

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](mailto:larry@bayes.wustl.edu)

October 21, 2016



# Contents

<b>Manual Status</b>	<b>16</b>
<b>1 An Overview Of The Bayesian Analysis Software</b>	<b>19</b>
1.1 The Server Software	19
1.2 The Client Interface	22
1.2.1 The Global Pull Down Menus	24
1.2.2 The Package Interface	24
1.2.3 The Viewers	27
<b>2 Installing the Software</b>	<b>29</b>
<b>3 the Client Interface</b>	<b>33</b>
3.1 The Global Pull Down Menus	35
3.1.1 the Files menu	35
3.1.2 the Packages menu	40
3.1.3 the WorkDir menu	45
3.1.4 the Settings menu	46
3.1.5 the Utilities menu	50
3.1.6 the Help menu	50
3.2 The Submit Job To Server area	51
3.3 The Server area	52
3.4 Interface Viewers	52
3.4.1 the Ascii Data Viewer	53
3.4.2 the fid Data Viewer	53
3.4.3 Image Viewer	59
3.4.3.1 the Image List area	59
3.4.3.2 the Set Image area	62
3.4.3.3 the Image Viewing area	62
3.4.3.4 the Grayscale area on the bottom	63
3.4.3.5 the Pixel Info area	63
3.4.3.6 the Image Statistics area	64
3.4.4 Prior Viewer	65
3.4.5 Fid Model Viewer	68
3.4.5.1 The fid Model Format	70

3.4.5.2	The Fid Model Reports . . . . .	71
3.4.6	Plot Results Viewer . . . . .	71
3.4.7	Text Results Viewer . . . . .	74
3.4.8	Files Viewer . . . . .	80
3.5	Common Interface Plots . . . . .	80
3.5.1	Data, Model And Residual Plot . . . . .	81
3.5.2	Posterior Probability For A Parameter . . . . .	82
3.5.3	Maximum Entropy Histograms . . . . .	83
3.5.4	Markov Monte Carlo Samples . . . . .	83
3.5.5	Probability Vs Parameter Samples plot . . . . .	86
3.5.6	Expected Log Likelihood Plot . . . . .	88
3.5.7	Scatter Plots . . . . .	88
3.5.8	Logarithm of the Posterior Probability Plot . . . . .	91
3.5.9	Fortran/C Code Viewer . . . . .	91
3.5.9.1	Fortran/C Model Viewer Popup Editor . . . . .	94
<b>4</b>	<b>An Introduction to Bayesian Probability Theory</b>	<b>99</b>
4.1	The Rules of Probability Theory . . . . .	99
4.2	Assigning Probabilities . . . . .	102
4.3	Example: Parameter Estimation . . . . .	109
4.3.1	Define The Problem . . . . .	110
4.3.1.1	The Discrete Fourier Transform . . . . .	110
4.3.1.2	Aliases . . . . .	113
4.3.2	State The Model—Single-Frequency Estimation . . . . .	114
4.3.3	Apply Probability Theory . . . . .	115
4.3.4	Assign The Probabilities . . . . .	118
4.3.5	Evaluate The Sums and Integrals . . . . .	120
4.3.6	How Probability Generalizes The Discrete Fourier Transform . . . . .	123
4.3.7	Aliasing . . . . .	126
4.3.8	Parameter Estimates . . . . .	132
4.4	Summary and Conclusions . . . . .	136
<b>5</b>	<b>Given Exponential Model</b>	<b>137</b>
5.1	The Bayesian Calculation . . . . .	139
5.2	Outputs From The Given Exponential Package . . . . .	141
<b>6</b>	<b>Unknown Number of Exponentials</b>	<b>143</b>
6.1	The Bayesian Calculations . . . . .	145
6.2	Outputs From The Unknown Number of Exponentials Package . . . . .	148
<b>7</b>	<b>Inversion Recovery</b>	<b>151</b>
7.1	The Bayesian Calculation . . . . .	153
7.2	Outputs From The Inversion Recovery Package . . . . .	154

<b>8</b>	<b>Bayes Analyze</b>	<b>155</b>
8.1	Bayes Model	159
8.2	The Bayes Analyze Model Equation	161
8.3	The Bayesian Calculations	167
8.4	Levenberg-Marquardt And Newton-Raphson	171
8.5	Outputs From The Bayes Analyze Package	176
8.5.1	The “bayes.params.nnnn” Files	177
8.5.1.1	The Bayes Analyze File Header	178
8.5.1.2	The Global Parameters	182
8.5.1.3	The Model Components	184
8.5.2	The “bayes.model.nnnn” Files	185
8.5.3	The “bayes.output.nnnn” File	186
8.5.4	The “bayes.probabilities.nnnn” File	190
8.5.5	The “bayes.log.nnnn” File	193
8.5.6	The “bayes.status.nnnn” and “bayes.accepted.nnnn” Files	196
8.5.7	The “bayes.model.nnnn” File	197
8.5.8	The “bayes.summary1.nnnn” File	198
8.5.9	The “bayes.summary2.nnnn” File	199
8.5.10	The “bayes.summary3.nnnn” File	200
8.6	Bayes Analyze Error Messages	200
<b>9</b>	<b>Big Peak/Little Peak</b>	<b>207</b>
9.1	The Bayesian Calculation	209
9.2	Outputs From The Big Peak/Little Peak Package	216
<b>10</b>	<b>Metabolic Analysis</b>	<b>219</b>
10.1	The Metabolic Model	223
10.2	The Bayesian Calculation	225
10.3	The Metabolite Models	228
10.3.1	The IPGD_D2O Metabolite	228
10.3.2	The Glutamate.2.0 Metabolite	232
10.3.3	The Glutamate.3.0 Metabolite	235
10.4	The Example Metabolite	236
10.5	Outputs From The Bayes Metabolite Package	238
<b>11</b>	<b>Find Resonances</b>	<b>239</b>
11.1	The Bayesian Calculations	241
11.2	Outputs From The Bayes Find Resonances Package	246
<b>12</b>	<b>Diffusion Tensor Analysis</b>	<b>247</b>
12.1	The Bayesian Calculation	249
12.2	Using The Package	254
<b>13</b>	<b>Big Magnetization Transfer</b>	<b>259</b>
13.1	The Bayesian Calculation	259
13.2	Outputs From The Big Magnetization Transfer Package	262

