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 12

Diffusion Tensor Analysis

The diffusion tensor package estimates the parameters associated with diffusion tensor models. The data analyzed by this package are Ascii and may be input from an Ascii file, they can be loaded from the amplitudes resulting from a Bayes Analyze run or they may be input from an image pixel. The diffusion tensor package is accessed by selecting the "Packages/Diffusion Tensor" menu. When this button is activated the interface window shown in Fig. 12.1 is displayed. To use the Diffusion Tensor package, you must do the following:

Check the abscissa options appropriate for your input data.

Load one or more Ascii data sets using the Files menu. This data must have three abscissa for both "b" and "g" vector data. The format of a diffusion tensor data set is: data point number, data value, A_r, A_p, A_s, where "A_r" is either a "B" or "G" value in the readout direction. Similarly, "A_p" is a "B" or "G" value in the phase encode direction "A_s" is a "B" or "G" value in the slice select direction. When a data set is successfully loaded it will be displayed in the Files Viewer.

Set the number of of tensors you wish to analyze. This number can be 1, 2 or 3 and is set using "Tensor Number" pull down menu. selection menu.

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

Check the Analysis Options/Find Outliers box if you suspect outliers are present in the data.

Review the prior probabilities for the the various parameters in the model using the Prior Viewer.

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.

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.

