Bayesian Analysis Users Guide
Release 4.00, Manual Version 1

G. Larry Bretthorst
Biomedical MR Laboratory
Washington University School Of Medicine,
Campus Box 8227
Room 2313, East Bldg.,
4525 Scott Ave.
St. Louis MO 63110
http://bayes.wustl.edu
Email: larry@bayes.wustl.edu

October 21, 2016

Contents

List of Figures

List of Tables

Chapter 11

Find Resonances

There are two frequency finding programs in the Bayesian Analysis Software: Bayes Analyze, Chapter 8, and Bayes Find Resonance. Bayes Analyze is a searching algorithm that uses the residuals from the current fit, to determine if there is evidence in the data for and additional resonance. While this procedure is implemented using Bayesian probability theory, it is still an approximation to computing the full Bayesian posterior probability for the number of resonances. We implemented Bayes Analyze in this way, so that it would be very fast. However, under some conditions, Bayes Analyze will miss resonances when they are either very close together or the signal-to-noise of the resonance is vary low. To solve these problems, we implemented a frequency finding program that uses Markov chain Monte Carlo to compute the posterior probability for the number of resonances. The interface to the find resonance package is shown in Fig. 11.1 To use this package, you must do the following:

Select the Bayes Find Resonances package from the Package menu.

Load the Fid data that is to be analyzed. If the Fid is arrayed, select the trace that is to be analyzed. The trace analyzed is the currently displayed trace. At the present time only a single Fid is processed by Bayes Find Resonances package.

Select the phase model, the choices are correlated, uncorrelated and automatic.

- If the phase model is "Common," the all resonances have the same phase.
- If the phase model is "Independent," the all resonances have a different phase.
- If the phase model is "Automatic," then the Bayes Find Resonances package computes the posterior probability for the phase of each resonance.

Check the "Constant" box if the data contains an offset.

Set the first and last Fids that are to be analyzed. Note that these Fids are analyzed separately, not jointly, so you will get an analysis for each selected Fid.

Set the maximum number of resonances that can be included in a model.

Select the server that is to process the analysis.