<b>14 Magnetization Transfer</b>	<b>265</b>
14.1 The Bayesian Calculation	267
14.2 Using The Package	271
<b>15 Magnetization Transfer Kinetics</b>	<b>275</b>
15.1 The Bayesian Calculation	277
15.2 Using The Package	281
<b>16 Given Polynomial Order</b>	<b>285</b>
16.1 The Bayesian Calculation	287
16.1.1 Gram-Schmidt	287
16.1.2 The Bayesian Calculation	288
16.2 Outputs From the Given Polynomial Order Package	290
<b>17 Unknown Polynomial Order</b>	<b>293</b>
17.1 Bayesian Calculations	295
17.1.1 Assigning Priors	296
17.1.2 Assigning The Joint Posterior Probability	297
17.2 Outputs From the Unknown Polynomial Order Package	299
<b>18 Errors In Variables</b>	<b>303</b>
18.1 The Bayesian Calculation	305
18.2 Outputs From The Errors In Variables Package	308
<b>19 Behrens-Fisher</b>	<b>311</b>
19.1 Bayesian Calculation	311
19.1.1 The Four Model Selection Probabilities	314
19.1.1.1 The Means And Variances Are The Same	315
19.1.1.2 The Mean Are The Same And The Variances Differ	317
19.1.1.3 The Means Differ And The Variances Are The Same	318
19.1.1.4 The Means And Variances Differ	319
19.1.2 The Derived Probabilities	320
19.1.3 Parameter Estimation	321
19.2 Outputs From Behrens-Fisher Package	322
<b>20 Enter Ascii Model</b>	<b>329</b>
20.1 The Bayesian Calculation	331
20.1.1 The Bayesian Calculations Using Eq. (20.1)	331
20.1.2 The Bayesian Calculations Using Eq. (20.2)	332
20.2 Outputs Form The Enter Ascii Model Package	335
<b>21 Enter Ascii Model Selection</b>	<b>337</b>
21.1 The Bayesian Calculations	339
21.1.1 The Direct Probability With No Amplitude Marginalization	340
21.1.2 The Direct Probability With Amplitude Marginalization	342
21.1.2.1 Marginalizing the Amplitudes	343
21.1.2.2 Marginalizing The Noise Standard Deviation	348

21.2	Outputs Form The Enter Ascii Model Package	349
<b>26</b>	<b>Phasing An Image</b>	<b>395</b>
26.1	The Bayesian Calculation	396
26.2	Using The Package	402
<b>27</b>	<b>Phasing An Image Using Non-Linear Phases</b>	<b>405</b>
27.1	The Model Equation	405
27.2	The Bayesian Calculations	407
27.3	The Interfaces To The Nonlinear Phasing Routine	409
<b>28</b>	<b>Analyze Image Pixel</b>	<b>411</b>
28.1	Modification History	413
<b>29</b>	<b>The Image Model Selection Package</b>	<b>415</b>
29.1	The Bayesian Calculations	417
29.2	Outputs Form The Image Model Selection Package	418
<b>A</b>	<b>Ascii Data File Formats</b>	<b>423</b>
A.1	Ascii Input Data Files	423
A.2	Ascii Image File Formats	424
A.3	The Abscissa File Format	425
<b>B</b>	<b>Markov chain Monte Carlo With Simulated Annealing</b>	<b>439</b>
B.1	Metropolis-Hastings Algorithm	440
B.2	Multiple Simulations	441
B.3	Simulated Annealing	442
B.4	The Annealing Schedule	442
B.5	Killing Simulations	443
B.6	the Proposal	444
<b>C</b>	<b>Thermodynamic Integration</b>	<b>445</b>
<b>D</b>	<b>McMC Values Report</b>	<b>449</b>
<b>E</b>	<b>Writing Fortran/C Models</b>	<b>455</b>
E.1	Model Subroutines, No Marginalization	455
E.2	The Parameter File	458
E.3	The Subroutine Interface	460
E.4	The Subroutine Declarations	462
E.5	The Subroutine Body	463
E.6	Model Subroutines With Marginalization	464
<b>F</b>	<b>the Bayes Directory Organization</b>	<b>469</b>
<b>G</b>	<b>4dfp Overview</b>	<b>471</b>

8

**H Outlier Detection**

**475**

**Bibliography**

**479**



# List of Figures

1.1	The Start Up Window . . . . .	23
1.2	Example Package Exponential Interface . . . . .	25
2.1	Installation Kit For The Bayesian Analysis Software . . . . .	31
3.1	The Start Up Window . . . . .	34
3.2	The Files Menu . . . . .	35
3.3	The Files/Load Image Submenu . . . . .	37
3.4	The Packages Menu . . . . .	41
3.5	The Working Directory Menu . . . . .	46
3.6	The Working Directory Information Popup . . . . .	47
3.7	The Settings Pull Down Menu . . . . .	47
3.8	The McMC Parameters Popup . . . . .	48
3.9	The Edit Server Popup . . . . .	49
3.10	The Submit Job Widgets . . . . .	51
3.11	The Server Widgets Group . . . . .	52
3.12	The Ascii Data Viewer . . . . .	54
3.13	The Fid Data Viewer . . . . .	55
3.14	Fid Data Display Type . . . . .	56
3.15	Fid Data Options Menu . . . . .	58
3.16	The Image Viewer . . . . .	60
3.17	The Image Viewer Right Mouse Popup Menu . . . . .	61
3.18	The Prior Probability Viewer . . . . .	66
3.19	The Fid Model Viewer . . . . .	69
3.20	The Plot Results Viewer . . . . .	72
3.21	Plot Information Popup . . . . .	73
3.22	The Text Results Viewer . . . . .	75
3.23	The Bayes Condensed File . . . . .	78
3.24	Data, Model, And Resid Plot . . . . .	81
3.25	The Parameter Posterior Probabilities . . . . .	82
3.26	The Maximum Entropy Histograms . . . . .	84
3.27	The Parameter Samples Plot . . . . .	85
3.28	Posterior Probability Vs Parameter Value . . . . .	86
3.29	Posterior Probability Vs Parameter Value, A Skewed Example . . . . .	87
3.30	The Expected Value Of The Logarithm Of The Likelihood . . . . .	89

