Bayesian Data-Analysis Toolbox
Release 4.23, Manual Version 3

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: gbretthorst@wustl.edu

September 18, 2018

Chapter 9

Big Peak/Little Peak

The Big Peal/Little peak package is designed to analyze a single FID containing a solvent resonance, the big peak, and a few small resonances. The solvent resonance is treated as a nuisance signal and marginalized from the problem; the frequencies, decay rate constants and amplitudes of the small resonances are the primary output from this analysis. The interface to this package is shown in Fig. 9.1. To use this package, you must do the following:

Select the "Big Peak/Little Peak" package from the Package menu.

Load the spectroscopic Fid you wish to analyze.

Display the trace you wish to analyze. Big Peak/Little peak analyzes a single trace and this trace must be displayed in the Fid viewer. Consequently, you must select the trace you wish to analyze by displaying it.

Use a double cursor to bracket the location of the solvent and hit the "Solvent" button. This will result in the location of the solvent being bracketed by a double set of lines on the display. Additionally, the solvent frequency and decay rate constant will be displayed in the parameter list. Clicking on these parameters will cause their associated prior probabilities to be displayed and you can adjust these values if desired.

Use a double cursor to bracket the locations of each metabolite to be included in the model and hit the "Metabolite" button. When the metabolite button is activated the frequency and decay rate constant for this metabolite are added to the list of parameters and as with the solvent parameters clicking on a parameter will display the parameter prior and you can adjust these prior probabilities as needed.

If needed a right click on a parameter can be used to remove a frequency and decay rate pair.

Select the server that is to process the analysis.