Figure 11.1: The Find Resonances Interface With The Ethyl Ether Spectrum

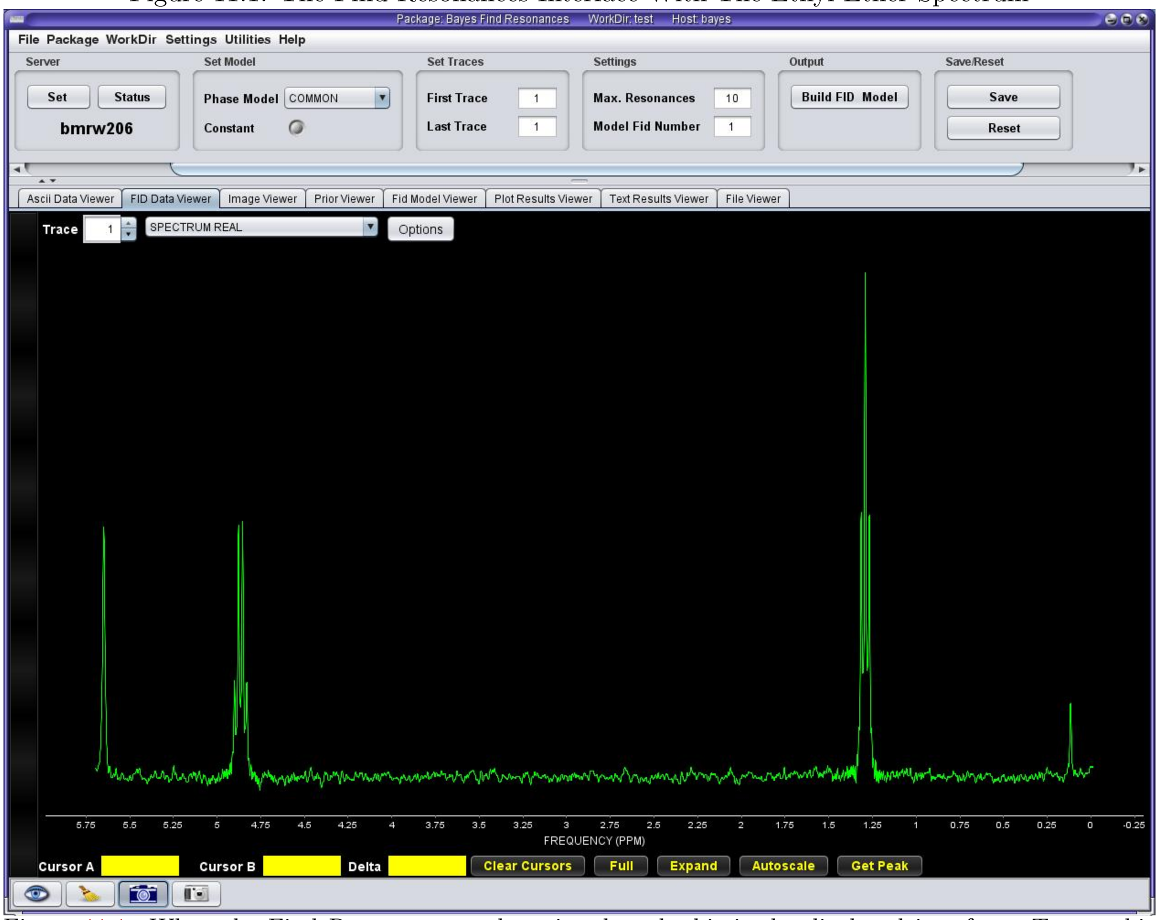

Figure 11.1: When the Find Resonances package is selected, this is the displayed interface. To use this package, load the Fid you wish to analyze. The spectrum of this Fid will be displayed in the Fid Data Viewer. Select the Fid you wish to analyze and display that Fid. At the present time only a single Fid may be processed at one time. Set the various optional feature of the model you wish to use and run the analysis. When the analysis finishes use the "Build Model" button to select and build a model of the Fid. The Fid Model Viewer can then be used to view this model.

Check the status of the selected server to determine if the server is busy, change to another server if the selected server is busy.

Run the analysis on the selected server by activating the "Run" button.

Get the results of the analysis by activating the "Get Job" button. If the analysis is running, this button will return the Accepted report containing the status of the current run. Otherwise, it will fetch and display the results from the current analysis.

Unlike Bayes Analyze which outputs parameter estimates computed using the values that maximized the joint posterior probability for the parameters, Bayes Find Resonances outputs mean and standard deviation parameter estimates computed from all high probability models. That is to say if the probability for the number of resonances was 50% for 9 and 50% for 10 resonances, then there will be mean and standard deviation parameter estimates for the frequencies and decay rate constants from both of these models. Additionally, Bayes Find Resonances runs multiple Fids one right after the other. Consequently, when a Fid model of the time domain data is generated, you must specify both the resonance model and the number of the Fid to model. The Fid number to model is indicated using the "Model Fid Number" entry box. If there are multiples high probability resonance models, clicking on the "Build FID Model" button will show you a list of these models and you can select which one you wish to use in generating a time domain Fid model.

11.1 The Bayesian Calculations

The first step in all Bayesian calculations is to define the problem. Here, the problem is essentially a parameter estimation calculations where one is estimating the frequencies, amplitudes, decay rate constants and phases of multiple exponentially decaying sinusoidal. The model will designated as $\mathbf{M(t_i)}$ where complex quantities will be in bold. Then symbolically the model which relates the complex data, model and noise is given by:

$$\mathbf{d_i} = \mathbf{M(t_i)} + \mathbf{n_i} \qquad (i \in \{1, \ldots, N\}), \tag{11.1}$$

where N is the total number of complex data values, $\mathbf{d_i}$ is a complex data values sampled at time t_i, and $\mathbf{n_i}$ is a complex noise value at time t_i. The complex model $\mathbf{M(t_i)}$ is given by:

$$\mathbf{M(t_i)} = [\mathbf{F}\delta(t_i) + \mathbf{C}]\,\delta(\nu) + \sum_{j=1}^{m} A_j \exp\left\{2\pi \mathrm{i} f_j(t_i - t_0) - \alpha_j t_i + \mathrm{i}\phi\delta(\xi_j) + \mathrm{i}\phi_j[1 - \delta(\xi_j)]\right\} \tag{11.2}$$

where $\mathbf{F}$ is a model of the first data value, the function $\delta(\cdot)$ is defined below, $\mathbf{C}$ is a complex offset, m is the unknown number of sinusoids, A_j is the amplitude of the jth sinusoid, f_j is the frequency of the jth sinusoid, t_0 is a first order phase correction, α_j is the decay rate constant of the jth sinusoid. The quantity "$\phi\delta(\xi_j) + \phi_j[1 - \delta(\xi_j)]$" is the phase of the j sinusoid and is either a common zero order phase, ϕ, or a unique phase specific to the jth sinusoid, ϕ_j. Whether or not the phase is common or unique depends on the value of the indicator function $\delta(\xi_j)$. The indicator function $\delta(\cdot)$ is defined as