3.31	The Scatter Plots . . . . .	90
3.32	The Logarithm Of The Posterior Probability By Repeat Plot . . . . .	92
3.33	The Fortran/C Model Viewer . . . . .	93
3.34	The Fortran/C Code Editor . . . . .	95
4.1	Frequency Estimation Using The DFT . . . . .	112
4.2	Aliases . . . . .	113
4.3	Nonuniformly Nonsimultaneously Sampled Sinusoid . . . . .	127
4.4	Alias Spacing . . . . .	128
4.5	Which Is The Critical Time . . . . .	130
4.6	Example, Frequency Estimation . . . . .	131
4.7	Estimating The Sinusoids Parameters . . . . .	133
5.1	The Given And Unknown Number Of Exponential Package Interface . . . . .	138
6.1	The Unknown Exponential Interface . . . . .	144
6.2	The Distribution Of Models . . . . .	149
6.3	The Posterior Probability For Exponential Model . . . . .	150
7.1	The Inversion Recovery Interface . . . . .	152
8.1	Bayes Analyze Interface . . . . .	156
8.2	Bayes Analyze Fid Model Viewer . . . . .	160
8.3	The Bayes Analyze File Header . . . . .	179
8.4	The bayes.noise File . . . . .	180
8.5	Bayes Analyze Global Parameters . . . . .	183
8.6	The Third Section Of The Parameter File . . . . .	184
8.7	Example Of An Initial Model In The Output File . . . . .	187
8.8	Base 10 Logarithm Of The Odds . . . . .	187
8.9	A Small Sample Of The Output Report . . . . .	188
8.10	Bayes Analyze Uncorrelated Output . . . . .	189
8.11	The bayes.proBABILITIES.nnnn File . . . . .	191
8.12	The bayes.log.nnnn File . . . . .	193
8.13	The bayes.status.nnnn File . . . . .	196
8.14	The bayes.model.nnnn File . . . . .	197
8.15	The bayes.model.nnnn File Uncorrelated Resonances . . . . .	198
8.16	Bayes Analyze Summary Header . . . . .	198
8.17	The Summary2 (Best Summary) . . . . .	199
8.18	The Summary3 Report . . . . .	201
9.1	The Big Peak/Little Peak Interface . . . . .	208
9.2	The Time Dependent Parameters . . . . .	218
10.1	The Bayes Metabolite Interface . . . . .	220
10.2	The Bayes Metabolite Viewer . . . . .	222
10.3	Bayes Metabolite Parameters And Probabilities List . . . . .	227
10.4	The IPGD_D20 Metabolite . . . . .	229

10.5	Bayes Metabolite IPGD_D20 Spectrum . . . . .	230
10.6	Bayes Metabolite, The Fraction of Glucose . . . . .	231
10.7	Glutamate Example Spectrum . . . . .	233
10.8	Estimating The $F_{c0}$ , $y$ and $F_{a0}$ Parameters . . . . .	236
10.9	Bayes Metabolite, The Ethyl Ether Example . . . . .	237
11.1	The Find Resonances Interface With The Ethyl Ether Spectrum . . . . .	240
12.1	The Diffusion Tensor Package Interface . . . . .	248
12.2	Diffusion Tensor Parameter Estimates . . . . .	256
12.3	Diffusion Tensor Posterior Probability For The Model . . . . .	257
13.1	The Big Magnetization Package Interface . . . . .	260
13.2	Big Magnetization Transfer Example Fid . . . . .	263
13.3	Big Magnetization Transfer Expansion . . . . .	263
13.4	Big Magnetization Transfer Peak Pick . . . . .	264
14.1	The Magnetization Transfer Package Interface . . . . .	266
14.2	Magnetization Transfer Package Peak Picking . . . . .	272
14.3	Magnetization Transfer Example Data . . . . .	273
14.4	Magnetization Transfer Example Spectrum . . . . .	274
15.1	Magnetization Transfer Kinetics Package Interface . . . . .	276
15.2	Magnetization Transfer Kinetics Package Arrhenius Plot . . . . .	282
15.3	Magnetization Transfer Kinetics Water Viscosity Table . . . . .	283
16.1	Given Polynomial Order Package Interface . . . . .	286
16.2	Given Polynomial Order Scatter Plot . . . . .	291
17.1	Unknown Polynomial Order Package Interface . . . . .	294
17.2	The Distribution of Models On The Console Log . . . . .	298
17.3	The Posterior Probability For The Polynomial Order . . . . .	300
18.1	The Errors In Variables Package Interface . . . . .	304
18.2	The McMC Values File Produced By The Errors In Variables Package . . . . .	310
19.1	The Behrens-Fisher Interface . . . . .	312
19.2	Behrens-Fisher Hypotheses Tested . . . . .	313
19.3	Behrens-Fisher Console Log . . . . .	323
19.4	Behrens-Fisher Status Listing . . . . .	324
19.5	Behrens-Fisher McMC Values File, The Preamble . . . . .	325
19.6	Behrens-Fisher McMC Values File, The Middle . . . . .	326
19.7	Behrens-Fisher McMC Values File, The End . . . . .	327
20.1	Enter Ascii Model Package Interface . . . . .	330
21.1	The Enter Ascii Model Selection Package Interface . . . . .	338

26.1	Absorption Model Images . . . . .	396
26.2	The Interface To The Image Phasing Package . . . . .	397
26.3	Linear Phasing Package The Console Log . . . . .	403
27.1	Nonlinear Phasing Example . . . . .	406
27.2	The Interface To The Nonlinear Phasing Package . . . . .	410
28.1	The Interface To The Analyze Image Pixels Package . . . . .	412
29.1	The Interface To The Image Model Selection Package . . . . .	416
29.2	Single Exponential Example Image . . . . .	419
29.3	Single Exponential Example Data . . . . .	420
29.4	Posterior Probability For The ExpOneNoConst Model . . . . .	421
A.1	Ascii Data File Format . . . . .	424
D.1	The McMC Values Report Header . . . . .	450
D.2	McMC Values Report, The Middle . . . . .	451
D.3	The McMC Values Report, The End . . . . .	452
E.1	Writing Models A Fortran Example . . . . .	456
E.2	Writing Models A C Example . . . . .	457
E.3	Writing Models, The Parameter File . . . . .	459
E.4	Writing Models Fortran Declarations . . . . .	463
E.5	Writing Models Fortran Example . . . . .	466
E.6	Writing Models The Parameter File . . . . .	467
G.1	Example FDF File Header . . . . .	473
H.1	The Posterior Probability For The Number of Outliers . . . . .	476
H.2	The Data, Model and Residual Plot With Outliers . . . . .	478