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

Figure 9.1: The Big Peak/Little Peak Interface

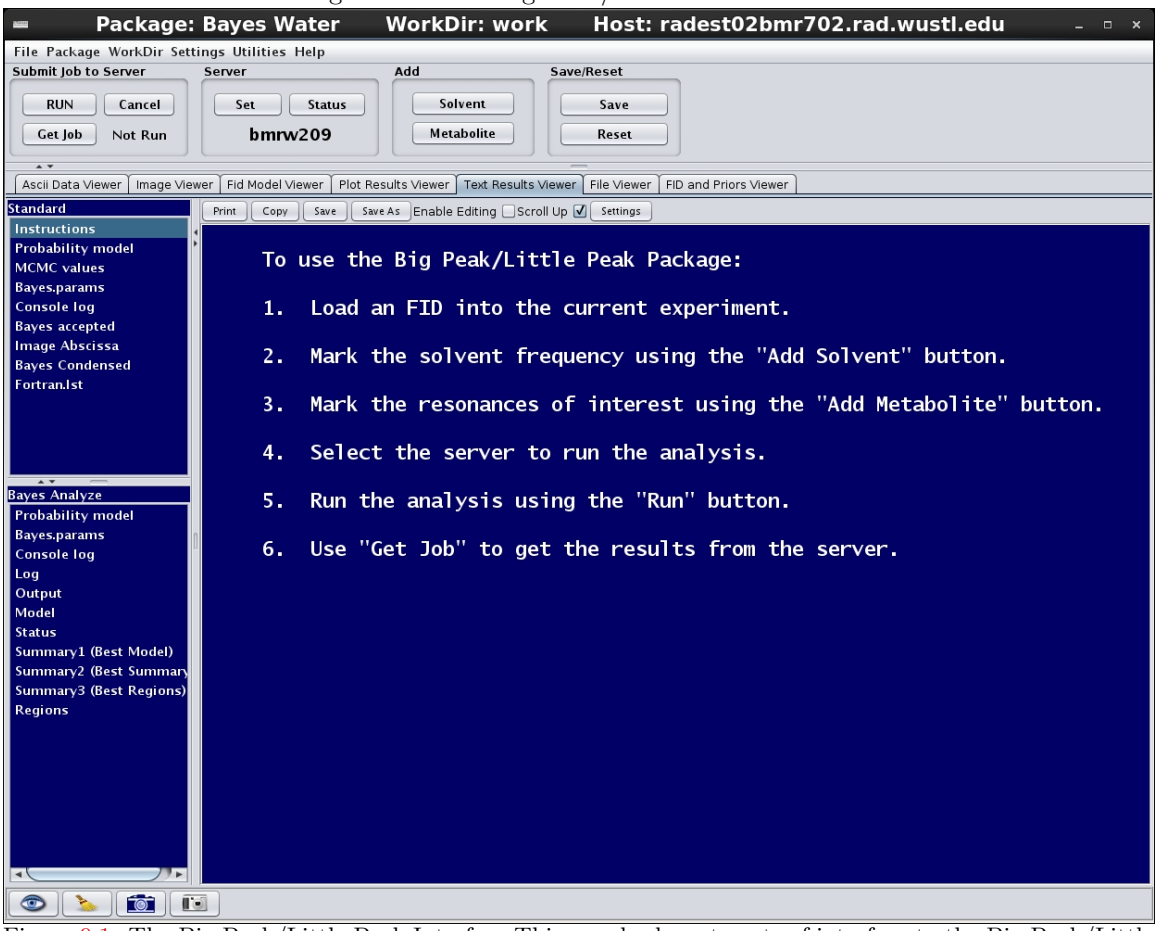

Figure 9.1: The Big Peak/Little Peak Interface This panel relevant parts of interface to the Big Peak/Little Peak package. To use the Big Peak/Little Peak package you must specify the frequency of the solvent, the big peak, and you must mark the locations of all of the small peaks you wish to analyze. The output from the program includes the estimated frequencies, decay rate constants, and amplitudes of these small resonances.

Get the 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.

View the results and the model using the Model Viewer.

9.1 The Bayesian Calculation

The Big Peak/Little Peak package analyzes a single Fid that contains a large solvent resonance or solvent suppression artifact and one or more smaller resonances of interest. In this problem we will assume these small resonances are of relatively low signal-to-noise and may be modeled as exponentially decaying sinusoids:

$$\text{Model of the Interesting Resonances} = \sum_{k=1}^{m} A_k \cos(2\pi f_k[t_i + t_0] + \theta) \exp\{-\alpha_k t_i\} \tag{9.1}$$

where m is the number of interesting resonances, A_k is the amplitude of the kth interesting resonance, f_k is its frequency, α_k is its decay rate constant, t_0 is a time offset and may be thought of as a frequency dependent phase correction, and θ is the constant or zero order phase.

For most purposes this is the model used in the Bayes Analyze package when it is modeling singlets. Unfortunately, the solvent peak cannot be modeled this way because of its dynamic range. Exponentially decaying sinusoidal models work very well for low to moderate signal-to-noise levels. However, when the signal-to-noise levels become larger than a few hundred, this model becomes inadequate. Inadequate in the sense that the difference between the data and the best fit model, the residual, contains a systematic artifact. The discrete Fourier transform of this artifact is often much larger than peaks associated with the resonances of interest. Consequently, programs that fit exponentially decaying sinusoidal to solvent peaks put more and more resonances in the location of the solvent and often never see the resonances of interest. Indeed if you run the Bayes Analyze program on the test data set "Bayes.test.data/BayesWater_test.fid" you will quickly discover that Bayes Analyze cannot really analyze this FID. It places multiple frequencies in the location of the solvent, and completely misses the small resonances on the right of the solvent. In order to fix this problem a better model of the solvent resonances must be used in the analysis.

Because we are using Bayesian probability theory and because we don't care about the solvent resonance we are free to model this resonance in any way that will fit it down to the noise level. The solvent model employed in this calculation is of an exponentially decaying sinusoid having an amplitude and phase that are slowly varying functions of time. If $S(t_i)$ represents the complex solvent signal, then the model of the solvent is given by

$$S(t_i) = A(t_i) \exp\{2\pi \mathrm{i} f_s t_i - \alpha_s t_i - \mathrm{i}\phi(t_i)\} \tag{9.2}$$

where f_s and α_s are the solvent frequency and decay rate constant and $A(t_i)$ and $\phi(t_i)$ are the unknown amplitude and phase modulation. Note that if the solvent frequency and decay rate constant are set to zero, then this model is just a complex trend.

This solvent model is a complex function and before it may be used in a Bayesian calculation it must separated into its real and imaginary parts:

$$\begin{aligned} S(t_i) &= [A_c(t_i)\cos(2\pi f_s t_i) - A_s(t_i)\sin(2\pi f_s t_i)]\exp\{-\alpha_s t_i\} \\ &- \mathrm{i}[A_c(t_i)\sin(2\pi f_s t_i) + A_s(t_i)\cos(2\pi f_s t_i)]\exp\{-\alpha_s t_i\} \end{aligned} \tag{9.3}$$

where we have written the amplitude, $A(t_i)$, and phase, $\phi(t_i)$, as a cosine and sine amplitude, $A_c(t_i)$ and $A_s(t_i)$.