$$\delta(\nu) = \begin{cases} 1 & \text{If } \nu = 0 \\ 0 & \text{Otherwise} \end{cases}, \tag{11.3}$$

so, for example, the quantity $\mathbf{F}\delta(t_i)$ is present only when $t_i = 0$. Similarly, the common phase ϕ is present only when $\xi_j = 0$, where ξ_j is a two value binary variable defined as:

$$\xi_j = \begin{cases} 1 & \text{If the } j\text{th sinusoid has a common zero order phase} \\ 0 & \text{Otherwise} \end{cases} . \tag{11.4}$$

Here common zero phase means that several sinusoids share the same zero order phase parameter.

The value of ξ_j is under user control. If the user selects the "Common" phase model, then $\xi_j = 0$ for all sinusoids and all sinusoids share the same common zero order phase parameter. If the user selects the "Independent" phase model, then $\xi_j = 1$ for all sinusoids and all sinusoids have a unique zero order phase parameter ϕ_j. Finally, if the user selects the phase model as "Independent" then the parameters ξ_j are binary variables that are simulated in the Markov chain Monte Carlo simulation, i.e., the Bayes Find Resonances package automatically determines which resonances have common zero order phase and which have a unique zero order phase.

Whether or not the constant models are present is also under user control. If the "Constant" check box is activated, then ν is set equal to zero by the package and $\delta(\nu) = 1$ and in Eq (11.2) the constant models are present. If the constant check box is not active then $\nu = 1$ and no constants are present in Eq (11.2). However, unlike the phase model, the Bayes Find Resonances package does not simulate the binary variable ν, this value is set by the user and the package uses the indicated value.

The Bayes Find Resonances package is a hybrid parameter estimation and model selection package. It is model selection in that it must determine how many resonances are present, and when the phase model is selected as "Independent" it must also estimate the binary parameters, ξ_j. So, the set of parameters estimated by Bayes Find Resonances package when all parameters are active, is:

F is the complex first point model, it contains two constants F_R and F_I, the real and imaginary first point parameters.

C is the complex constant offset model, it contains two constants C_R and C_I, which are the real and imaginary offset parameters.

m is the unknown number of sinusoids in the data.

A is the collection of amplitudes in the m sinusoids, so $A \equiv \{A_1, \ldots, A_m\}$.

f is the collection of frequencies in the m sinusoids, so $f \equiv \{f_1, \ldots, f_m\}$.

α is the collection of decay rate constants in the m sinusoids, so $\alpha \equiv \{\alpha_1, \ldots, \alpha_m\}$.

t_0 is the first order phase in in the m sinusoids.

ξ is the collection of phase model indicators in the m sinusoids, so $\xi \equiv \{\xi_1, \ldots, \xi_m\}$.

ϕ is the common zero order phase in the jth sinusoids when $\delta(\xi_j) = 1$.

ϕ_j is the zero order phase in the jth sinusoids when $\delta(\xi_j) = 0$.

We are going to designate this collection of parameters as Φ and then proceed with the Bayesian calculations.

The Bayesian calculations are for the posterior probability for the number of resonances in the data set. This posterior probability is designated as $P(m|DI)$, where D represents all of the data and I stands for all of the prior information. This posterior probability is computed by application of Bayes' Theorem:

$$P(m|DI) = \frac{P(m|I)P(D|mI)}{P(D|I)} \tag{11.5}$$

where $p(m|I)$ is the prior probability for the number of resonances, $P(D|mI)$ is the marginal direct probability for the data given the model order and $P(D|I)$ is a marginal direct probability for the data given only the prior information I. The direct probability for the data given only the prior information, $P(D|I)$, is a normalization constant and is given by:

$$\begin{aligned} P(D|I) &= \sum_{m=1}^{\text{Max}} P(Dm|I) \\ &= \sum_{m=1}^{\text{Max}} P(m|I)P(D|mI) \end{aligned} \tag{11.6}$$

where the maximum number of resonances is designated as "Max." Comparing, Eq. (11.6) to Eq. (11.5) it is easy to see that $P(D|I)$ is a normalization constant. If we normalize the posterior probability for the number of resonances, $P(m|DI)$, at the end of the calculations, then Eq (11.5) becomes:

$$P(m|DI) \propto P(m|I)P(D|mI). \tag{11.7}$$

The prior probability for the number of resonances, $P(m|I)$, is sufficiently simplified that we could assign it a numerical value. For now we will simply leave it in this symbolic form. However, the direct probability for the data given the number of resonances and the prior information, $P(D|mI)$, is not yet sufficiently simplified so that its value can be assigned. To proceed with the calculation, one introduces the collection of parameters Φ into this probability, the the direct probability for the data, $P(D|mI)$, the posterior probability for the number of resonances becomes

$$P(m|DI) \propto P(m|I) \int P(D\Phi|mI)d\Phi. \tag{11.8}$$

One proceeds by applying the product rule to the right-hand side of this equation:

$$P(m|DI) \propto P(m|I) \int P(\Phi|I)P(D|\Phi mI)d\Phi. \tag{11.9}$$

Factoring the prior probability for all of the parameters, $P(\Phi|I)$ into individual prior probabilities for each parameter, one obtains:

$$\begin{aligned} P(m|DI) &\propto P(m|I)\int P(F_R|I)P(F_I|I)P(C_R|I)P(C_I|I) \\ &\times P(\phi|I)P(t_0|I)P(D|\Phi mI) \\ &\times \left[\prod_{j=1}^{m} P(A_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(f_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(\alpha_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(\xi_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(\phi_j|I)^{\delta(\xi_j)}\right] d\Phi. \end{aligned} \tag{11.10}$$

However, this is the marginal posterior probability for the number of resonances, and is the main output from the Bayes Find Resonances package. But the quantity sampled in the Markov chain Monte Carlo simulation is the joint posterior probability for all of the parameters. While this posterior probability is very similar, its not quite the same. The joint posterior probability for all of the parameters is given by:

$$\begin{aligned} P(m\Phi|DI) &\propto P(m|I)P(F_R|I)P(F_I|I)P(C_R|I)P(C_I|I) \\ &\times P(\phi|I)P(t_0|I)P(D|\Phi mI) \\ &\times \left[\prod_{j=1}^{m} P(A_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(f_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(\alpha_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(\xi_j|I)\right] \\ &\times \left[\prod_{j=1}^{m} P(\phi_j|I)^{\delta(\xi_j)}\right] \end{aligned} \tag{11.11}$$

which is Eq. (11.10) without the integrations. The prior probabilities are assigned as follows:

$P(m|I)$ is assigned as an Exponential prior with $(0 \leq m \leq 50)$ where a model containing 50 resonances is a hard-coded maximum.

$P(F_R|I)$ is assigned as a bounded Gaussian prior whose range is $\pm 6\times$ the amplitude of the first data value and whose standard deviation is three times the magnitude of the first data value.

$P(F_I|I)$ is assigned the same was $P(F_R|I)$ was assigned.

$P(C_R|I)$ is assigned as a bounded Gaussian whose mean is equal to the average of the last 10 real data values and whose upper and lower bounds is 5 times the mean. Finally, the standard deviation of this prior is one fifth of the prior range.

$P(C_I|I)$ is assigned like $P(C_R|I)$ except the means and bounds are taken from the imaginary channel.

$P(\phi|I)$ is assigned a uniform prior probability ranging from zero to 2π.

$P(t_0|I)$ essentially sets a time when the phases of all the sinusoids are the same. The prior probability for this time offset has a mean of zero, meaning that the phases are all expected to be the same at the start of the acquisition. However, we allow this parameter to range over ± 5 dwell or sampling times, i.e., the zero of time could occur the equivalent of 5 data values before the start of the acquisition and up to 5 data value past the start of acquisition. However, the standard deviation of this Gaussian prior probability is only 0.3 data values. Indicating, that while T_0 is allowed to have a large range, in fact it is strongly suspected that its value is near zero.

$P(A_j|I)$ is assigned a bounded Gaussian prior probability for one of the amplitudes. Its mean is zero, it's upper and lower bounds are 3 times the average magnitude of the first 10 complex data values and its standard deviation is two times that average. This prior is used for the prior probability for each resonance amplitude in the model.

$P(f_j|I)$ is assigned a uniform prior probability ranging over the entire sweep width of the data. This prior is used for the prior probability for each resonance frequency in the model.