# List of Tables

8.1	Multiplet Relative Amplitudes . . . . .	165
8.2	Bayes Analyze Models . . . . .	181
8.3	Bayes Analyze Short Descriptions . . . . .	195

## Chapter 21

# Enter Ascii Model Selection

The Enter Ascii Model Selection Package allows you to load one or more Ascii model and then use Bayesian probability theory to compute the posterior probability for the loaded model, thus allowing you to determine which model best accounts for the data.<sup>1</sup> To use this package you do not have to have Fortran or C installed on your server. However, if you do not have Fortran or C installed, you must use the system models. Consequently, installing both Fortran and C is strongly recommended. The interface to this package is shown in Fig. 21. To use this package, you must do the following:

**Select** the “Enter Ascii Model Selection” package from the Package menu.

**Load** one or more Fortran or C model using the “System” or “User” buttons in the “Load And Build Model” widget group.

**Load** one or more Ascii data sets using the Files menu. When a data set is successfully loaded the data is plotted in the Ascii Data viewer. The format of the Ascii data that must be loaded is dependent on the model. Usually the data are two column Ascii, however, in general this package takes multicolumn Ascii data with a multicolumn abscissas. See Appendix A for a detailed description of the Ascii data files used by the Bayesian Analysis software.

**Build** the model using the “Build” button would normally be the next step. However, in this package it is assumed that you have previously compiled and tested each Ascii model, so this package does not have a “Build” button.

**Check** the Find Outliers box if you suspect outliers are present in the data.

**Review** the prior probabilities for the loaded model using the Prior Viewer would normally be the next step. However, because multiple models are loaded and thus multiple parameter sets are loaded, we require the user to test his models prior to running this package. So there is no review of the prior probabilities in this package. You can test your models using either the Enter Ascii Model Package or the Test Ascii Model package and you can review and update the prior information using those packages.

---

<sup>1</sup>I would like to build a system library of predefined models. If you have models that you think would be of general use, I would like to hear from you. To have one of your models included, I would need the source code, the parameter file, a brief description of the model equations and data requirements.

Figure 21.1: The Enter Ascii Model Selection Package Interface

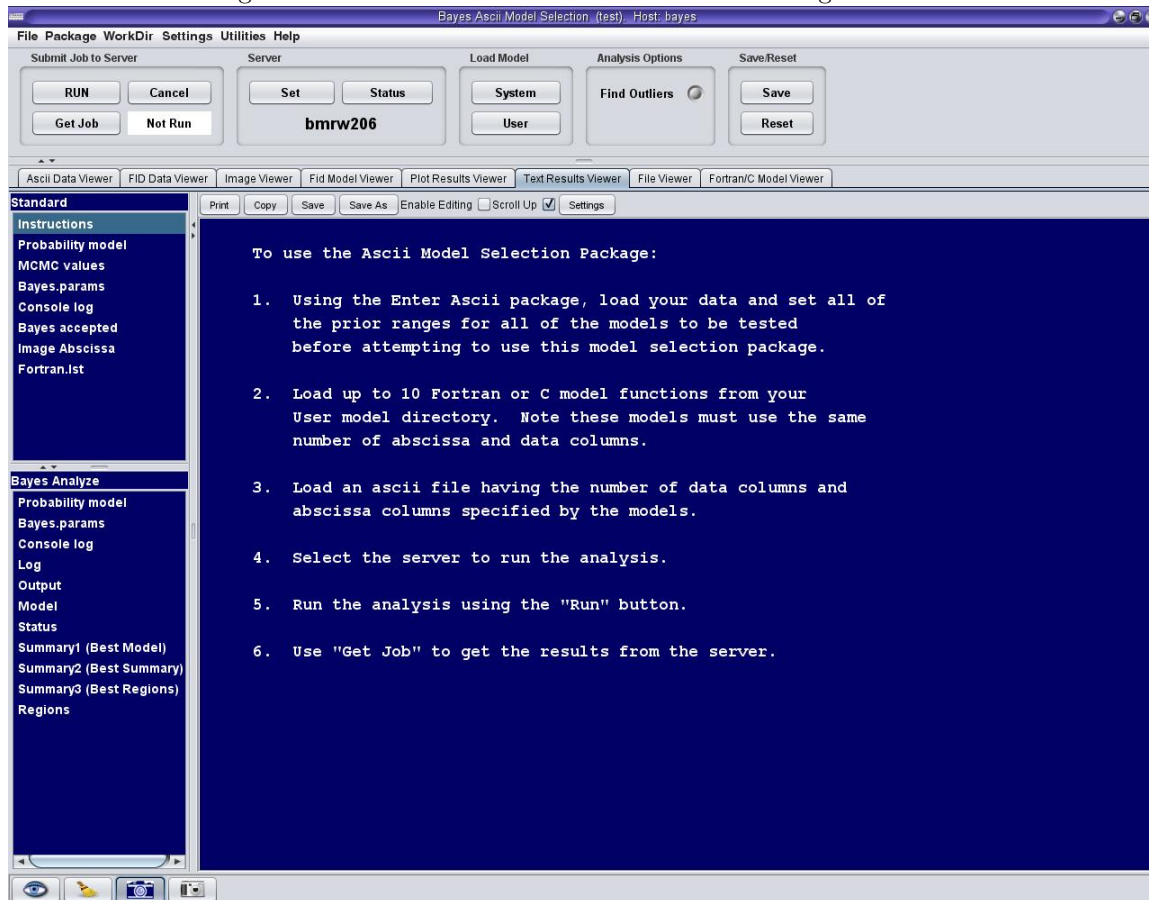


Figure 21.1: This is the interface to the Enter Ascii Model Selection package. This package allows you to load multiple Fortran or C models and then run the the model selection calculation using these models. The program will select the best model from within the set of loaded models. For more on the actual calculations and the widgets see the text.