The two functions $A_c(t_i)$ and $A_s(t_i)$ must have a few simple properties if they are to be physically meaningful. First, when the shimming is good or the signal-to-noise is low, the expansion order will go to 1, and the polynomials reduce to constants. Second, the solvent resonance does not usually deviate all that much from an exponentially decaying sinusoid, so we do not expect the functions $A_c(t_i)$ and $A_s(t_i)$ to be rapidly varying. Indeed if these two functions were rapidly varying we would reject their use because they would be able to represent the small resonances of interest. The simplest functions that have the desired properties are polynomial expansions:

$$A_c(t_i) = \sum_{j=1}^{n_c} B_i \mathcal{L}_j(t_i) \qquad \text{and} \qquad A_s(t_i) = \sum_{k=1}^{n_s} C_i \mathcal{L}_k(t_i) \tag{9.4}$$

where n_c and n_s are the number of the polynomials in the cosine and sine expansions respectively. The number of polynomials, n_c and n_s, and the amplitudes, B_i and C_i, are unknowns and must be inferred as part of the Bayesian calculations.

The polynomials $\mathcal{L}_j(t_i)$ will be chosen so that they are orthogonal on the discretely sampled times:

$$\sum_{i=1}^{N} \mathcal{L}_j(t_i)\mathcal{L}_k(t_i) = \delta_{jk} \tag{9.5}$$

where N is the total number of complex data values, and $\delta_{jk} = 0$ if $j \neq k$ and it is equal to one if $j = k$. We make this choice for computational reasons because it helps stabilize the numerical computation.

Having specified the model for the small resonances of interest, Eq. (9.1) and the model for the solvent, Eq. (9.3), we are now in a position to relate the FID data to the model signal. The model for the real data will be written as

$$d_R(t_i) = M_R(t_i) + \text{Noise in the real data} \tag{9.6}$$

and similarly for the imaginary data

$$d_I(t_i) = M_I(t_i) + \text{Noise in the imaginary data.} \tag{9.7}$$

The model of the real part of the FID data, $M_R(t_i)$, will be written as

$$M_R(t_i) = \sum_{j=1}^{n_{\mathcal{A}}} \mathcal{A}_j R_j(t_i) \tag{9.8}$$

and similarly the model for the quadrature, or imaginary, part of the FID will be written as

$$M_I(t_i) = \sum_{j=1}^{n_{\mathcal{A}}} \mathcal{A}_j I_j(t_i) \tag{9.9}$$

where the total number of amplitudes, $n_{\mathcal{A}}$, is given by

$$n_{\mathcal{A}} = m + n_c + n_s. \tag{9.10}$$

In these equations we have relabeled the amplitudes, A_l, B_j and C_k, as $\mathcal{A}_j$:

$$\mathcal{A} \equiv \{A_1, \ldots, A_m, B_1, \ldots, B_{n_c}, C_1, \ldots, C_{n_s}\}. \tag{9.11}$$

The functions $R_j(t_i)$ and $I_j(t_i)$ are defined as

$$R_j(t_i) = \begin{cases} \cos(2\pi f_j[t_i + t_0] + \theta)\exp\{-\alpha_j t_i\} & \text{if } 1 \le j \le m \\ \mathcal{L}_{(j-m)}(t_i)\cos(2\pi f_s t_i)\exp\{-\alpha_s t_i\} & \text{if } m+1 \le j \le m+n_c \\ -\mathcal{L}_{(j-m-n_c)}(t_i)\sin(2\pi f_s t_i)\exp\{-\alpha_s t_i\} & \text{if } m+n_c+1 \le j \le m+n_c+n_s \end{cases} \tag{9.12}$$

and

$$I_j(t_i) = \begin{cases} -\sin(2\pi f_j[t_i + t_0] + \theta)\exp\{-\alpha_j t_i\} & \text{if } 1 \le j \le m \\ -\mathcal{L}_{(j-m)}(t_i)\sin(2\pi f_s t_i)\exp\{-\alpha_s t_i\} & \text{if } m+1 \le j \le m+n_c \\ -\mathcal{L}_{(j-m-n_c)}(t_i)\cos(2\pi f_s t_i)\exp\{-\alpha_s t_i\} & \text{if } m+n_c+1 \le j \le m+n_c+n_s \end{cases} . \tag{9.13}$$

The calculation implemented in the Markov chain Monte Carlo simulation is a combined parameter estimation and model selection calculation. The target distribution of the Monte Carlo simulation is the joint posterior probability for the nonlinear parameters and the expansion orders, $P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s|D_R D_I I)$. This is a marginal posterior probability where the amplitudes, $\mathcal{A} \equiv \{\mathcal{A}_1, \ldots, \mathcal{A}_{n_\mathcal{A}}\}$, and the standard deviation of the noise prior probability, σ, were removed using the rules of probability theory. We could have left these parameters in the joint posterior probability for all of the parameters and then targeted this distribution in the Monte Carlo simulations. However, we choose not to do this because when the order of one of the expansions is changed, the amplitudes that maximize the joint posterior probability for the parameters tend to change rather abruptly. Consequently, sampling them is very difficult; it was easier to remove them using marginalization. By removing them, we simply allowed probability to determine them for us automatically.

