppy Model perspective adopted in the report. In terms of \mathbf{B}_o the spectral vector \mathbf{s}_o is found by optimizing the following expressions

$$\mathbf{y} = \mathbf{B}_o \mathbf{s}_o \quad (2.19)$$

$$\mathbf{d} = \mathbf{B}_o \mathbf{s}_o + \boldsymbol{\varepsilon} \quad (2.20)$$

Parameterizations using either \mathbf{B} or \mathbf{B}_o convey identical information because the bases and parameters can be transformed one into the other.

It is worthwhile to consider if the orthogonal basis parameters have a physical interpretation. Figure 3 plots orthogonal basis functions and parameters to illustrate their properties. Panel A shows the shapes of the orthonormalized basis vectors stored in the rows of \mathbf{B}_0 . As the number of pure exponential basis curves in \mathbf{B} increases, the shapes of the orthogonalized basis curves in \mathbf{B}_0 approach a limit for finely-spaced basis curves. This limiting shape is already approximated very well using a relatively small numbers of basis curves. Figure 3 Panel C shows that the importance of the orthonormal basis functions decreases rapidly, and the behavior is quite similar regardless of the number of basis functions used.

Figure 3, Panel B shows the corresponding orthonormal vectors that weigh the basis curves to give curves in Panel A. These basis weightings suggest that the different orthogonal basis functions can be ascribed qualitative physical meaning relative to the underlying distribution. Because it has no zero crossings (blue curve), the first rate-domain coefficient integrates intensity from the entire distribution. Because it has one zero crossing (orange curve), the second parameter is related to the difference between contributions of fast and slow contributions. Third and higher parameters reflect contributions from three and more regions. By invoking increasing numbers of orthogonal basis parameters, the underlying distribution of components is described with increasing precision. This is somewhat analogous to the way increasing the number of Fourier basis functions increases the fineness with which frequencies of an oscillating signal can be localized. When Fourier weightings (not shown) are overtly used to weigh the basis decay curves, they lead to time domain weighting functions shown in Figure 3 Panel D. These functions are not simple and have the disadvantage of not being orthogonal along the time axis. They have advantages, however, for regularization as discussed below.

Monte Carlo Exploration of Component Selection Heuristics and Uncertainty in Parameters Estimated using Autoregressive Methods

Monte Carlo (MC) explorations of likely parameter values under different assumptions about the number of definable components have been reported (Prange and Song, 2010) to gain insight into the best ways to evaluate multicomponent exponential decay data. In this report, MC is used to validate and illuminate predictions of ARM+LSQ analysis with component selection heuristics. The approach can also help describe parameter uncertainties. To explore this, a decay curve was created using the starting distribution shown in Figure 1, and Gaussian random noise was added to the level of 0.3% of the first point.

The ARM+LSQ analysis suggests that three components are appropriate for parameterizing these synthesized noisy data. For the case of three components, the starting parameter values in the MC calculations were taken from the ARM+LSQ analysis. Local minima for four- and five-component cases are not available from the methods applied to noisy data because the additional components are oscillatory and unphysical. Initial guesses were therefore taken from analysis of the decay curve prior to adding noise, and these parameters were used as starting values in Monte Carlo analysis of the noisy data. This allows exploration of situations where, by hypothesis or inspiration, values of parameters beyond those available from ARM+LSQ are considered.

Figure 7 displays parameter values and their uncertainties determined using MC explorations compared to the Cramér-Rao lower bounds on the parameters given by ARM+LSQ. The orange curves were generated using Gaussian components having positions and integrated amplitudes coming from the ARM+LSQ fits. Widths of the components are the marginalized standard deviations of the rates from the Cramér-Rao bounds given by the inverse Fisher information matrices. The blue curves are determined

from binning the MC calculations. Amplitudes of the components for each step are added in the bin corresponding to the relaxation rate sampled by the algorithm. For both types of curves, the total amplitude was normalized to unity prior to plotting.