**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.

## 21.1 The Bayesian Calculations

In the model selection calculation done by the Enter Ascii Model Selection package, it assumes one has a set of models,  $U_j \equiv \{U_1, \dots, U_m\}$ , and one wishes to compute the posterior probability for each of the  $U_j$  models. These models can be loaded from the System directory or they can be loaded from the user directory. They don't need to have common parameters, but they do need to have common data requirements. Each of these models will have a set of parameters associated it and these parameters will be designated as  $\Omega_j$ . The subscript  $j$  indicating that these are the parameters associated with model  $U_j$ . The equation that relates model  $U_j$  to the data is given by

$$d_k(t_i) = U_j(t_i, \Omega_j) + n_k(t_i) \tag{21.1}$$

where  $d_k(t_i)$  represents a data item in the  $k$ th data set, sampled at abscissa value  $t_i$ . The  $t_i$  may be vector valued or it may be a single column of numbers. However, it must have the same number of columns in all loaded models. The noise is represented symbolically by  $n_k(t_i)$  and is the noise value in the  $k$ th data set sampled at abscissa value  $t_i$ .

The posterior probability for each model  $U_j$  is given by Bayes' theorem:

$$P(U_j|DI) = \frac{P(U_j|I)P(D|U_jI)}{P(D|I)} \tag{21.2}$$

where  $P(U_j|DI)$  is the posterior probability for model  $U_j$  given the data and the prior information,  $P(U_j|I)$  is the prior probability for model  $U_j$  given only the prior information,  $P(D|U_jI)$  is the marginal direct probability for all of the data given the model  $U_j$  and the prior information. This term was called a direct probability because it is a probability for the data and it is a marginal probability because none of the parameters from model  $U_j$  appear, so they have been marginalized out.  $P(D|I)$  is a normalization constant that ensures the probabilities sum to one:

$$\sum_{j=1}^m P(U_j|DI) = 1, \tag{21.3}$$

so

$$\sum_{j=1}^m \frac{P(U_j|I)P(D|U_jI)}{P(D|I)} = 1, \tag{21.4}$$

and

$$\sum_{j=1}^m P(U_j|I)P(D|U_jI) = P(D|I). \tag{21.5}$$



Equation (21.2) cannot be computed in this form because the marginal direct probability,  $P(D|U_j I)$ , cannot be assigned. However, the prior probability for the models,  $P(U_j|I)$  can be assigned and in this calculation it is assigned as a uniform prior:

$$P(U_j|I) = \frac{1}{m}. \quad (21.6)$$

The marginal direct probability for the data given the model,  $P(D|U_j I)$ , can be computed if we reintroduce the model parameters  $\Omega_j$  and then use the sum rule of probability theory to marginalize out these parameters

$$P(D|U_j I) = \int P(D\Omega_j|U_j I) d\Omega_j. \quad (21.7)$$

The probability on the right-hand side of this equation is the joint prior probability for the data and the  $\Omega_j$  parameters given the model  $U_j$  and the prior information  $I$ . Applying the product rule to factor the right-hand side of this equation, one obtains

$$P(D|U_j I) = \int P(\Omega_j|I) P(D|\Omega_j U_j I) d\Omega_j \quad (21.8)$$

where  $P(D|\Omega_j U_j I)$  is the direct probability for the data given the parameters and the model, this direct probability is also called a likelihood and  $P(\Omega_j|I)$  is the joint prior probability for the  $\Omega_j$  parameters.

In the Bayesian Analysis software that implements this calculation, there are two types of Ascii models: those that do not use marginalization to remove the amplitudes and those that do. The functional form of the Ascii model is very different between these two types of Fortran/C codes. Consequently, the calculation for the marginal direct probability for the data given the model,  $P(D|U_j I)$ , must be split into two separate calculations: one that uses amplitude marginalization and one that does not.

### 21.1.1 The Direct Probability With No Amplitude Marginalization

In this subsection it will be assumed that the input model is one that has not defined any amplitudes and consequently there is no amplitude marginalization in the posterior probability for the parameters. So in this calculation the  $\Omega_j$  parameters are given by  $\Omega_j \in \{\omega_{j1}, \dots, \omega_{j\nu_j}\}$  where  $\nu_j$  is the number of parameters. The individual Ascii model parameters are designated by  $\omega_{jk}$  and these parameters may include amplitudes, but if it does, the program that implements the calculation has no knowledge of them. Assuming logical independence of the parameters, i.e., knowing the value of one parameter would not help us in making inferences about the other parameters, then the joint prior probability for the  $\Omega_j$  parameters can be factored into an independents prior probability for each parameter,

$$P(\Omega_j|I) = \prod_{k=1}^{\nu_j} P(\omega_{jk}|\Omega_j I) \quad (21.9)$$

where  $P(\omega_{jk}|\Omega_j I)$  is the prior probability for the  $k$  parameter in the  $j$ th model. The model  $\Omega_j$  was include in the prior information because what prior we assign to the parameters will definitely be dependent on the model.

The likelihood,  $P(D|\Omega_j U_j I)$ , is the likelihood function for all of the data  $D$ . However, the data  $D$  is made up of multiple data sets,  $D \equiv \{D_1, \dots, D_n\}$ , where  $n$  is the input number of data sets,

and each data set consists of  $N_j$  data values, so  $D_j \equiv \{d_j(t_1), \dots, d_j(t_{N_j})\}$ , where  $d_j(t_i)$  is a data item in the  $j$ th data set sampled at abscissa value  $t_i$ . The number of data values in a given data set,  $N_j$ , need not be the same from one data set to another. Assuming logical independence of the various data sets, the joint direct probability for the data,  $P(D|\Omega U_j I)$ , can be written as

$$P(D|\Omega_j U_j I) = \prod_{k=1}^n P(D_k|\Omega_j U_j I). \quad (21.10)$$

Each of the probabilities,  $P(D_k|\Omega_j U_j I)$ , is a direct probability or likelihood for the data in the  $k$ th data set given the model, the parameters, and the prior information. Such direct probabilities are usually assigned using a Gaussian prior probability for the noise, then the likelihood for a single data set becomes