Figure 12.1: The Diffusion Tensor Package Interface

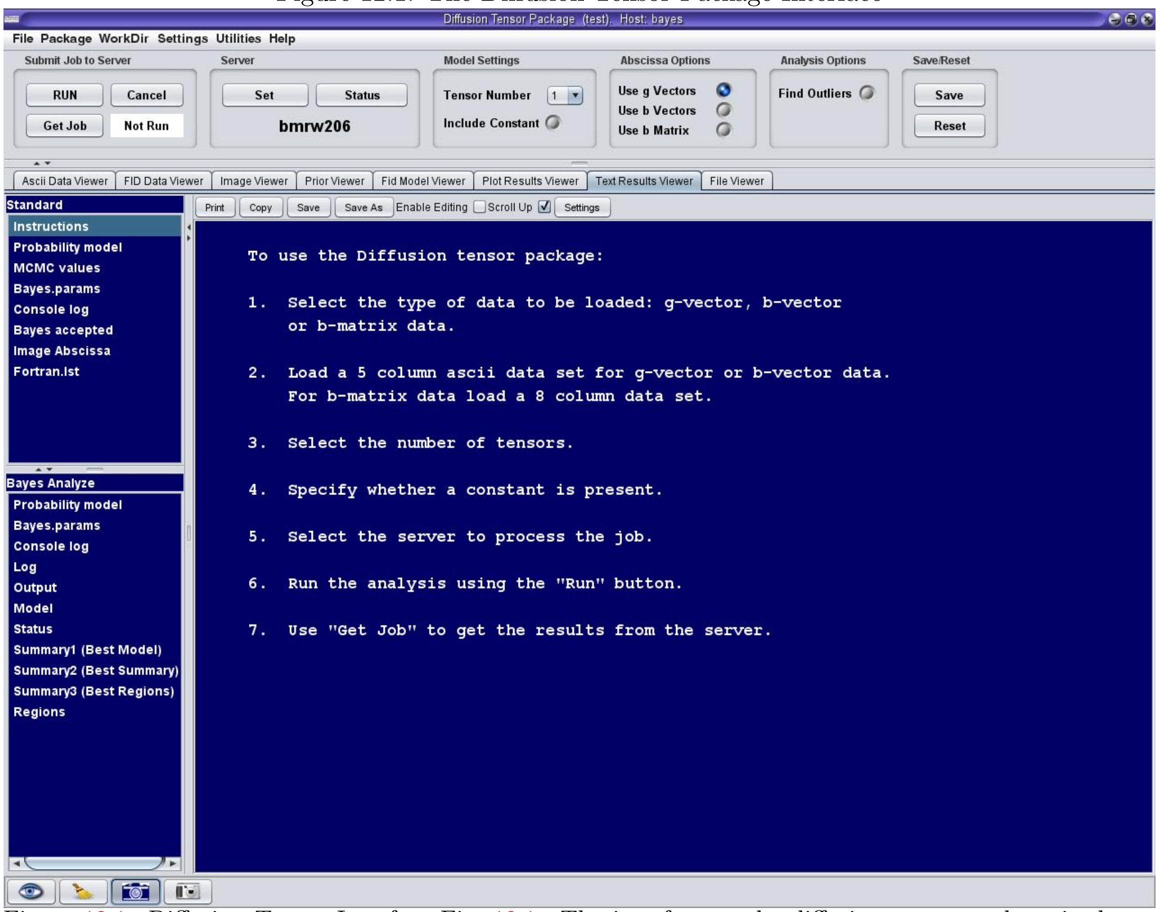

Figure 12.1: Diffusion Tensor Interface Fig. 12.1. The interface to the diffusion tensor package is shown here. To use this package one first selects the type diffusion tensor data to be analyzed, using the "Abscissa Options." Depending on what check box is activated, the input data will have either 3 or 6 abscissa values. After the Abscissa Options are set, the data may be loaded, the number of tensors can then be selected and whether or not a constant offset is present. Finally, the analysis can be run using the "Run" button.

12.1 The Bayesian Calculation

A diffusion Tensor **D** is a 3×3 symmetric matrix having six independent elements:

$$\mathbf{D} = \begin{pmatrix} D_{rr} & D_{pr} & D_{sr} \\ D_{rp} & D_{pp} & D_{sp} \\ D_{rs} & D_{ps} & D_{ss} \end{pmatrix} \tag{12.1}$$

where there are three redundant elements, $D_{rp} = D_{pr}$, $D_{ps} = D_{sp}$ and $D_{rs} = D_{sr}$. Using this diffusion tensor, the data and the model are related to each other by

$$d_i = A \exp\left\{ -\text{Conv} \sum_{j=1}^{3} \sum_{k=1}^{3} g_{ij} D_{jk} g_{ik} \right\} + n_i \tag{12.2}$$

where d_i is the ith data value acquired using gradients (g_{i1}, g_{i2}, and g_{i3}), where the first, second and third directions are x, y and z respectively, "Conv" is a conversion constant that ensures the argument of the exponential is unitless, and n_i represents the noise. For a rectangular gradient this conversion factor is given by:

$$\text{Conv} = \Gamma^2 \delta^2 (\Delta - \delta/3) \tag{12.3}$$

where Γ is the gyromagnetic ratio, usually of water, δ is the duration of the gradient pulse, Δ is the time between gradient pulse.

The diffusion tensor model, Eq. (12.2) without the noise, is what is used in most diffusion tensor analyzes. To estimate the diffusion matrix, the experimenter takes a number of diffusion weighted data values, typically 6, plus a data value having no diffusion gradients. This nondiffusion weighted data value is used as an estimate of the amplitude A, and the data are divided by this amplitude estimate. If the measurement are noiseless, and there are no other effects in the data, this cancel the amplitude from the right-hand side of Eq. (12.2). Thus if one takes the logarithm, the resulting equations are linear in the elements of the diffusion matrix. One then solves the resulting set of linear equations for the diffusion matrix. Finally, an eigenvalue decomposition of the diffusion matrix results in an estimate of the magnitude of the diffusion along three primary directions. Here we are going to call these principal diffusion magnitudes λ_1, λ_2 and λ_3 respectively. The eigenvectors are a rotation matrix **R** characterized by three Euler angles and these Euler angles can also be estimated from the eigenvectors. Unfortunately, this procedure has many problems including, but not limited to the fact that the amplitude does not cancel because the measured amplitude is not noiseless. In some samples the diffusion data does not go to zero and, more importantly, this procedure can generate negative eigenvalues, a completely unphysical result because a real symmetric matrix must have three positive eigenvalues. Consequently, a different approach is needed to resolve these problems.

The negative eigenvalue problem can be solved by introducing the eigenvalues directly into the problem, and then using Bayesian probability theory and prior probabilities to prohibit negative eigenvalues. Here is how this is done. First, one introduces a diagonal diffusion tensor **U** defined as:

$$\mathbf{U} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{12.4}$$

where λ_1, λ_2 and λ_3 are the unknown magnitudes of the diffusion along an as yet unknown set of directions. Next one introduces the rotation matrix $\mathbf{R}$ mentioned earlier. In this matrix the three Euler angles, θ, ϕ and ψ, are unknowns that must be inferred:

$$\mathbf{R} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \times \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{12.5}$$

where we have expressed this rotation matrix as a product of the three rotations need rotate a diagonal matrix into any 3×3 symmetric matrix. Multiplying the three matrices, one obtains:

$$\mathbf{R} = \begin{pmatrix} R_{rr} & R_{pr} & R_{sr} \\ R_{rp} & R_{pp} & R_{sp} \\ R_{rs} & R_{ps} & R_{ss} \end{pmatrix}, \tag{12.6}$$

with

$$R_{rr} = \cos(\psi)\cos(\theta)\cos(\phi) - \sin(\psi)\sin(\phi), \tag{12.7}$$
$$R_{pr} = \cos(\psi)\cos(\theta)\sin(\phi) + \sin(\psi)\cos(\phi), \tag{12.8}$$
$$R_{sr} = -\cos(\psi)\sin(\theta), \tag{12.9}$$
$$R_{rp} = -\sin(\psi)\cos(\theta)\cos(\phi) - \cos(\psi)\sin(\phi), \tag{12.10}$$
$$R_{pp} = -\sin(\psi)\cos(\theta)\sin(\phi) + \cos(\psi)\cos(\phi), \tag{12.11}$$
$$R_{sp} = \sin(\theta)\sin(\psi), \tag{12.12}$$
$$R_{rs} = \sin(\theta)\cos(\phi), \tag{12.13}$$
$$R_{ps} = \sin(\theta)\sin(\phi), \tag{12.14}$$

and

$$R_{ss} = \cos(\theta). \tag{12.15}$$
$$\tag{12.16}$$

If we now rotate the diagonal tensor of eigenvalues, Eq. (12.4), into a nondiagional space, one obtains

$$\mathbf{D} = \mathbf{R}^{\mathrm{T}} \times \mathbf{U} \times \mathbf{R} \tag{12.17}$$

where $\mathbf{D}$ is the same tensor defined in Eq. (12.1). Explicitly working out this rotation results in:

$$D_{rr} = R_{rr}\lambda_1 R_{rr} + R_{pr}\lambda_2 R_{pr} + R_{sr}\lambda_3 R_{sr}, \tag{12.18}$$
$$D_{rp} = R_{rr}\lambda_1 R_{rp} + R_{pr}\lambda_2 R_{pp} + R_{sr}\lambda_3 R_{sp}, \tag{12.19}$$
$$D_{rs} = R_{rr}\lambda_1 R_{rs} + R_{pr}\lambda_2 R_{ps} + R_{sr}\lambda_3 R_{ss}, \tag{12.20}$$
$$D_{pr} = D_{rp}, \tag{12.21}$$
$$D_{pp} = R_{rp}\lambda_1 R_{rp} + R_{pp}\lambda_2 R_{pp} + R_{sp}\lambda_3 R_{sp}, \tag{12.22}$$
$$D_{ps} = R_{rp}\lambda_1 R_{rs} + R_{pp}\lambda_2 R_{ps} + R_{sp}\lambda_3 R_{ss}, \tag{12.23}$$
$$D_{sr} = D_{rs}, \tag{12.24}$$
$$D_{sp} = D_{ps}, \tag{12.25}$$

and

$$D_{ss} = R_{rs}\lambda_1 R_{rs} + R_{ps}\lambda R_{ps} + R_{ss}\lambda_3 R_{ss}, \tag{12.26}$$
$$\tag{12.27}$$

which means that we can apply Bayesian probability theory to infer the eigenvalues and Euler angles directly from the data without taking any logarithms and without attempting to divide the amplitude of the tensor from the problem. Additionally, because we can use the eigenvalues and Euler angles directly, we will be able to infer these quantities even when the data contain multiple diffusion tensors; something not possible with the procedures described earlier.

In the Bayesian calculations that follow, it will be assumed that there are multiple diffusion tensors in a given data set. Consequently, the notation used up to now must be modified. Here we introduce an index that represents the ℓth diffusion tensor. Additionally, we are going to include an constant offset in the diffusion tensor model. The presence of this constant will be under user control, so it may be included or excluded at the user's discretion. Equation (12.2) will be rewritten as:

$$d_i = C\delta(\nu) + \sum_{\ell=1}^{m} A_\ell \exp\left\{\text{Conv}\sum_{j=1}^{3}\sum_{k=1}^{3} g_{ij} D_{jk\ell} g_{ik}\right\} + n_i \tag{12.28}$$

where m is the number of diffusion tensors, "Conv" is the conversion factor defined in Eq. (12.3), $D_{jk\ell}$ is the jth column, the k row, of the ℓth diffusion tensor, g_{ik} is the kth component of the ith gradient, A_ℓ is the amplitude of the ℓth tensor, C is the constant offset, and $\delta(\nu)$ is an indicator function defined as:

$$\delta(\nu) = \begin{cases} 1 & \text{If the user included a constant} \\ 0 & \text{Otherwise} \end{cases}. \tag{12.29}$$