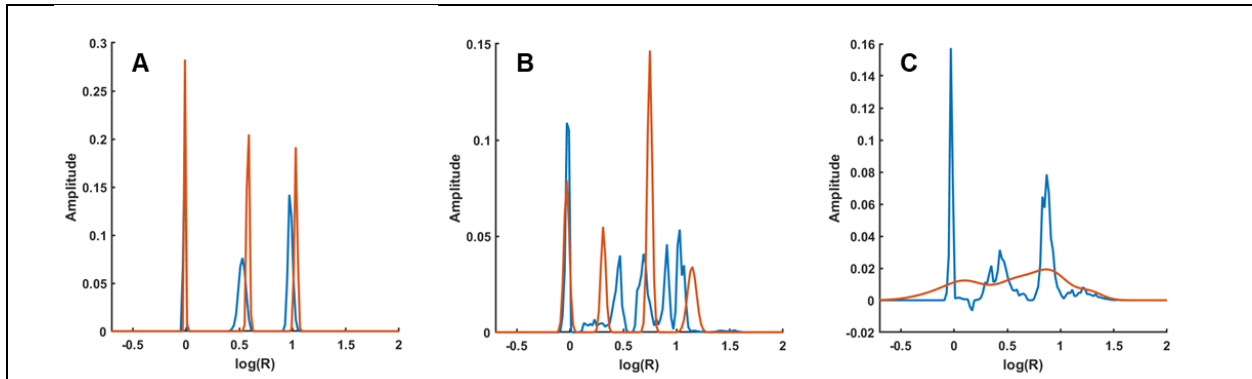


Figure 7. Parameter uncertainties depend on the number of parameters extracted. Three (A), four (B) and five (C) parameters were fitted to noisy decay curves generated using the starting distribution described in Figure 1. Orange lines use Gaussian distributions to convey uncertainties determined from Fisher Information Matrices. Blue lines convey uncertainties as amplitude-weighted distributions of relaxation rates sampled during Monte Carlo explorations.

Using three-components as suggested by ARM+LSQ using component selection heuristics, the components appear mono-modal in both types of display. For this example, the MC method found slightly different maximum likelihood values compared to ARM+LSQ, though that is not always the case. Monte Carlo also finds higher uncertainties compared to the Cramér-Rao lower bound. This is not surprising if the response of χ^2 to parameter variations is slightly non-parabolic. For the four-component case, MC methods find several different parameter combinations that are similarly likely to account for the data. Standard deviations from the Cramér-Rao lower-bound do not accommodate this behavior because of the underlying assumption of a locally parabolic χ^2 surface. For the five-component case, the covariances of the parameters lead to Cramér-Rao lower-bound uncertainties that are very large relative to the differences in the optimized relaxation rates. This is supported by broad but jagged distributions from MC. Inspection of the parameter value trajectories from MC (not shown) indicate that the parameter values do not settle into well-defined ranges, but instead appear to be diffusing in a very large parameter space. This indicates that only a small region of the available likely parameter space is explored by the algorithm in the allotted number of steps (25,000).

Taken together, results of the MC calculations and the Cramér-Rao lower bound suggest that ARM+LSQ, in conjunction with physically defensible component selection heuristics, gives an appropriate parameterization of the data (Fricke et al., 2020). Using fewer components gives inferior fit quality, while using more components opens large regions of parameter space, allowing mutual compensation and giving ambiguous parameter values.

Quasi-Continuous Rendering via Fourier Domain Regularization

A regularization protocol that builds on the parsimonious parameterization methods described above, and that meets reasonable user expectations, was introduced qualitatively in Part 1 of this report. The

approach can be further justified and understood through consideration of the distortion and resolution loss that results from truncated SVD analysis starting with a known, sharp distribution.