The joint posterior probability for the nonlinear parameters is computed from the joint posterior probability for all of the parameters, $P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s\sigma\mathcal{A}|D_R D_I I)$, by application of Bayes' theorem and the sum rule,

$$P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s|D_R D_I I) \propto \int d\mathcal{A}d\sigma P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s\sigma\mathcal{A}|D_R D_I I). \tag{9.14}$$

By repeatedly applying the product rule, the right-hand side of this equation may be factored to obtain

$$\begin{aligned} P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s|D_R D_I I) \quad &\propto \quad P(\theta|I)P(t_0|I)P(f_s|I)P(\alpha_s|I) \\ &\times \quad P(n_c|I)P(n_s|I)\left[\prod_{l=1}^{m} P(f_l|I)P(\alpha_l|I)\right] \\ &\times \quad \int d\mathcal{A}d\sigma\ P(\sigma|I)\left[\prod_{j=1}^{n_\mathcal{A}} P(\mathcal{A}_j|I)\right] \\ &\times \quad P(D_R|f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s\mathcal{A}I) \\ &\times \quad P(D_I|f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s\mathcal{A}I) \end{aligned} \tag{9.15}$$

where $P(\theta|I)$ is the prior probability for the phase, $P(t_0|I)$ is the prior probability for the time offset, $P(f_s|I)$ and $P(\alpha_s|I)$ are the prior probabilities for the solvent frequency and decay rate constant, $P(n_c|I)$ and $P(n_s|I)$ are the prior probability for the cosine and sine expansion orders and $P(f_l|I)$ and $P(\alpha_l|I)$ are the prior probability for the small peak frequencies and decay rate constants. These prior probabilities do not depend on the amplitudes or the standard deviation of the noise; consequently, we have removed them from the integral. However, $P(\sigma|I)$, $P(\mathcal{A}_j|I)$, $P(D_R|f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s \mathcal{A}I)$, and $P(D_I|f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s \mathcal{A}I)$, do participate in the integral.

All of the terms in Eq. (9.15) have been simplified to the point that they may now be assigned. The prior probability for the phase, $P(\theta|I)$, was assigned a uniform prior probability:

$$P(\theta|I) = \begin{cases} \dfrac{1}{360} & \text{If } 0 \leq \theta \leq 360, \\ 0 & \text{Otherwise} \end{cases} . \tag{9.16}$$

The prior probability for the delay time, $P(t_0|I)$, was assigned as a Gaussian:

$$P(t_0|I) = (2\pi\sigma_{t_0}^2)^{-\frac{1}{2}} \exp\left\{-\frac{t_0^2}{2\sigma_{t_0}^2}\right\}, \tag{9.17}$$

where $\sigma_{t_0} = 3\Delta T$ and ΔT is the sampling rate. In words we are saying that at one standard deviation we think we could miss the first three time points, but we think it would be very unlikely to miss the first 10.

The prior probability for the frequency of the solvent resonance, $P(f_s|I)$, is assigned as a Gaussian using input from the user. Using a double cursor the user is required to mark the low, f_{low}, and high, f_{high}, solvent frequency and the prior then constrains the solvent frequency to be within these marked values. The mean of this Gaussian is the mean of the low and high, so $\hat{f}_s = (f_{low}+f_{high})/2$, while the interval high minus low is taken to be a three standard deviation interval, so $\hat{\sigma}_f = (f_{high} - f_{low})/3$. The prior probability for the solvent resonance is then given by

$$P(f_s|I) \propto \begin{cases} \exp\left\{-\dfrac{(\hat{f}_s - f_s)^2}{2\hat{\sigma}_f^2}\right\} & \text{if } f_{low} \leq f_s \leq f_{high} \\ 0 & \text{otherwise} \end{cases} \tag{9.18}$$

Because of its width and because it is centered at the peak in the Fourier transform, this prior ensures that the Monte Carlo simulations do not wander into a nonphysical region of the parameter space. Additionally, the user has the option of not specifying a solvent resonance. When this is done the solvent frequency and decay rate constant, f_s and α_s, are set equal to zero and the priors are not used in the calculation. Under these conditions the solvent model reduces to a trend in the real and imaginary FID data.

In a similar vane the prior probability for the solvent decay rate constant is also assigned as a Gaussian:

$$P(\alpha_s|I) \propto \begin{cases} \exp\left\{-\dfrac{(\hat{\alpha}_s - \alpha_s)^2}{2\hat{\sigma}_\alpha^2}\right\} & \text{if } 0 \leq \alpha_s \leq \alpha_{high} \\ 0 & \text{otherwise} \end{cases} \tag{9.19}$$

where $\alpha_{high} = 10(f_{high} - f_{low})$, $\hat{\alpha}_s = (f_{high} - f_{low})/2$ and $\sigma_\alpha = f_{high} - f_{low}$. This prior also serves little purpose other than to keep the Monte Carlo simulations from wandering into a nonphysical region of parameter space.