We will combine the constant offset and the amplitudes into a single set of amplitudes represented by: $B_\ell \in \{C, A_1, \ldots, A_m\}$. The number of B amplitudes is n with $n = m + \delta(\nu)$. Similarly, we will define a new model function G_ℓ that combines the diffusion tensors and the constant:

$$G_\ell(\Theta, i) \equiv \begin{cases} i^0 & \text{If } \ell = 1 \\ \exp\left[\text{Conv}\displaystyle\sum_{j=1}^{3}\sum_{k=1}^{3} g_{ij} D_{jkm} g_{ik}\right] & \text{otherwise, with } m = \ell - 1 \end{cases}, \tag{12.30}$$

where we intentionally wrote the constant model as i^0 as a reminder to the reader that the constant model is 1 for all i. Additionally, we have included i in the argument list of the function G as a reminder that G is a function of the data point number. In the above we are using Θ to stand for all of the diffusion tensor parameters, so

$$\Theta \equiv \{\lambda_{11}, \lambda_{12}, \lambda_{13}, \theta_1, \phi_1, \psi_1, \ldots, \lambda_{m1}, \lambda_{m2}, \lambda_{m3}, \theta_m, \phi_m, \psi_m\}. \tag{12.31}$$

In this notation, the Eq. (12.2) can be written as

$$d_i = \sum_{\ell=1}^{n} B_\ell G_\ell(\Theta, i) + n_i \tag{12.32}$$

and it is in this form that it will be used in the Bayesian calculations for the posterior probability for the Θ parameters.

The Diffusion Tensor package computes the marginal posterior probability for each of the parameters appearing in the model. For computational convenience we are going to marginalize out the amplitudes, the constant offset and the standard deviation of the noise. Consequently, the target distribution for the Markov chain will be the joint marginal posterior probability for all of the diffusion parameters Θ. This joint posterior probability is computed from the joint posterior probability for all of the parameters, including the amplitudes and the standard deviation of the noise:

$$P(\Theta|DI) = \int P(B_1 \ldots B_n \sigma\Theta|DI) dB_1 \ldots dB_n d\sigma. \tag{12.33}$$

The right-hand side of this equation is factored using Bayes' theorem

$$P(\Theta|DI) \propto \int P(B_1 \ldots B_m \sigma\Theta|I) P(D|B_1 \ldots B_n \sigma\Theta I) dB_1 \ldots dB_n d\sigma \tag{12.34}$$

where $P(B_1 \ldots B_n \sigma\Theta|I)$ is the joint prior probability for all of the parameters and the other term, the direct probability for the data, $P(D|B_1 \ldots B_n \sigma\Theta I)$, is often called the likelihood function. Next the product rule of probability theory is used to factor the joint prior probability for the parameters into a series of independent prior probabilities, one for each parameter,

$$\begin{aligned} P(B_1 \ldots B_n \sigma\Theta|I) \quad &\propto \quad P(\sigma|I) \prod_{\ell=1}^{n} P(B_\ell|I) \\ &\times \quad \prod_{j=1}^{m} [P(\lambda_{j1}|I) P(\lambda_{j2}|I) P(\lambda_{j3}|I)] \\ &\times \quad \prod_{j=1}^{m} [P(\theta_j|I) P(\phi_j|I) P(\psi_j|I)] \end{aligned} \tag{12.35}$$

where to make this factorization we assumed logical independents, i.e., the prior probability we would assign to one parameter does not depend on any other parameters. Substituting, the prior probability for the parameters, Eq. (12.35), into the marginal posterior probability for the Θ parameters, Eq. (12.34),one obtains:

$$\begin{aligned} P(\Theta|DI) \quad &\propto \quad \int \prod_{\ell=1}^{n} [P(B_\ell|I)] \\ &\times \quad \prod_{j=1}^{m} [P(\lambda_{j1}|I) P(\lambda_{j2}|I) P(\lambda_{j3}|I)] \\ &\times \quad \prod_{j=1}^{m} [P(\theta_j|I) P(\phi_j|I) P(\psi_j|I)] \\ &\times \quad P(D|B_1 \ldots B_n \sigma\Theta I) dB_1 \ldots dB_n d\sigma \end{aligned} \quad . \tag{12.36}$$