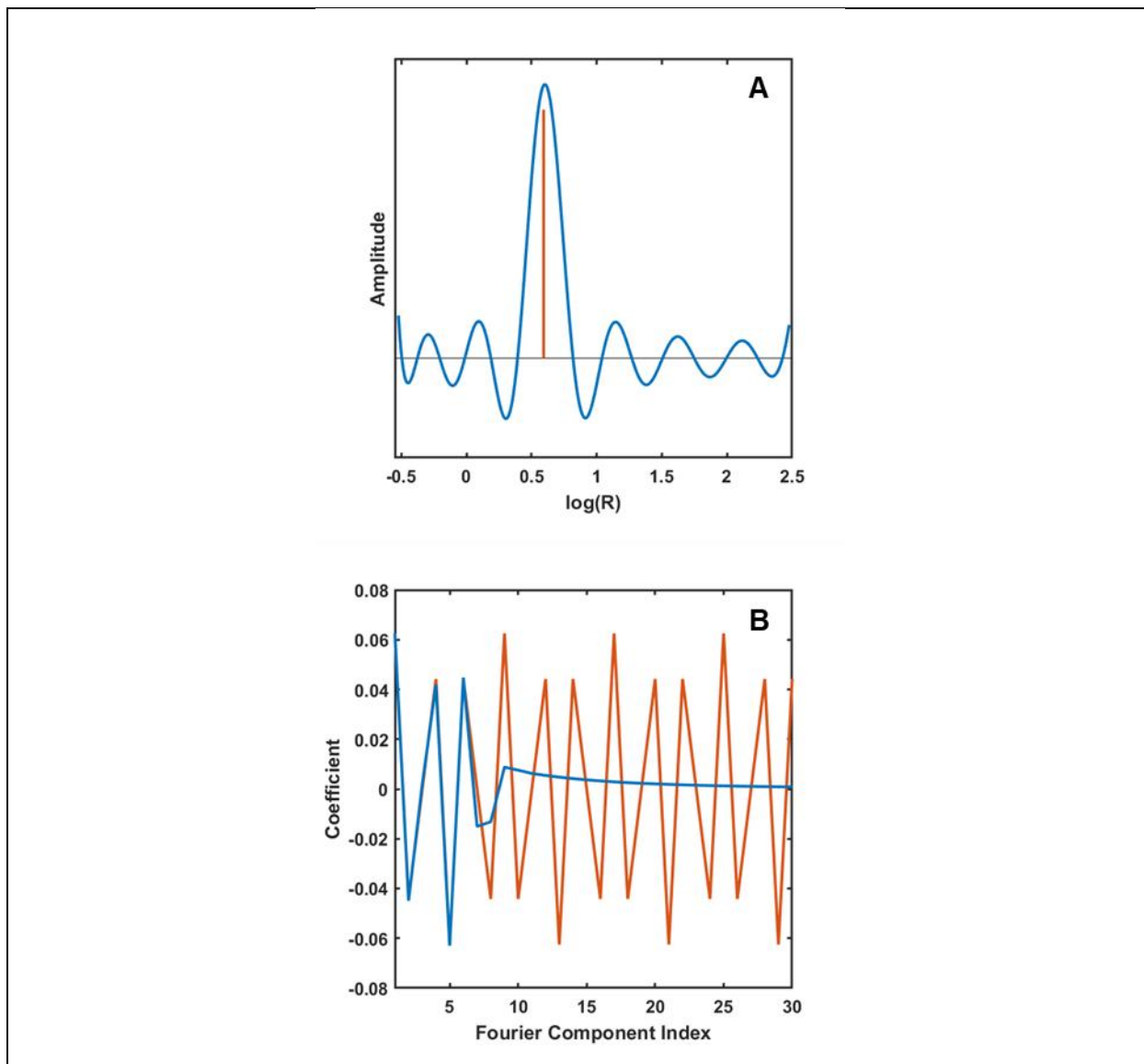


Figure 8. Fourier analysis of a spike distribution and the distribution determined using truncated SVD to analyze the single-component decay curve. (A) Spike distribution (orange) in the spectral domain leads to sinc-like baseline oscillations after fitting the decay curve using t-SVD. (B) Fourier index domain representation of the curves in panel A. (Only real components are shown.) In the Fourier index domain, the fitted distribution is abruptly truncated and distorted for high-index components. This suggests that regularization can be easily achieved in the Fourier index domain by controlling phases, amplitudes, and fall-off of the coefficients.