The prior probabilities for the expansion orders, $P(n_c|I)$ and $P(n_s|I)$, were assigned using an exponential prior probability:

$$P(n_c|I) \propto \exp\{-n_c\} \qquad \text{and} \qquad P(n_s|I) \propto \exp\{-n_s\}\,. \tag{9.20}$$

Note that we have indicated that these priors are proportional to an exponential. Normally in model selection problems one must use fully normalized prior probabilities. However, for this problem, the normalization constant for these priors cancel so we do not bother deriving it.

The prior probability for the standard deviation of the noise prior probability, $P(\sigma|I)$ will be assigned as a Jeffreys' prior:

$$P(\sigma|I) \propto \frac{1}{\sigma} \tag{9.21}$$

and we will not bound this prior. Normally this is a bad idea because it effectively introduces an infinity into the Bayesian calculation. We can get away with it here, because that same infinity is introduced into every model we consider in exactly the same way and so cancels out of the calculation when the distributions are normalized. However in many cases this cannot be done, and one must explicitly bound the above Jeffreys' prior and then integrate only over that bound.

The target distribution of the Monte Carlo simulation is a marginal distribution, Eq. (9.15), where the amplitudes have been removed by integration. When we assigned the prior probabilities for these amplitudes we assigned a broad unbounded Gaussian of the form:

$$P(\mathcal{A}_j|I) = \left(\frac{2\pi\sigma^2}{\beta^2 G_{jj}}\right)^{-\frac{1}{2}} \exp\left\{-\frac{\beta^2 G_{jj}\mathcal{A}_j^2}{2\sigma^2}\right\} \tag{9.22}$$

where σ is the standard deviation of the noise prior probability. The quantity G_{jj}, defined in Eq. (9.34) below, is the squared length of a model signal and serves to ensure that the prior information about a particular amplitude remains roughly the same. The hyperparameter β was set at $\beta = 0.01$, so when the marginalization occurs this prior affects the estimated amplitudes in the fourth decimal place: this prior stabilizes the matrix inversion that occurs when these amplitudes are marginalized.

The prior probabilities for the frequencies of the resonances are also generated from the input low and high frequency range. The interface computes a mean frequency, $\hat{f}_l$, as the average of the low and high frequency. And it takes the interval, high-low, as a 3 standard deviation interval. The prior probability for the frequency of the lth resonance of interest are then assigned as

$$P(f_l|I) \propto \exp\left\{-\frac{\left(\hat{f}_l - f_l\right)^2}{2\hat{\sigma}_l^2}\right\}, \tag{9.23}$$

subject to the condition $f_{low} \leq f_l \leq f_{high}$, where f_{low} and f_{high} are the input low and high parameter ranges.

Similarly, the prior probability for the decay rate constants are also assigned as Gaussians using the input linewidth. If we designate this linewidth as lb then the program generates

$$P(\alpha_l|I) \propto \exp\left\{-\frac{\alpha_l^2}{2\times(3lb)^2}\right\} \tag{9.24}$$

subject to the condition $0 \leq \alpha_l \leq 10lb$. So this prior keeps the decay rate small, and allows the program to explore a wide range of decay rate constants, while keeping the rate constants near the input lb value.

Finally, Gaussian noise prior probabilities were used to assign the two likelihoods. The likelihood for the real part of the FID data was assigned as

$$P(D_R|f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s\sigma\mathcal{A}I) = \left(2\pi\sigma^2\right)^{-\frac{N}{2}} \exp\left\{-\sum_{i=1}^{N}\frac{\left[d_R(t_i)-M_R(t_i)\right]^2}{2\sigma^2}\right\}. \tag{9.25}$$

Similarly, the likelihood for the imaginary part of the FID data was assigned as

$$P(D_I|f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s\sigma\mathcal{A}I) = \left(2\pi\sigma^2\right)^{-\frac{N}{2}} \exp\left\{-\sum_{i=1}^{N}\frac{\left[d_I(t_i)-M_I(t_i)\right]^2}{2\sigma^2}\right\}. \tag{9.26}$$