$$P(D_k|\sigma_k \Omega_j U_j I) = (2\pi\sigma_k^2)^{-\frac{N_k}{2}} \exp\left\{-\frac{Q_{jk}^2}{2\sigma_k^2}\right\} \quad (21.11)$$

where the standard deviations have been added to  $P(D_k|\Omega_j U_j I)$  because the right-hand side of this equation cannot be computed unless the standard deviations are known. The total squared residual,  $Q_{jk}^2$ , is defined as:

$$Q_{jk}^2 \equiv \sum_{i=1}^{N_k} [d_k(t_i) - U_j(t_i, \Omega_j)]^2 \quad (21.12)$$

in the  $k$ th data set given the  $j$ th model and its parameters  $\Omega_j$ . Substituting Eq. (21.11) into Eq. (21.10) one obtains

$$P(D|\sigma_1 \cdots \sigma_n \Omega_j U_j I) = \prod_{k=1}^n (2\pi\sigma_k^2)^{-\frac{N_k}{2}} \exp\left\{-\frac{Q_{jk}^2}{2\sigma_k^2}\right\} \quad (21.13)$$

as the direct probability for all of the data, where we have modified the notation to indicate that all of the  $\sigma_j$  must be given. Finally, substituting Eq. (21.13), Eq. (21.9) and Eq. (21.6) into Eq. (21.2) one obtains

$$P(U_j|\sigma_1 \cdots \sigma_n DI) \propto \int \frac{1}{m} \left[ \prod_{k=1}^{\nu_j} P(\omega_{jk}|I) \right] \left[ \prod_{k=1}^n (2\pi\sigma_k^2)^{-\frac{N_k}{2}} \exp\left\{-\frac{Q_{jk}^2}{2\sigma_k^2}\right\} \right] d\Omega_j \quad (21.14)$$

as the marginal posterior probability for model  $U_j$ . In the package that processes this analysis, the prior probabilities are specified by the user. And because the prior probability for the model was assigned as uniform, it will cancel when this probability density function is normalized. A Markov chain Monte Carlo simulation is used to numerically integrate this equation and thus obtain the posterior probability for the models.

Often the standard deviation of the noise prior probability are not known, so Eq. (21.14) cannot be used. When the standard deviation are not known, one can remove the standard deviations using the rules of probability theory:

$$P(U_j|DI) = \int P(\sigma_1 \cdots \sigma_n U_j|DI) d\sigma_1 \cdots d\sigma_n d\Omega_j. \quad (21.15)$$

Factoring the right-hand side using the product rule of probability theory, one obtains

$$P(U_j|DI) = \int P(\sigma_1 \cdots \sigma_n | I) P(U_j | \sigma_1 \cdots \sigma_n DI) d\sigma_1 \cdots d\sigma_n d\Omega_j. \quad (21.16)$$

Assuming logical independence, the joint prior probability for the standard deviations can be factored to obtain:

$$P(U_j|DI) = \int \left[ \prod_{k=1}^n P(\sigma_k | I) \right] P(U_j | \sigma_1 \cdots \sigma_n DI) d\sigma_1 \cdots d\sigma_n d\Omega_j. \quad (21.17)$$

Substituting Eq. (21.14) into Eq. (21.17) one obtains

$$P(U_j|DI) \propto \frac{1}{m} \int \left[ \prod_{k=1}^{\nu_j} P(\omega_{jk} | I) \right] \left[ \prod_{k=1}^n (2\pi\sigma_k)^{-\frac{N_k+1}{2}} \exp \left\{ -\frac{Q_{jk}^2}{2\sigma_k^2} \right\} \right] d\sigma_1 \cdots d\sigma_n d\Omega_j \quad (21.18)$$

as the posterior probability for model  $U_j$  given the data and the prior information. The integrals of the standard deviations of the noise prior probability are gamma function integrals. We address those integrals in the next section when we consider the same calculation but with marginalization over the amplitudes. Here we are just going to give the result of marginalizing over the  $\sigma_k$ :

$$P(U_j|DI) \propto \frac{1}{m} \int \left[ \prod_{k=1}^{\nu_j} P(\omega_{jk} | I) \right] \prod_{k=1}^n \left[ \frac{1}{2} \Gamma \left( \frac{N_k}{2} \right) Q_{jk}^{-\frac{N_k}{2}} \right] d\Omega_j. \quad (21.19)$$

For the details of how this integration is performed see subsection 21.1.2.2. When the standard deviation for the noise prior probability is known Eq. (21.18) is used in the numerical simulation that implements this calculation. Knowledge of the standard deviation of the noise prior probability usually comes about in image processing applications because in those applications it is possible to directly sample the noise. However, more often than not, no information is available about the standard deviation of the noise prior probability, then the posterior probability for the model independent of the standard deviations of the noise prior probability, Eq. (21.19), is used. The functional form of this equation is that of a Student's  $t$ -distribution. The Markov chain Monte Carlo simulation that is used to evaluate the remaining integrals targets either Eq. (21.18) or Eq. (21.19) using a Monte Carlo simulation to draw samples from this posterior probability and then uses Monte Carlo integration to evaluate the posterior probability for the models.

### 21.1.2 The Direct Probability With Amplitude Marginalization

The above calculation assumes the amplitudes are not analytically marginalized out. However, the Bayesian Analysis software uses two types of models. One that does not explicitly marginalize out the amplitudes and one that does. The functional form of these models that marginalize out the amplitudes is a special subset of the form assumed above. The models that do not marginalize out the amplitudes may contain amplitudes, but they are not identified as such in the parameter file associated with the Fortran/C model; rather the amplitudes are just lumped in with the other nonlinear parameters. However, when the amplitudes are marginalized out, the amplitudes are made explicit in the Fortran/C model parameter file and the functional form of the Fortran/C model is different. Consequently, the calculation for the marginal posterior probability for the model is a bit different for these models.