Figure 8A shows the result of using truncated SVD to fit a single component spike distribution. The spike (orange) was used to generate a noiseless decay curve. Analyzing that decay curve using a geometrically incremented (256 components) amplitude-only basis set in conjunction with the truncated ($k=15$) SVD algorithm gives the result plotted in blue. The result appears much like the convolution of the starting

spike with a *sinc* ($\sin(x)/x$) function. Since *sinc*-like distortions are frequently caused by truncation of signals in a Fourier domain, this suggests that truncated SVD processing has roughly the same effect as calculating the inverse Fourier transform of the input spike, truncating that representation sharply at a cutoff, and then performing the forward Fourier transform to regenerate the original spectral domain.

In the following, domains shown in Figure 8 will be denoted as the spectral domain (Panel A), and the Fourier index domain (Panel B). In the Fourier index domain, inspection of the spectral vector from truncated SVD compared to that of the spike confirms that truncation is a good approximation to the net effect of the analysis protocol. Besides near truncation, higher index but small components are found in the Fourier index representation. In NMR spectroscopy, artifacts like these are readily repaired using known techniques. This suggests that it would be simpler and more physically meaningful if multicomponent exponential relaxation data are fitted and/or regularized in the Fourier index domain.

The Fourier regularization approach is formally described starting with Equation (1.6), and replacing \mathbf{Q} and its inverse with the unitary matrix \mathbf{F} that performs Fourier transformation, and its inverse (Hermitian transpose) \mathbf{F}^\dagger , giving

$$\mathbf{y} = \mathbf{B}_F \mathbf{s}_F, \quad (2.21)$$

where $\mathbf{B}_F = \mathbf{B}\mathbf{F}$ and $\mathbf{s}_F = \mathbf{F}^\dagger \mathbf{s}$. This rearranges to

$$\mathbf{B}_F^{-1} \mathbf{d} = \mathbf{s}_F. \quad (2.22)$$

Working in the Fourier index domain enables simple manipulations to directly suppress oscillations and unphysical features, and to introduce desired features, before converting back to the relaxation rate domain. Conventional regularization methods also suppress oscillations and create smoothness in the spectral domain. This is achieved, for example, by imposing non-negativity. Also, instead of truncating the singular values and vectors, Tikhonov regularization imposes smoothness and suppresses baseline distortions using a regularization constant α and by replacing reciprocals of singular values in Equation (2.11) by

$$\Sigma_i^{-1} \rightarrow \frac{\Sigma_i}{\Sigma_i^2 + \alpha}. \quad (2.23)$$

Regularization in the Fourier index domain is more versatile and easier to design, and the consequences are easier to understand physically. For example, phases of components can be fixed at 0° or 180° to accommodate negative intensities that arise in some measurements. Amplitudes in the Fourier index domain might be extended or damped at high indices using a function that conveys the resolution appropriate to the measurement. Resolution might even be chosen specifically for individual rate components if appropriate prior knowledge is available. The use of Equation (2.23) could also be examined in this context. From the sloppy model perspective, the proposed regularization protocol achieves desired behavior of sloppy parameter combinations in the spectral domain by imposing a smooth decay and eventual full zeroing of components in the Fourier index domain.

Uncertainty Analysis for Two-Dimensional Measurements

The Fisher Information Matrix can be defined for a two-dimensional data set using a double summation

$$I_{\mu,\nu} = \frac{1}{\sigma^2} \sum_{i,j} \frac{\partial Y_{i,j}}{\partial \mu} \frac{\partial Y_{i,j}}{\partial \nu} \quad (2.24)$$

The only difference from the one-dimensional case is that the sum is over all elements of the predicted data matrix rather than over a vector. Note, however, that the indices in Equation (2.24) are μ and ν which refer to variables, not α and β which refer to components. For example μ may refer to an amplitude associated with component α , while ν may refer to the relaxation rate belonging to the same component α . Symbolic expressions for Fisher information matrix elements can in principle be found using computer algebra packages. However, this report explores parameter uncertainties from two-dimensional measurements using Monte Carlo explorations rather than symbolically. Within the assumptions that lead to the usefulness of the Fisher Information Matrix, this is equivalent to exploring its inverse.