We have now assigned all of the probabilities appearing in Eq. (9.15), and so are now able to evaluate the integrals over the amplitudes and the standard deviation of the noise prior probability. To begin this evaluation we substitute Eqs. (9.16,9.17,9.18,9.19, 9.20,9.21,9.22,9.23,9.24, 9.25,9.26) into Eq. (9.15) to obtain

$$\begin{aligned}
P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0\theta n_c n_s|D_R D_I I) \quad &\propto \quad \tfrac{1}{360} \\
&\times \quad (2\pi\sigma_{t_0}^2)^{-\frac{1}{2}} \exp\left\{-\tfrac{t_0^2}{2\sigma_{t_0}^2}\right\} \\
&\times \quad \exp\left\{-\tfrac{(\hat{f}_s-f_s)^2}{2\hat{\sigma}_f^2}\right\} \\
&\times \quad \exp\left\{-\tfrac{(\hat{\alpha}_s-\alpha_s)^2}{2\hat{\sigma}_\alpha^2}\right\} \\
&\times \quad \exp\left\{-n_c\right\} \\
&\times \quad \exp\left\{-n_s\right\} \\
&\times \quad \textstyle\prod_{l=1}^{m}(2\pi\hat{\alpha}_l^2)^{-\frac{1}{2}} \exp\left\{-\tfrac{(\hat{f}_l-f_l)^2}{2\hat{\alpha}_l^2}\right\} \\
&\times \quad \textstyle\prod_{l=1}^{m}(2\pi\hat{\alpha}_l^2)^{-\frac{1}{2}} \exp\left\{-\tfrac{(\hat{\alpha}_l-\alpha_l)^2}{2\hat{\alpha}_l^2}\right\} \\
&\times \quad \int d\mathcal{A}d\sigma \\
&\times \quad \tfrac{1}{\sigma} \\
&\times \quad \textstyle\prod_{j=1}^{n_{\mathcal{A}}}\left(\tfrac{2\pi\sigma^2}{\beta^2 G_{jj}}\right)^{-\frac{1}{2}} \exp\left\{-\tfrac{\beta^2 G_{jj}\mathcal{A}_j^2}{2\sigma^2}\right\} \\
&\times \quad (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\textstyle\sum_{i=1}^{N}\tfrac{[d_R(t_i)-M_R(t_i)]^2}{2\sigma^2}\right\} \\
&\times \quad (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\textstyle\sum_{i=1}^{N}\tfrac{[d_I(t_i)-M_I(t_i)]^2}{2\sigma^2}\right\}
\end{aligned} \tag{9.27}$$

where each of the prior probabilities have been written on a separate line for easy identification.

Now in preparation of evaluating the integrals over the amplitudes and standard deviation of the noise prior probability, we are going to drop all constants that are common to each model, one obtains

$$\begin{aligned} P(f_s\alpha_s f_1\alpha_1\ldots f_m\alpha_m t_0\theta n_c n_s|D_R D_I I) &\propto \exp\left\{-\tfrac{t_0^2}{2\sigma_{t_0}^2}-\tfrac{(\hat{f}_s-f_s)^2+(\hat{\alpha}_s-\alpha_s)^2}{2\hat{\alpha}_s^2}-n_c-n_s\right\} \\ &\times \exp\left\{-\textstyle\sum_{l=1}^m \tfrac{(\hat{f}_l-f_l)^2+(\hat{\alpha}_l-\alpha_l)^2}{2\hat{\alpha}_l^2}\right\} \\ &\times \int d\mathcal{A}d\sigma\sigma^{-2N-1}\prod_{j=1}^{n_\mathcal{A}}\left(\frac{2\pi\sigma^2}{\beta^2 G_{jj}}\right)^{-\frac{1}{2}}\exp\left\{-\frac{\beta^2\mathcal{A}_j^2 G_{jj}}{2\sigma^2}\right\} \\ &\times \exp\left\{-\textstyle\sum_{i=1}^N \tfrac{[d_R(t_i)-M_R(t_i)]^2+[d_I(t_i)-M_I(t_i)]^2}{2\sigma^2}\right\}. \end{aligned} \tag{9.28}$$

Now designating the integral as $\mathcal{I}$ and working only on that integral by first substituting Eqs. (9.12) and (9.13) into $\mathcal{I}$ one obtains

$$\mathcal{I}=\prod_{j=1}^{n_\mathcal{A}}\left(\frac{2\pi}{\beta^2 G_{jj}}\right)^{-\frac{1}{2}}\int d\mathcal{A}d\sigma\sigma^{-2N-1-n_\mathcal{A}}\exp\left\{-\frac{Q}{2\sigma^2}\right\} \tag{9.29}$$

where

$$Q\equiv 2N\overline{d^2}-2\sum_{l=1}^{n_\mathcal{A}}\mathcal{A}_l T_l+\sum_{j=1}^{n_\mathcal{A}}\sum_{k=1}^{n_\mathcal{A}}\mathcal{A}_j\mathcal{A}_k g_{jk}. \tag{9.30}$$