The signal model used in the calculation when the amplitudes are marginalized out, starts out similar to the calculation where no marginalization occurs. The data are related to the  $j$ th model  $U_j$  by

$$d_k(t_i) = U_j(t_i, \mathcal{B}_j, \Omega_j) + n_k(t_i) \tag{21.20}$$

where  $d_k(t_i)$  represents a data item in the  $k$ th data set that was sampled at abscissa value  $t_i$ ,  $n_k(t_i)$  represents the noise in the  $k$ th data set sampled at abscissa value  $t_i$ ,  $\mathcal{B}_j$  is the set of all amplitudes appearing in the  $j$ th model, and  $\Omega_j$  is the set of all nonamplitude parameters,  $\Omega_j \in \{\omega_{j1}, \dots, \omega_{j\nu_j}\}$  and  $\nu_j$  is the number of nonamplitude parameters in the  $j$ th data set. These parameters will be referred to as the nonlinear parameters because they almost always appear in the model in a nonlinear fashion. The model,  $U_j \in \{U_1, \dots, U_m\}$ , is one from the set of models that are being analyzed. When the amplitudes are marginalized out, it is assumed that the  $U_j(t_i, \mathcal{B}_j, \Omega_j)$  have the following functional form:

$$U_j(t_i, \mathcal{B}_j, \Omega_j) = \sum_{\ell=1}^{u_j} B_{jk\ell} G_{j\ell}(t_i, \Omega_j) \tag{21.21}$$

where  $B_{jk\ell}$  is the  $\ell$ th amplitude of the  $j$ th model in the  $k$ th data set, and  $G_{j\ell}(t_i, \Omega_j)$  is the  $\ell$ th subcomponent of the  $j$ th model. Note that the  $G_{j\ell}(t_i, \Omega_j)$  do not depend on the amplitudes in any way; and in this model the only dependence on the amplitudes is a linear dependence. To give an example of this, suppose we are estimating a sum of  $u_j$  sinusoids, the model  $U_j(t_i, \mathcal{B}_j, \Omega_j)$  would be given by

$$U_j(t_i, \mathcal{B}_j, \Omega_j) = \sum_{\ell=1}^{u_j} B_{jk\ell} \cos(2\pi\omega_{j\ell}t_i + \theta_{j\ell}) \tag{21.22}$$

and each submodel component, the  $G_{j\ell}(t_i, \Omega_j)$  are given by

$$G_{j\ell}(t_i, \Omega_j) = \cos(2\pi\omega_{j\ell}t + \theta_{j\ell}). \tag{21.23}$$

In this sinusoidal model, the sinusoids depend on the data sets through the amplitude and phase while the frequencies are common to all data sets.

### 21.1.2.1 Marginalizing the Amplitudes

To compute the posterior probability for the model when the amplitudes are marginalized out, we note that that the calculation begins where the previous calculation ends with Eq. (21.14). Rewriting Eq. (21.14) so as to separate out the amplitudes integrals from the integrals over the nonlinear parameters, one obtains

$$P(U_j|DI) \propto \frac{1}{m} \int \left[ \prod_{\ell=1}^{\nu_j} P(\omega_{j\ell}|I) \right] \left[ \prod_{k=1}^n \left( \prod_{z=1}^{u_j} P(B_{jkz}|I) \right) (2\pi\sigma_k)^{-\frac{N_k}{2}} \exp \left\{ -\frac{Q_{jk}^2}{2\sigma_k^2} \right\} \right] dB_j d\Omega_j \tag{21.24}$$

where  $u_j$  is the number of amplitudes in the  $j$ th model. The prior probability for the  $l$ th nonlinear parameter in the  $j$ th model is designated by  $P(\omega_{jl}|I)$ . Finally,  $P(B_{jkz}|I)$  is the prior probability for the  $z$ th amplitude in the  $j$ th model of the  $k$ th data set. Substituting the model, Eq. (21.21), into

the definition of  $Q_{jk}^2$ , Eq. (21.12), we can obtain

$$\begin{aligned} Q_{jk}^2 &\equiv \sum_{i=1}^{N_k} [d_k(t_i) - U_k(t_i, \mathcal{B}_j, \Omega_j)]^2 \\ &= \sum_{i=1}^{N_k} \left[ d_k(t_i) - \sum_{\ell=1}^{u_j} B_{jk\ell} G_{j\ell}(t_i, \Omega_j) \right]^2. \end{aligned} \quad (21.25)$$

Multiplying out the brackets one obtains

$$Q_{jk}^2 = d_k(t_i) \cdot d_k(t_i) - 2 \sum_{\ell=1}^{u_j} B_{jk\ell} d_k(t_i) \cdot G_{j\ell}(t_i) + \sum_{\ell=1}^{u_j} \sum_{v=1}^{u_j} B_{jk\ell} B_{jkv} G_{j\ell}(t_i) \cdot G_{jv}(t_i) \quad (21.26)$$

where we are using the “.” notation to mean sum over time, so

$$d_k(t_i) \cdot G_{j\ell} = \sum_{i=1}^{N_k} d_k(t_i) G_{j\ell}(t_i). \quad (21.27)$$

And to compress the notation still further, we define

$$d_k^2 = d_k(t_i) \cdot d_k(t_i), \quad (21.28)$$

$$T_{jk\ell} = d_k(t_i) \cdot G_{j\ell}(t_i) \quad (21.29)$$

and

$$g_{j\ell v} = G_{j\ell}(t_i) \cdot G_{jv}(t_i) \quad (21.30)$$

so Eq. (21.26) becomes

$$Q_{jk}^2 = d_k^2 - 2 \sum_{\ell=1}^{u_j} B_{jk\ell} T_{jk\ell} + \sum_{\ell=1}^{u_j} \sum_{v=1}^{u_j} B_{jk\ell} B_{jkv} g_{j\ell v}. \quad (21.31)$$

To evaluate the integrals in Eq. (21.24), we must assign a prior probability for the amplitudes. A zero-mean Gaussian prior probability will be assigned for each amplitudes. This Gaussian prior probability is given by

$$P(B_{jku}|I) \propto \left( \frac{2\pi\sigma_j^2}{\gamma^2 g_{juu}} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{B_{jku}^2 \gamma^2 g_{juu}}{2\sigma_j^2} \right\} \quad (21.32)$$

where  $\gamma$  is used to control the width of this prior probability,  $\sigma_j$  is the standard deviation of the noise prior probability in the  $j$ th data set, and  $g_{jku}$  was defined in Eq. (21.30). The reason for this particular form is that it allows one to evaluate the integrals over the amplitudes in a concise functional form that aids in doing the numerical calculations.