Special Case: T1-T2 Measurement

For some circumstances, including the common T1-T2 measurement where equilibrium magnetization and the preparation flip angle must be determined from the data, the spectral matrix components $S_{\alpha,\beta}$ are not linearly independent. This leads immediately to a rank-deficient Fisher information matrix, or to unbounded covariances in its reciprocal. Though SVD analysis can identify interdependent parameters, it is also possible to examine the measurement protocol directly to see the origin of the parameter interdependencies. Doing this gives physical insight and defines a set of parameters that can be determined using ARM+LSQ and explored using Monte-Carlo calculations.

To analyze the T1-T2 measurement, consider a case with two spin systems labeled α and β . Assume that each spin system has unique relaxation rates in the longitudinal relaxation period, $r_{1,\alpha}$ and $r_{1,\beta}$, as well as in the transverse relaxation period $r_{2,\alpha}$ and $r_{2,\beta}$. Magnetization can exchange between the spin systems during the experiment. For the magnetization component that begins on spin system α and that is destined to end on spin system β , denoted $\alpha \rightarrow \beta$, the behavior during the longitudinal relaxation time period as the magnetization decays towards its equilibrium value is described by $A_{\alpha,\beta} (1 - f e^{-r_{1,\alpha} t})$, where the flip angle factor $f = 1 - \cos(\theta)$ is a parameter that varies between 0 and 2, and θ is the net flip angle that depends on the instrument and its adjustments. In the subsequent transverse relaxation time period, further relaxation proceeds towards zero according to $e^{-r_{2,\beta} t}$. The net behavior of magnetization from this subset of spin systems is

$$Y_{\alpha,\beta} = A_{\alpha,\beta} e^{-r_{2,\beta} t} - f A_{\alpha,\beta} e^{-r_{1,\alpha} t} e^{-r_{2,\beta} t} \quad (2.25)$$

We refer to coefficients like $A_{\alpha,\beta}$ as *component matrix amplitudes*. Similar expressions hold for the $\alpha \rightarrow \alpha$, $\beta \rightarrow \beta$ and $\beta \rightarrow \alpha$ pathways.

These expressions can be put into the form of Equation (1.12) if fictitious spin systems having zero relaxation rates are included in each dimension. For the longitudinal relaxation dimension, this zero-relaxation rate component accommodates the non-zero equilibrium magnetization. For the transverse

relaxation dimension, this zero-relaxation rate component can accommodate an offset that is often added to the data to facilitate numerical analysis. In the implementation of the Matrix Pencil Method used here, ambiguities sometimes arise when decay curves cross zero. The difficulty originates within the Matrix Pencil implementation during a singular value decomposition step, where zero crossings in the data sometimes lead to problematic cusps in numerically identified singular vectors. Adding an otherwise inconsequential shift so the entire data set is non-negative eliminates this effect.

For T₁-T₂ data, using μ and ν to enumerate components in the longitudinal and transverse relaxation dimensions, respectively, the elements of the spectral matrix are as follows.

$$S_{0,0} = A_{0,0} \quad (2.26)$$

$$S_{\mu,0} = 0 \quad (2.27)$$

$$S_{0,\nu} = \sum_{\mu} A_{1\mu,2\nu} \quad (2.28)$$

$$S_{\mu,\nu} = -f A_{1\mu,2\nu} \quad (2.29)$$

The adjustable parameters are the component matrix amplitudes $A_{\mu,\nu}$, the flip angle factor f , and the relaxation rates $r_{1\mu}$ and $r_{2,\nu}$. Subscripts 1 and 2 refer to longitudinal and transverse relaxation periods, respectively. The sum of all elements in the zeroth row, excluding the offset, gives the sum of all the component transfer amplitudes. The sum of all the elements of the matrix except the zeroth row and column gives the sum of all the component transfer intensities scaled by the flip angle factor f . The flip angle factor is an independent parameter because it can, for example, be found from the two sums. Analysis of T₁-T₂ data can therefore determine coefficients for all resolvable component transfer amplitudes.