$P(\alpha_j|I)$ is assigned a positive prior probability whose low is 0.001 in dimensionless units. The peak value is set to the current value of the "lb" parameter. The maximum value is set so that if the signal were decaying at this maximum, it will go through roughly 3 e-foldings in the first 9 data values. So the maximum value is strongly dependent on the sampling rate.

$P(\xi_j|I)$ is assigned a discrete uniform prior probability having two possible values, zero or one with zero indicating an independent phase and one indicating a common phase.

$P(\phi_j|I)$ is assigned a uniform prior probability ranging from zero to 2π. In the above equations, 11.10 and 11.11, this probability is written as $P(\phi_j|I)^{\delta(\xi_j)}$, which is just a notational mechanism to indicate that the prior is either present or not. When $\delta(\xi_j) = 1$, $\xi_j = 0$, the phase model is independent and the prior is present. Similarly, when $\xi_j = 1$ the phase model is common and this prior is not present because the prior is raised to the zero power.

$P(D|\Phi mI)$ is the direct probability for the data given all of the parameters and the prior information and is assigned using a Gaussian prior probability for the noise. This probability is often called a likelihood or likelihood function.

The Markov chain simulation that implements this calculation targets the joint posterior probability for all of the parameters in the model, Eq. (11.11). It then uses Monte Carlo integration to obtain samples from the marginal posterior probability for each parameter appearing in the model. The samples from the marginal posterior probability for each parameter are then used to generate mean and standard deviation estimates of each parameter appearing in the model. Additionally, the samples are used to generate histograms. These histograms are crude estimates of the posterior probability for each parameter. In addition to outputting the histograms, the samples are also output and these samples can be used to generate Maximum Entropy histograms of the samples, see

Section (**??**). Finally, these samples are used to computes the posterior probability for the number of resonances, Eq. (11.10).

If there are multiple high probability models in this posterior, i.e., the posterior probability for the number of resonances, $P(m|DI)$, has significant weight on several values of m, then samples are drawn for each high probability model and these samples are used to generate histograms and parameter estimates given each high probability model. Consequently, then the algorithm finishes it is possible to have not one set of outputs but several. These outputs are viewed using the standard widgets. Additionally, it is possible to generate time domain Fid models for each high probability model.

11.2 Outputs From The Bayes Find Resonances Package

The Text outputs files from the Find Resonance packages consist of: "Bayes.prob.model," "BayesFind-Res.mcmc.values," "Bayes.params," "Console.log," "Bayes.accepted" and a "Bayes.Condensed.File." These output files can be viewed using the Text Viewer or they can be viewed using File Viewer by navigating to the current working directory and then selecting the files. The format of the mcmc.values report is discussed in Appendix D and the other reports are discussed in Chapter 3. Additionally, the "Plot Results Viewer" can be used to view the output probability density functions. In addition to the standard data, model and residual plots there are probability density functions for the frequencies and decay rate constants, and the amplitudes for each resonance for each high probability model.

Finally, for each high probability model, there is an output "bayes.params.nnnn" and "bayes.model.nnnn" file where this file has exactly the same format as the Bayes Analyze Model file, see Chapter 8.5.7. These files are used in conjunction with the "Build FID Model" button to generate a time domain model of the input Fid data. When this button is activated the Fid and the selected bayes.model.nnnn file are sent to the server and the Bayes Model program, see Chapter 8.1 for a description of this program, is run. During this time the interface waits for the Bayes Model program to finish. When the interface detects that the model has been built, the interface fetches the model from the server, Fourier transforms the model and then displays the model using the Fid Model Viewer.

Bibliography

[1] Rev. Thomas Bayes (1763), "An Essay Toward Solving a Problem in the Doctrine of Chances," *Philos. Trans. R. Soc. London*, **53,** pp. 370-418; reprinted in *Biometrika*, **45**, pp. 293-315 (1958), and *Facsimiles of Two Papers by Bayes,* with commentary by W. Edwards Deming, New York, Hafner, 1963.

[2] G. Larry Bretthorst (1988), "Bayesian Spectrum Analysis and Parameter Estimation," in *Lecture Notes in Statistics,* **48**, J. Berger, S. Fienberg, J. Gani, K. Krickenberg, and B. Singer (eds), Springer-Verlag, New York, New York.

[3] G. Larry Bretthorst (1990), "An Introduction to Parameter Estimation Using Bayesian Probability Theory," in *Maximum Entropy and Bayesian Methods,* Dartmouth College 1989, P. Fougère *ed.,* pp. 53-79, Kluwer Academic Publishers, Dordrecht the Netherlands.