The mean-square data value, $\overline{d^2}$, is given by

$$\overline{d^2}=\frac{1}{2N}\sum_{i=1}^{2N}\left[d_R(t_i)^2+d_I(t_i)^2\right]. \tag{9.31}$$

The projection of the data onto the lth model function, T_l, is given by

$$T_l=\sum_{i=1}^{N}\left[d_R(t_i)R_l(t_i)+d_I(t_i)I_l(t_i)\right]. \tag{9.32}$$

The matrix g_{jk} is given by

$$g_{jk}=(1+\beta^2\delta_{jk})G_{jk} \tag{9.33}$$

where δ_{jk} is the Kronecker delta function and

$$G_{jk}\equiv\sum_{i=1}^{N}\left[R_j(t_i)R_k(t_i)+I_j(t_i)I_k(t_i)\right]. \tag{9.34}$$

The amplitude integrals are multivariate Gaussian integrals and evaluating such integrals is straight-forward, one obtains

$$\mathcal{I}=|g_{jk}|^{-\frac{1}{2}}\prod_{j=1}^{n_\mathcal{A}}\left(\beta^2 G_{jj}\right)^{\frac{1}{2}}\int d\sigma\sigma^{-2N-1}\exp\left\{-\frac{2N\overline{d^2}-n_\mathcal{A}\overline{h^2}}{2\sigma^2}\right\} \tag{9.35}$$

where the factors involving 2π have canceled,

$$\overline{h^2} \equiv \frac{1}{n_{\mathcal{A}}} \sum_{j=1}^{n_{\mathcal{A}}} T_j \hat{\mathcal{A}}_j \tag{9.36}$$

and the $\hat{\mathcal{A}}_j$ are given by the solution to

$$\sum_{j=1}^{n_{\mathcal{A}}} g_{jk} \hat{\mathcal{A}}_j = T_k. \tag{9.37}$$

The amplitudes, $\hat{\mathcal{A}}$, are the maximum posterior probability estimates given the parameter values.

The integral over the standard deviation of the noise prior probability may be transformed into a Gamma function and the integral is then easily evaluated. Evaluating the integral, one obtains

$$\mathcal{I} = |g_{jk}|^{-\frac{1}{2}} \prod_{j=1}^{n_{\mathcal{A}}} \left(\beta^2 G_{jj}\right)^{\frac{1}{2}} \left[\frac{2N\overline{d^2} - n_{\mathcal{A}}\overline{h^2}}{2}\right]^{-N} \tag{9.38}$$

where a factor of $2\Gamma(N)$ was dropped because it cancels when the posterior probability for the nonlinear parameters is normalized.

Finally inserting $\mathcal{I}$ back into Eq. (9.28), one obtains

$$\begin{aligned} P(f_s\alpha_s f_1\alpha_1 \ldots f_m\alpha_m t_0 \theta n_c n_s | D_R D_I I) \quad &\propto \quad \exp\left\{-\frac{t_0^2}{2\sigma_{t_0}^2} - \frac{(\hat{f}_s - f_s)^2}{2\sigma_f^2} - \frac{(\hat{\alpha}_s - \alpha_s)^2}{2\hat{\sigma}_\alpha^2} - n_c - n_s\right\} \\ &\times \quad \exp\left\{-\sum_{l=1}^{m} \frac{(\hat{f}_l - f_l)^2 + (\hat{\alpha}_l - \alpha_l)^2}{2\hat{\alpha}_l^2}\right\} \\ &\times \quad |g_{jk}|^{-\frac{1}{2}} \prod_{j=1}^{n_{\mathcal{A}}} \left(\beta^2 G_{jj}\right)^{\frac{1}{2}} \left[\frac{2N\overline{d^2} - n_{\mathcal{A}}\overline{h^2}}{2}\right]^{-N} \end{aligned} \tag{9.39}$$

as the posterior probability for the nonlinear parameters. It is this posterior probability that is targeted by the Monte Carlo simulation. For a bit more on this topic, see [18]. There you will find some examples of this calculation and more on how the various prior probabilities were assigned.

9.2 Outputs From The Big Peak/Little Peak Package

The text outputs files from the Big Peak/Little Peak packages consist of: "Bayes.prob.model," "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 **??**. 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 each of the marked frequencies including the solvent frequency. These probability density plots include plots for the frequency, decay rate constant and the amplitudes. In the case of the solvent, there are a number of additional plots that help one determine how the water resonance is changing as a function of time. These plots are shown in Fig. 9.2. Here is a description of these additional plots:

- Fig. 9.2 Panel **a** is a two line plot. The red line is a plot of the logarithm of the estimated solvent envelope as a function of acquisition time. The solvent envelope is defined in Eq. (9.2) as the function $A(t)\exp\{-\alpha_s t\}$. If a resonance is a pure exponentially decaying sinusoid, this function should be a straight line whose slope would be estimated to be $-\alpha'_s$, the blue line, where α'_s is the estimated slope given an exponentially decaying sinusoidal model.