We have reached the point in the calculation where we must assign probabilities to represent the various terms in this equation. The prior probability for the amplitudes B_ℓ will be assigned as a Gaussians given by

$$P(B_\ell|I) = \left(\frac{2\pi\sigma^2}{g_{\ell\ell}\beta^2}\right)^{-\frac{1}{2}} \exp\left\{-\frac{g_{\ell\ell}\beta^2 B_\ell^2}{2\sigma^2}\right\} \tag{12.37}$$

where σ is the standard deviation of the prior probability for the errors, β is hyperparameter and expresses how strongly we think we can guess the value of the amplitudes relative to how well it can be determined from the data and $g_{\ell\ell}$, the squared-length of the function $G_\ell(\Theta, i)$, is defined in Eq. (12.44) below. The prior probability for the noise standard deviation, $P(\sigma|I)$, was assigned a Jeffreys' prior

$$P(\sigma|I) \propto \frac{1}{\sigma}. \tag{12.38}$$

The prior probabilities for the eigenvalues are defined by the user. When the user loads a set of data, the interface will make its best guess for the prior probabilities and set a default Gaussian prior. The user, using the prior viewer, can modify or completely replace these default prior probabilities.

The model equation is symmetric under relabeling of the amplitudes and tensors parameters and each tensor is symmetric under reordering the rows or columns of the tensors. These symmetry cause the joint posterior probability for the tensor parameters to be symmetric in the sense that if there is a peak at $\lambda_{11} = \beta$ and $\lambda_{12} = \gamma$, then there is also a peak at $\lambda_{11} = \gamma$ and $\lambda_{12} = \beta$; this symmetry occurs because the model does not tell us which eigenvalue corresponds to to which model component. Consequently, a convention must be introduced which brakes this symmetry by identifying a model component with a signal component. In the calculations implemented here, we break this symmetry by ordering the eigenvalues within and across tensors. By ordering the eigenvalues this manner, we effectively tell probability theory what we mean when we say, "eigenvalue one," we mean the largest eigenvalue. We do impose one additional condition on these eigenvalues, if there are multiple tensors, then we impose the additional condition: $0 < \lambda_{11} < \lambda_{21} < \cdots < \lambda_{m1} <$ High, where "High" is the upper bound in the prior probability for the eigenvalues. In effect telling probability theory that when we say tensor 1, we mean the tensor having the smallest principle eigenvalue, etc. Finally, we will assign the prior probabilities for the angles as uniform prior probabilities of the form:

$$P(\text{angle}|I) = \left\{ \begin{array}{ll} & \text{If } 0 \leq \text{angle} \leq \pi \\ 0 & \text{Otherwise} \end{array} \right. , \tag{12.39}$$

where "angle" is any of the Euler angles in any of the tensors. The direct probability for the data was assigned using a Gaussian likelihood. Gathering up the prior probabilities and assigning the likelihood, the joint posterior probability for the parameters, Eq. (12.34), is given by:

$$P(\Theta|DI) \propto \int (2\pi\sigma^2)^{-\frac{N+m}{2}} \prod_{j=1}^{m} \sqrt{g_{jj}}\beta \prod_{k=1}^{3} P(\lambda_{jk}|I) \exp\left\{-\frac{Q}{2\sigma^2}\right\} dB_1 \dots dB_n d\sigma \tag{12.40}$$

where we have left the prior probabilities for the eigenvalues in symbolic form because they are user defined. The quantity Q is given by

$$Q \equiv N\overline{d^2} - 2N \sum_{j=1}^{m} B_j T_j(\Theta) + N \sum_{j=1}^{m} \sum_{k=1}^{m} B_j B_k g_{jk}(1 + \beta^2 \delta_{jk}) \tag{12.41}$$

and is essentially Chi-squared where

$$\overline{d^2} = \frac{1}{N} \sum_{i=1}^{N} d_i^2 \tag{12.42}$$

is the mean-square data value. The T_j

$$T_j = \frac{1}{N}\sum_{i=1}^{N} d_i G_j(\Theta, i) \tag{12.43}$$

are the mean projection of the data onto the model. Finally, the g_{jk} are defined as

$$g_{jk} = \frac{1}{N}\sum_{i=1}^{N} G_j(\Theta, i) G_k(\Theta, i). \tag{12.44}$$