[4] G. Larry Bretthorst (1990), "Bayesian Analysis I. Parameter Estimation Using Quadrature NMR Models" *J. Magn. Reson.*, **88**, pp. 533-551.

[5] G. Larry Bretthorst (1990), "Bayesian Analysis II. Signal Detection And Model Selection" *J. Magn. Reson.*, **88**, pp. 552-570.

[6] G. Larry Bretthorst (1990), "Bayesian Analysis III. Examples Relevant to NMR" *J. Magn. Reson.*, **88**, pp. 571-595.

[7] G. Larry Bretthorst (1991), "Bayesian Analysis. IV. Noise and Computing Time Considerations," *J. Magn. Reson.*, **93**, pp. 369-394.

[8] G. Larry Bretthorst (1992), "Bayesian Analysis. V. Amplitude Estimation for Multiple Well-Separated Sinusoids," *J. Magn. Reson.*, **98**, pp. 501-523.

[9] G. Larry Bretthorst (1992), "Estimating The Ratio Of Two Amplitudes In Nuclear Magnetic Resonance Data," in *Maximum Entropy and Bayesian Methods,* C. R. Smith *et al.* (eds.), pp. 67-77, Kluwer Academic Publishers, the Netherlands.

[10] G. Larry Bretthorst (1993), "On The Difference In Means," in Physics & Probability Essays in honor of Edwin T. Jaynes, W. T. Grandy and P. W. Milonni (eds.), pp. 177-194, Cambridge University Press, England.

[11] G. Larry Bretthorst (1996), "An Introduction To Model Selection Using Bayesian Probability Theory," in *Maximum Entropy and Bayesian Methods*, G. R. Heidbreder, ed., pp. 1-42, Kluwer Academic Publishers, Printed in the Netherlands.

[12] G. Larry Bretthorst (1999), "The Near-Irrelevance of Sampling Frequency Distributions," in *Maximum Entropy and Bayesian Methods,* W. von der Linden *et al.* (eds.), pp. 21-46, Kluwer Academic Publishers, the Netherlands.

[13] G. Larry Bretthorst (2001), "Nonuniform Sampling: Bandwidth and Aliasing," in *Maximum Entropy and Bayesian Methods in Science and Engineering,* Joshua Rychert, Gary Erickson and C. Ray Smith *eds.*, pp. 1-28, American Institute of Physics, USA.

[14] G. Larry Bretthorst, Christopher D. Kroenke, and Jeffrey J. Neil (2004), "Characterizing Water Diffusion In Fixed Baboon Brain," in Bayesian Inference And Maximum Entropy Methods In Science And Engineering, Rainer Fischer, Roland Preuss and Udo von Toussaint *eds.*, AIP conference Proceedings, **735**, pp. 3-15.

[15] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.H. Ackerman (2005), "Exponential parameter estimation (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 55-63.

[16] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J. H. Ackerman (2005), "Exponential model selection (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 64-72.

[17] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.H. Ackerman (2005), "How accurately can parameters from exponential models be estimated? A Bayesian view," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 73-83.

[18] G. Larry Bretthorst, W. C. Hutton, J. R. Garbow, and Joseph J. H. Ackerman (2008), "High Dynamic Range MRS Time-Domain Signal Analysis," *Magn. Reson. in Med.*, **62**, pp. 1026-1035.

[19] V. Chandramouli, K. Ekberg, W. C. Schumann, S. C. Kalhan, J. Wahren, and B. R. Landau (1997), "Quantifying gluconeogenesis during fasting," *American Journal of Physiology*, **273**, pp. H1209-H1215.

[20] R. T. Cox (1961), "The Algebra of Probable Inference," Johns Hopkins Univ. Press, Baltimore.

[21] André d'Avignon, G. Larry Bretthorst, Marlyn Emerson Holtzer, and Alfred Holtzer (1998), "Site-Specific Thermodynamics and Kinetics of a Coiled-Coil Transition by Spin Inversion Transfer NMR," *Biophysical Journal*, **74**, pp. 3190-3197.

[22] André d'Avignon, G. Larry Bretthorst, Marlyn Emerson Holtzer, and Alfred Holtzer (1999), "Thermodynamics and Kinetics of a Folded-Folded Transition at Valine-9 of a GCN4-Like Leucine Zipper," *Biophysical Journal*, **76**, pp. 2752-2759.