In practice, it is quite convenient to use this format with the matrix pencil method applied to inversion-recovery T₁-T₂ data. The component selection heuristics reliably find components that accommodate the non-decaying basis curves and assign negligible relaxation rates to those components. The resulting analyses give expected offset and zeros to within the uncertainty. As with the one-dimensional case, the parameters yielded by autoregressive methods can be refined using least-squares fitting.

DISCUSSION AND CONCLUSIONS

All the information in a multicomponent exponential relaxation curve having signal-to-noise ratios reasonably achievable in magnetic resonance measurements can be captured in a small number of parameters. Different parameterization protocols may present that information in different forms, each having its own advantages, disadvantages, and sometimes hazards, with respect to the various purposes of the measurement.

A measure of the resolution of noisy multicomponent exponential decay data can be expressed as a geometric interval between relaxation rates used to construct a basis matrix. Using a larger increment than defined by the resolution may not capture all the information in the data, whereas using a smaller increment does not capture any additional information. The resolution of a data set depends on the

sampling rate, the duration of the acquisition, the range of possible decay rates contained in the data, and the noise.

Small, geometrically incremented basis parameterizations can easily be obtained numerically without regularization or concern for problems associated with ill-conditioned matrices. The geometric increment to be used can be determined using the resolution determined from plots such as Figure 2. When a small basis set is orthonormalized using eigenvalue decomposition of the Fisher information matrix or singular value decomposition of the exponential decay basis set, it establishes a unique set of functionals for generating parameters with many uses and advantageous properties. Beyond being statistically independent, these parameters have physical meaning. As their index increases, the distribution of relaxation rates in the source is defined more precisely. Once a rate range and resolution are set, these parameters can be calculated automatically and rapidly for an unlimited number of measurements and samples. The parameters can, in turn, be used for automated data interpretation tasks like classification, automated screening and control, diagnostics, and more. Their versatility and robustness arise because they capture the available information regardless of the underlying distribution of relaxation rates that lead to the measured data, they are commensurate from measurement-to-measurement, and they have useful statistical properties including equal uncertainty ranges and uncorrelated values.

Autoregressive analysis, in conjunction with component selection heuristics and least-squares refinement, here called ARM+LSQ, leads to parameterizations that are optimal in the following sense. Empirical Monte Carlo investigation of simulated data sets shows that using fewer components than determined by the heuristics ignores some information in the data since the parameterizations cannot fit the data within the noise. Using larger numbers of components than suggested by the heuristics leads to multiple local minima, or wide parameter ranges, all giving comparable agreement with the data. Similarly, the Cramér-Rao uncertainty bound shows that the parameters overlap extensively when too many components are considered. Physical meaning can be ascribed to the resulting parameters as follows. A relaxation component represents an underlying distribution of components whose width is no larger than the resolution of the experiment. That distribution may be sharp or broad up to the resolution limit, and it is not possible to distinguish these possibilities from the ARM+LSQ parameterization or any other method without additional knowledge or assumptions.

Recipients of the analyzed data frequently expect pseudo-continuous, spectrum-like renderings. Constructing these renderings requires assigning values to more parameters than can be determined from the data. From the sloppy model perspective, some parameter combinations, classified as stiff, can be well determined. However, most parameter combinations are classified as sloppy, and their values can vary widely, having negligible or no effect on the quality of the fit of the parameters to the data but having very large effects on the appearance of the rendered distributions. Creating meaningful spectrum-like renderings requires assigning values to these sloppy parameter combinations by using regularization methods. Examination of the parameterizations in a Fourier index domain underlies a regularization method that assigns values to these sloppy parameter combinations in a way that generates a useful pseudo-continuous representation. These plots satisfy a reasonable set of user expectations concerning what information is conveyed and how it is to be interpreted from the representations.