The integrals over the amplitudes are Gaussians quadrature and the standard deviation the standard deviation is essentially a gamma function integral. Both of which are straight forward to evaluate and we omit the details, one obtains

$$P(\Theta|DI) \propto |g_{jk}|^{-\frac{1}{2}} \prod_{j=1}^{m} \sqrt{g_{jj}}\beta \prod_{k=1}^{3} P(\lambda_{jk}|I) \left[\frac{Q}{2}\right]^{-\frac{N}{2}}. \tag{12.45}$$

This probability density function is of the form of Students t-distribution, and it is this t-distribution that is targeted by the Markov chain Monte Carlo simulation.

For some examples of diffusion tensors imaging, see [14] and for more on the origins of diffusion tensors imaging, see [38, 44, 62, 63]. For those wishing to know more about how these integrals were done see [2, 61]. For a brief overview of this subject, WikipediA has a surprisingly good discussion.

12.2 Using The Package

The type of data loaded by the diffusion tensor package is dependent on the type of abscissa one selects, there are three options that may be chosen in the "Abscissa Options" box, by checking one of the three check boxes:

Use g Vectors tells the package that the abscissa of the input data will contain gradient or g-vectors. When g-vectors are used, the package converts them into b-vectors using

$$b_j = \Gamma\delta\sqrt{(\Delta - \delta/3)}g_j. \tag{12.46}$$

where j should be replaced by x, y or z. This conversion requires three additional parameters to be entered, Γ the gyromagnetic ratio in units of 1/(Gauss Sec), δ the duration of the gradient pulse in units of seconds and Δ the time between the two gradient pulses in units of seconds. These parameters are specified by three prior probabilities of prior type parameter, i.e., constants. So when this option is selected, three additional prior probabilities are included in prior list. Note that gradients are three dimensional vectors, so the input Ascii data file must contain: a data point number, the data, the gradients g_r, g_p and g_s, where the subscripts stand for the readout, phase encode and the slice select directions. Their appearance in the Ascii data file is a convention and they must appear in the given order, see Chapter A for more on Ascii file formats.

Use b Vectors indicates that the input Abscissa are b-vectors and consequently, no conversion to b-vectors is required. In this case, the Diffusion Tensor package sets the conversion factor to one. The input Ascii data file must contain: a data point number, the data, b_r, b_p and b_s in that order.

Use b Matrix indicates that the input data has a b-matrix for the abscissa. A b-matrix is a real symmetric 3×3 matrix having six independent elements:

$$B \text{ Matrix} \equiv \begin{pmatrix} b_{rr} & b_{pr} & b_{sr} \\ b_{rp} & b_{pp} & b_{sp} \\ b_{rs} & b_{ps} & b_{ss} \end{pmatrix}. \tag{12.47}$$

Consequently, for b-matrix data, the input file format is: a data point number, the data followed by b_{rr}, b_{pp}, b_{ss}, b_{rp}, b_{rs} and b_{ps}. Note the order of the elements in the b-matrix data file is a convention and must be in the given order.

After selecting the "Abscissa Options" one can load the diffusion data. As explained above, the type of data you must load is dependent on which Abscissa Options you selected. The model is specified by selecting the number of diffusion tensors, and indicating whether or not a constant offset is present. Setting the number of diffusion tensors is done using the "Tensor number" pull-down menu, see Fig. 12.1. Currently, you can select one, two or three tensors. To include a constant offset in the model, one checks the "Include Constant" check box.

The Text outputs files from the Diffusion Tensor packages consist of: "Bayes.prob.model," "BayesDiffTensor.mcmc.values," "Bayes.params," "Console.log," "Bayes.accepted" and finally 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 a small sample of this file is shown in Fig. 12.2. 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 eigenvalues, the three Euler angles, and the standard deviation of the noise.

The mean and standard deviation estimates of the various parameters are written to the "Bayes-DiffusionTensor.mcmc.values" output file, see Fig. 12.2 for a sample of this file. This file may be printed, or viewed using the interface. It consist of four general sections, the first section just identifies the model that was processed and relates some information about the posterior probability, the prior, and the likelihood. The second section contains the parameters that maximized the joint posterior probability. The third section contains mean and standard estimates of the parameters appearing in the Diffusion Tensor package. This third section also contains the logarithm of the posterior probability for the model. This probability is used to update the probabilities file and may be used to do model selection.