[23] David Freedman, and Persi Diaconis (1981), "On the histogram as a density estimator: L_2 theory," *Zeitschrift f¨r Wahrscheinlichkeitstheorie und verwandte Gebiete*, **57**, 4, pp. 453-476.

[24] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (1996), "Markov Chain Monte Carlo in Practice," Chapman & Hall, London.

[25] Paul M. Goggans, and Ying Chi (2004), "Using Thermodynamic Integration to Calculate the Posterior Probability in Bayesian Model Selection Problems," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 23rd International Workshop*, **707**, pp. 59-66.

[26] Marlyn Emerson Holtzer, G. Larry Bretthorst, D. André d'Avignon, Ruth Hogue Angelette, Lisa Mints, and Alfred Holtzer (2001), "Temperature Dependence of the Folding and Unfolding Kinetics of the GCN4 Leucine Lipper via 13C alpha-NMR," *Biophysical Journal*, **80**, pp. 939-951.

[27] E. T. Jaynes (1968), "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, pp. 227-241; reprinted in [30].

[28] E. T. Jaynes (1978), "Where Do We Stand On Maximum Entropy?" in *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus *Eds.*, pp. 15-118, Cambridge: MIT Press, Reprinted in [30].

[29] E. T. Jaynes (1980), "Marginalization and Prior Probabilities," in *Bayesian Analysis in Econometrics and Statistics*, A. Zellner *ed.,* North-Holland Publishing Company, Amsterdam; reprinted in [30].

[30] E. T. Jaynes (1983), "Papers on Probability, Statistics and Statistical Physics," a reprint collection, D. Reidel, Dordrecht the Netherlands; second edition Kluwer Academic Publishers, Dordrecht the Netherlands, 1989.

[31] E. T. Jaynes (1957), "How Does the Brain do Plausible Reasoning?" unpublished Stanford University Microwave Laboratory Report No. 421; reprinted in *Maximum-Entropy and Bayesian Methods in Science and Engineering* **1,** pp. 1-24, G. J. Erickson and C. R. Smith *Eds.,* 1988.

[32] E. T. Jaynes (2003), "Probability Theory—The Logic of Science," edited by G. Larry Bretthorst, Cambridge University Press, Cambridge UK.

[33] Sir Harold Jeffreys (1939), "Theory of Probability," Oxford Univ. Press, London; Later editions, 1948, 1961.

[34] John G. Jones, Michael A. Solomon, Suzanne M. Cole, A. Dean Sherry, and Craig R. Malloy (2001) "An integrated ^{2}H and ^{13}C NMR study of gluconeogenesis and TCA cycle flux in humans," *American Journal of Physiology, Endocrinology, and Metabolism,* **281,** pp. H848-H856.

[35] John Kotyk, N. G. Hoffman, W. C. Hutton, G. Larry Bretthorst, and J. J. H. Ackerman (1992), "Comparison of Fourier and Bayesian Analysis of NMR Signals. I. Well-Separated Resonances (The Single-Frequency Case)," *J. Magn. Reson.*, **98**, pp. 483–500.

[36] Pierre Simon Laplace (1814), "A Philosophical Essay on Probabilities," John Wiley & Sons, London, Chapman & Hall, Limited 1902. Translated from the 6th edition by F. W. Truscott and F. L. Emory.

[37] N. Lartillot, and H. Philippe (2006), "Computing Bayes Factors Using Thermodynamic Integration," *Systematic Biology*, **55** (2), pp. 195-207.

[38] D. Le Bihan, and E. Breton (1985), "Imagerie de diffusion in-vivo par rsonance," Comptes rendus de l'Acadmie des Sciences (Paris), **301** (15), pp. 1109-1112.

[39] N. R. Lomb (1976), "Least-Squares Frequency Analysis of Unevenly Spaced Data," *Astrophysical and Space Science*, **39**, pp. 447-462.

[40] T. J. Loredo (1990), "From Laplace To SN 1987A: Bayesian Inference In Astrophysics," in *Maximum Entropy and Bayesian Methods,* P. F. Fougere (ed), Kluwer Academic Publishers, Dordrecht, The Netherlands.

[41] Craig R. Malloy, A. Dean Sherry, and Mark Jeffrey (1988), "Evaluation of Carbon Flux and Substrate Selection through Alternate Pathways Involving the Citric Acid Cycle of the Heart by 13C NMR Spectroscopy," *Journal of Biological Chemistry*, **263** (15), pp. 6964-6971.