Results of ARM+LSQ, Cramér-Rao uncertainties, and pseudo-continuous renderings can be conveyed on a single plot to provide a great deal of useful information in a form easily understood by expert and non-

expert audiences. Uncertainties in parameters from a discrete analysis method such as ARM+LSQ, the resolution of the measurement, and the nature and width of the underlying distribution, are all different concepts. Representations generated using superimposed spike plots, parameter uncertainty distributions, and quasi-continuous renderings convey what is knowable—the parsimonious parameter set from ARM+LSQ, the parameter uncertainties, and the resolution of the measurement—in a familiar format. The precise nature of the underlying distribution cannot be determined and should not be implied without additional considerations.

All the methods described for one-dimensional measurements apply with minor and obvious modifications to two-dimensional measurements.

REFERENCES

- Amari, S., 2016. Information Geometry and its Applications, Applied Mathematical Sciences. Springer Japan.
- Bevington, P.R., 1969. Data Reduction and Error Analysis for the Physical Sciences. McGraw-Hill, New York.
- Borgia, G.C., Brown, R.J.S., Fantazzini, P., 1998. Uniform-penalty inversion of multiexponential decay data. *J. Magn. Reson.* 132, 65–77.
- Butler, J.P., Reeds, J.A., Dawson, S.V., 1981. Estimating Solutions of First Kind Integral Equations with Nonnegative Constraints and Optimal Smoothing. *SIAM J. Numer. Anal.* 18, 381–397. <https://doi.org/10.1137/0718025>
- Callaghan, P.T., Arns, C.H., Galvosas, P., Hunter, M.W., Qiao, Y., Washburn, K.E., 2007. Recent Fourier and Laplace perspectives for multidimensional NMR in porous media. *Magn. Reson. Imaging* 25, 441–444. <https://doi.org/10.1016/j.mri.2007.01.114>
- Caticha, A., 2015. The basics of information geometry, arXiv:1412.5633 [physics.data-an].
- Coe, D., 2009. Fisher Matrices and Confidence Ellipses: A Quick-Start Guide and Software. ArXiv:0906.4123 Astro-Ph.
- English, A.E., Whittall, K.P., Joy, M.L.G., Henkelman, R.M., 1991. Quantitative Two-Dimensional time Correlation Relaxometry. *Magn. Reson. Med.* 22, 425–434.
- Fricke, S.N., Seymour, J.D., Battistel, M.D., Freedberg, D.I., Eads, C.D., Augustine, M.P., 2020. Data processing in NMR relaxometry using the matrix pencil. *J. Magn. Reson.* 313, 106704. <https://doi.org/10.1016/j.jmr.2020.106704>
- Greer, M., Chen, C., Mandal, S., 2018. Automated classification of food products using 2D low-field NMR. *J. Magn. Reson.* 294, 44–58. <https://doi.org/10.1016/j.jmr.2018.06.011>
- Hua, Y., Sarkar, T.K., 1990. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Acoust. Speech Signal Process.* 38, 814–824.
- Istratov, A.A., Vyvenko, O.F., 1999. Exponential analysis in physical phenomena. *Rev. Sci. Instrum.* 70, 1233–1257.
- Koehl, P., 1999. Linear prediction spectral analysis of NMR data. *Prog. Nucl. Magn. Reson. Spectrosc.* 34, 257–299.
- Lee, J.H., Labadie, C., Springer Jr, C.S., Harbison, G.S., 1993. Two-dimensional inverse Laplace transform NMR: altered relaxation times allow detection of exchange correlation. *J. Am. Chem. Soc.* 115, 7761–7764.
- Machta, B.B., Chachra, R., Transtrum, M.K., Sethna, J.P., 2013. Parameter Space Compression Underlies Emergent Theories and Predictive Models. *Science* 342, 604–607.
- Mailhot, S.E., Zong, F., Maneval, J.E., June, R.K., Galvosas, P., Seymour, J.D., 2018. Quantifying NMR relaxation correlation and exchange in articular cartilage with time domain analysis. *J. Magn. Reson.* 287, 82–90. <https://doi.org/10.1016/j.jmr.2017.12.014>
- Mandelstam, V.A., 2001. FDM: the filter diagonalization method for data processing in NMR experiments. *Prog. Nucl. Magn. Reson. Spectrosc.* 38, 159–196.
- Parker, R.L., Song, Y.-Q., 2005. Assigning uncertainties in the inversion of NMR relaxation data. *J. Magn. Reson.* 174, 314–324. <https://doi.org/10.1016/j.jmr.2005.03.002>
- Petrov, O.V., Stapf, S., 2017. Parameterization of NMR relaxation curves in terms of logarithmic moments of the relaxation time distribution. *J. Magn. Reson.* 279, 29–38. <https://doi.org/10.1016/j.jmr.2017.04.009>
- Prange, M., Song, Y.-Q., 2010. Understanding NMR spectral uncertainty. *J. Magn. Reson.* 204, 118–123. <https://doi.org/10.1016/j.jmr.2010.02.010>