- Fig. 9.2 panel **b** is the difference between the two lines shown in panel **a**. The smaller this difference is, the more exponential the decay is.

- Panel **c** is the phase of the solvent frequency as a function of the time. The phase of the solvent is defined in Eq. (9.2) as the function $\phi(t_i)$ and this plot starts at zero because the constant phase has been removed.

- Panel **d** in Fig. 9.2 is the accumulated phase of the water resonance. If the water frequency is a constant, this plot would be the water frequency times time.

Figure 9.2: The Time Dependent Parameters

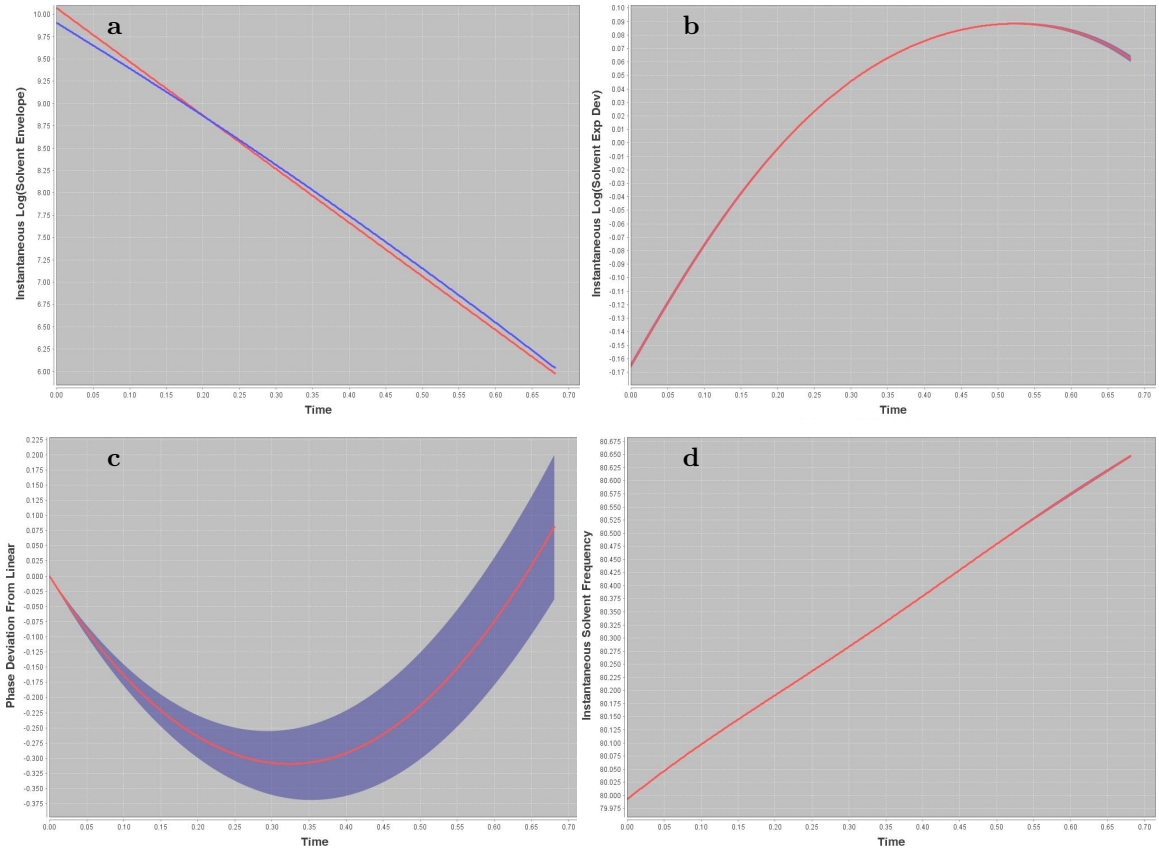

Figure 9.2: Panel **a** shows the logarithm of the estimated instantaneous decay envelope of the water (blue) as a function of time. The red line is the logarithm of the best fit decay envelope assuming an exponentially decaying sinusoid. Note that the instantaneous water frequency starts low, and then gets faster. Eventually the difference between these two lines goes to a constant. Panel **b** is the difference between the two curves shown in Panel **a**, i.e., the deviation from linear. Panel **c** is a deviation in the phase from a constant. Panel **d** is the accumulated solvent frequency as a function of time. As one can see, this curve is highly linear indicating that these phase variations are small.

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, Marilyn 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, Marilyn 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] Marilyn 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.

Index