[42] Craig R. Malloy, Dean Sherry, and Mark Jeffrey (1990), "Analysis of tricarboxylic acid cycle of the heart using ^{13}C isotope isomers," *American Journal of Physiology*, **259**, pp. H987-H995.

[43] Lawrence R. Mead and Nikos Papanicolaou, "Maximum entropy in the problem of moments," *J. Math. Phys.* **25**, 2404–2417 (1984).

[44] K. Merboldt, Wolfgang Hanicke, and Jens Frahm (1969), "Self-diffusion NMR imaging using stimulated echoes," Journal of Magnetic Resonance, **64** (3), pp. 479-486.

[45] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller (1953), "Equation of State Calculations by Fast Computing Machines," Journal of Chemical Physics. The previous link is to the Americain Institute of Physics and if you do not have access to Science Sitations you many not be able to retrieve this paper.

[46] Radford M. Neal (1993), "Probabilistic Inference Using Markov Chain Monte Carlo Methods," technical report CRG-TR-93-1, Dept. of Computer Science, University of Toronto.

[47] Jeffrey J. Neil, and G. Larry Bretthorst (1993), "On the Use of Bayesian Probability Theory for Analysis of Exponential Decay Data: An Example Taken from Intravoxel Incoherent Motion Experiments," *Magn. Reson. in Med.*, **29**, pp. 642–647.

[48] H. Nyquist (1924), "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, **3**, pp. 324-346.

[49] H. Nyquist (1928), "Certain Topics in Telegraph Transmission Theory," *Transactions AIEE*, **3**, pp. 617-644.

[50] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery (1992), "Numerical Recipes The Art of Scientific Computing Second Edition," Cambridge University Press, Cambridge UK.

[51] Emanuel Parzen (1962), "On Estimation of a Probability Density Function and Mode," *Annals of Mathematical Statistics* **33**, 1065–1076

[52] Karl Pearson (1895), "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material," *Phil. Trans. R. Soc. A* **186**, 343–326.

[53] Murray Rosenblatt, "Remarks on Some Nonparametric Estimates of a Density Function," *Annals of Mathematical Statistics* **27**, 832–837 (1956).

[54] Jeffery D. Scargle (1981), "Studies in Astronomical Time Series Analysis I. Random Process In The Time Domain," *Astrophysical Journal Supplement Series*, **45**, pp. 1-71.

[55] Jeffery D. Scargle (1982), "Studies in Astronomical Time Series Analysis II. Statistical Aspects of Spectral Analysis of Unevenly Sampled Data," *Astrophysical Journal*, **263**, pp. 835-853.

[56] Jeffery D. Scargle (1989), "Studies in Astronomical Time Series Analysis. III. Fourier Transforms, Autocorrelation Functions, and Cross-correlation Functions of Unevenly Spaced Data," *Astrophysical Journal*, **343**, pp. 874-887.

[57] Arthur Schuster (1905), "The Periodogram and its Optical Analogy," *Proceedings of the Royal Society of London*, **77**, p. 136-140.

[58] Claude E. Shannon (1948), "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, **27**, pp. 379-423.

[59] John E. Shore, and Rodney W. Johnson (1981), "Properties of cross-entropy minimization," *IEEE Trans. on Information Theory*, **IT-27,** No. 4, pp. 472-482.

[60] John E. Shore and Rodney W. Johnson (1980), "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Trans. on Information Theory*, **IT-26** (1), pp. 26-37.

[61] Devinderjit Sivia, and John Skilling (2006), "Data Analysis: A Bayesian Tutorial," Oxford University Press, USA.

[62] Edward O. Stejskal and Tanner, J. E. (1965), "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient." Journal of Chemical Physics, **42** (1), pp. 288-292.

[63] D. G. Taylor and Bushell, M. C. (1985), "The spatial mapping of translational diffusion coefficients by the NMR imaging technique," Physics in Medicine and Biology, **30** (4), pp. 345-349.

[64] Myron Tribus (1969), "Rational Descriptions, Decisions and Designs," Pergamon Press, Oxford.

[65] P. M. Woodward (1953), "Probability and Information Theory, with Applications to Radar," McGraw-Hill, N. Y. Second edition (1987); R. E. Krieger Pub. Co., Malabar, Florida.

[66] Arnold Zellner (1971), "An Introduction to Bayesian Inference in Econometrics," John Wiley and Sons, New York.