The Bayesian calculation for joint posterior probability for all of the parameters is done using Markov chain Monte Carlo with simulated annealing. In the annealing phase of the calculation the posterior probability for the model is computed. An example of this file is shown in Fig. 12.3. The "Bayes.prob.model" file contains the name of the model, the natural logarithm of the posterior probability for the model, the normalized probability for the model and the date and time the model was run.

The example shown in Fig. 12.3 was done using the "bayes/Bayes.test.data/DiffTensorAscii.Data". This simulated data set contains one diffusion tensor plus a constant. To illustrate how to use the

Figure 12.2: Diffusion Tensor Parameter Estimates

```
    ------------------- Eigenvectors -------------------
Tensor: 1    X                 Y                 Z
Vec: 1   -4.67473323E-01   -2.86124440E-01   -8.36421842E-01
Vec: 2   -7.92821559E-01   -2.82826221E-01    5.39854892E-01
Vec: 3   -3.91027707E-01    9.15501029E-01   -9.46319111E-02

  The expected parameter values (mean value of the probability distributions):

  Parameter Description             Mean Value      Std. Dev.     Peak Value
  Lambda 1 Tensor 1                 4.66911E-04    2.21524E-06    4.66811E-04
  Lambda 2 Tensor 1                 1.55124E-04    2.58916E-06    1.55834E-04
  Lambda 3 Tensor 1                 6.50448E-07    6.54884E-07    3.50572E-08
  Average Lambda 1                  2.07562E-04    1.19007E-06    2.07560E-04
  RA Tensor 1                       1.93932E-04    1.03217E-06    1.94039E-04
  FA Tensor 1                       9.34352E-01    5.63672E-03    9.34857E-01
  Theta Tensor 1                    9.53311E+01    8.87027E-01    9.54301E+01
  Phi Tensor 1                      1.13092E+02    3.09494E-01    1.13128E+02
  Psi Tensor 1                      3.27776E+01    4.09517E-01    3.28396E+01
  Dxx Tensor 1                      1.14992E-04    1.70590E-06    1.14795E-04
  Dxy Tensor 1                      1.85249E-04    1.16739E-06    1.85606E-04
  Dxz Tensor 1                      4.48854E-05    2.11995E-06    4.45133E-05
  Dyy Tensor 1                      3.06394E-04    2.01249E-06    3.05897E-04
  Dyz Tensor 1                      1.04314E-04    2.36628E-06    1.04367E-04
  Dzz Tensor 1                      2.01300E-04    1.74817E-06    2.01988E-04
  Amp Tensor 1 Set 1                1.00122E+02    6.10312E-02    1.00124E+02

  Const 1                           1.00030E+01    1.43077E-01    1.00079E+01
```

Figure 12.2: The output from the Diffusion Tensor Package consists of four parts. The first part identifies the model that was processed and then relates some information about the priors and the likelihood. This is followed by the parameters that maximized the joint posterior probability for the parameters. The maximum posterior probability parameters are followed by the eigenvectors that maximized the posterior probability. And these are followed by the mean and standard deviation estimates for the parameters show here.

Figure 12.3: Diffusion Tensor Posterior Probability For The Model

```
Model Name          Log(e) Prob      Probability  Date/Time Run
One D               -1.68351E+02      0.00000     Mon Jun  9 10:24:23 2003
One D + Const       -1.31755E+02      0.99997     Mon Jun  9 10:25:05 2003
Two D               -1.72142E+02      0.00000     Mon Jun  9 10:26:49 2003
Two D + Const       -1.42238E+02      0.00003     Mon Jun  9 10:31:02 2003
```

Figure 12.3: During the annealing phase the Diffusion Tensor Package computes the posterior probability for the model. This probability is written to the "Bayes.prob.model" file located in the experiment. If you run multiple models, the probabilities for the various models are appended to this file and may be used for model selection.

Diffusion Tensor package to do model selection, the one tensor model, then the one plus constant, etc. up to the two tensors plus a constant were run. Note that the logarithm of the posterior probability reaches a peak at the one tensor plus a constant model, and then decreases when the two tensor models were run.

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.