- Provencher, S.W., 1982. A constrained regularization method for inverting data represented by linear algebraic or integral equations. *Comput. Phys. Commun.* 27, 213–227.
- Provencher, S.W., 1976. An eigenfunction expansion method for the analysis of exponential decay curves. *J. Chem. Phys.* 64, 2772–2777.
- Reci, A., Sederman, A.J., Gladden, L.F., 2017. Retaining both discrete and smooth features in 1D and 2D NMR relaxation and diffusion experiments. *J. Magn. Reson.* 284, 39–47. <https://doi.org/10.1016/j.jmr.2017.08.017>
- Sarkar, T.K., Pereira, O., 1995. Using the matrix pencil method to estimate the parameters of a sum of complex exponentials. *IEEE Antennas Propag. Mag.* 37, 48–55.
- Silletta, E.V., Franzoni, M.B., Monti, G.A., Acosta, R.H., 2018. Probing numerical Laplace inversion methods for two and three-site molecular exchange between interconnected pore structures. *J. Magn. Reson.* 286, 82–90. <https://doi.org/10.1016/j.jmr.2017.11.016>
- Sivia, D.S., Skilling, J., 2006. *Data Analysis A Bayesian Tutorial*, Second. ed. Oxford University Press.
- Song, Y.-Q., Venkataramanan, L., Hürlimann, M.D., Flaum, M., Frulla, P., Straley, C., 2002. T₁–T₂ Correlation Spectra Obtained Using a Fast Two-Dimensional Laplace Inversion. *J. Magn. Reson.* 154, 261–268. <https://doi.org/10.1006/jmre.2001.2474>
- Su, G., Zhou, X., Wang, L., Wang, X., Yang, P., Nie, S., Zhang, Y., 2019. Improved Butler–Reeds–Dawson Algorithm for the Inversion of Two-Dimensional NMR Relaxometry Data. *Math. Probl. Eng.* 2019, 1–11. <https://doi.org/10.1155/2019/2102343>
- Transtrum, M.K., Machta, B.B., Brown, K.S., Daniels, B.C., Myers, C.R., Sethna, J.P., 2015. Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *J. Chem. Phys.* 143, 010901. <https://doi.org/10.1063/1.4923066>
- Van Landeghem, M., Haber, A., D’espinoze De Lacaillerie, J.-B., Blümich, B., 2010. Analysis of multisite 2D relaxation exchange NMR. *Concepts Magn. Reson. Part A* 36A, 153–169. <https://doi.org/10.1002/cmr.a.20157>
- Whittall, K.P., 1996. ANALYSIS OF NMR RELAXATION DATA. *Data Handl. Sci. Technol.* 18, 44–67.
- Windig, W., Antalek, B., 1997. Direct exponential curve resolution algorithm (DECRA): a novel application of the generalized rank annihilation method for a single spectral mixture data set with exponentially decaying contribution profiles. *Chemom. Intell. Lab. Syst.* 37, 241–254.