Substituting Eq. (21.32) and Eq. (21.31) into the posterior probability for the model, Eq. (21.24), one obtains

$$\begin{aligned}
 P(U_j|DI) &= \frac{1}{m} \int \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \\
 &\times \prod_{k=1}^n \left[ \frac{1}{\sigma_k} \prod_{\beta=1}^{u_{jk}} \left( \frac{2\pi\sigma_k^2}{\gamma^2 g_{j\beta\beta}} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{B_{jk\beta}^2 \gamma^2 g_{j\beta\beta}}{2\sigma_k^2} \right\} \right] \\
 &\times \left( 2\pi\sigma_k^2 \right)^{-\frac{N_k}{2}} \exp \left\{ -\frac{1}{2\sigma_k^2} \left[ d_k(t_i)^2 - 2 \sum_{\ell=1}^{u_j} B_{jk\ell} T_{jk\ell} + \sum_{\ell=1}^{u_j} \sum_{v=1}^{u_j} B_{jk\ell} B_{jkv} g_{j\ell v} \right] \right\} \\
 &\times dB_{jk1} \cdots dB_{jk u_j} d\sigma_k d\Omega_j
 \end{aligned} \tag{21.33}$$

as the posterior probability for model  $U_j$ . Expanding the product over  $\beta$ , and simplifying this equation somewhat, one obtains

$$\begin{aligned}
 P(U_j|DI) &\propto \int \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \prod_{k=1}^n \left[ \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k+u_j}{2}} \gamma^{u_j} g_{j11} \cdots g_{j u_j u_j} \right. \\
 &\times \left. \exp \left\{ -\frac{1}{2\sigma_k^2} \left[ d_k(t_i)^2 - 2 \sum_{\ell=1}^{u_j} B_{jk\ell} T_{jk\ell} + \sum_{\ell=1}^{u_j} \sum_{v=1}^{u_j} B_{jk\ell} B_{jkv} g'_{j\ell v} \right] \right\} \right] \\
 &\times dB_{jk1} \cdots dB_{jk u_j} d\sigma_1 \cdots d\sigma_n d\Omega_j
 \end{aligned} \tag{21.34}$$

where

$$g'_{j\ell v} = g_{j\ell v} (1 + \delta_{\ell v} \gamma^2) \tag{21.35}$$

and  $\delta_{\ell v}$  is the Kronecker delta function. The Kronecker delta function is zero if  $\ell \neq v$  and its one if  $\ell = v$ . In order to evaluate the amplitude integrals a change of variables and functions will be made. The  $g'_{j\ell v}$  matrix is positive definite and of rank  $u_j$ . Let  $e_{\ell v}$  represent the  $v$ th component of the  $\ell$ th normalized eigenvector of  $g'_{j\ell v}$ , then

$$\sum_{v=1}^{u_j} g'_{j\beta v} e_{\ell v} = \lambda_{\ell} e_{\ell\beta} \tag{21.36}$$

where  $\lambda_{\ell}$  is the  $\ell$ th eigenvalue of  $g_{j\ell v}$ . The functions  $H_{j\ell}(t_i)$ , defined as

$$H_{j\ell}(t_i) = \frac{1}{\sqrt{\lambda_{\ell}}} \sum_{v=1}^{u_j} e_{\ell v} G_{jv}(t_i), \tag{21.37}$$

are orthonormal, i.e.,

$$\sum_{i=1}^{N_j} H_{j\ell}(t_i) H_{jv}(t_i) = \delta_{\ell v} \tag{21.38}$$

where  $\delta_{\ell v}$  is the Kronecker delta function and it is these  $H_j$  functions that will be used in the calculation. The model Eq. (21.21) can be rewritten in terms of these orthonormal functions as

$$U_j(t_i, \mathcal{A}_j, \Omega_j) = \sum_{\ell=1}^{u_j} A_{jk\ell} H_{j\ell}(t_i) \tag{21.39}$$

and it is these  $A_{jk\ell}$  that will be used in a change of variables. The amplitudes  $B_{jk\ell}$  are linearly related to the  $A_{jk\ell}$  by

$$B_{jk\ell} = \sum_{\beta=1}^{u_j} \frac{A_{jk\beta} e_{\beta\ell}}{\sqrt{\lambda_\beta}} \quad \text{and} \quad A_{jk\ell} = \sqrt{\lambda_\ell} \sum_{\beta=1}^{u_\beta} B_{jk\beta} e_{\ell\beta}. \quad (21.40)$$

The volume elements in the  $k$ th data set is given by

$$\begin{aligned} dB_{jk_1} \cdots dB_{jk_{u_j}} &= \left| \frac{e_{\ell v}}{\sqrt{\lambda_v}} \right| dA_{jk_1} \cdots dA_{jk_{u_j}} \\ &= \lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}} dA_{jk_1} \cdots dA_{jk_{u_j}}. \end{aligned} \quad (21.41)$$

The Jacobin is a function of the  $\Omega_j$  parameters but it is a constant as long as we are not integrating over the  $\Omega_j$  parameters. At the end of the calculation the linear relations between the  $A$ 's and  $B$ 's can be used to calculate the expected values of the  $B$ 's from the expected value of the  $A$ 's,

$$E(B_{jk\ell} | \Omega_j DI) = \langle B_{jk\ell} \rangle = \sum_{\beta=1}^{u_j} \frac{\langle A_{jk\beta} \rangle e_{\beta\ell}}{\sqrt{\lambda_\beta}} \quad (21.42)$$

and the same is true of the second posterior moments:

$$E(B_{jk\ell} B_{jkl} | \Omega_j DI) = \langle B_{jk\ell} B_{jkl} \rangle = \sum_{\beta=1}^{u_j} \sum_{v=1}^{u_j} \frac{e_{\beta\ell} e_{v\ell} \langle A_{jk\beta} A_{jkv} \rangle}{\sqrt{\lambda_\beta \lambda_v}} \quad (21.43)$$

where  $E(B_{jk\ell} | \Omega_j DI)$  means the expected value of the  $\ell$ th amplitude  $B_{jk\ell}$  in the  $j$ th model of the  $k$ th data set given the nonlinear  $\Omega_j$  parameters, the data  $D$  and the prior information  $I$ . Substituting, Eq. (21.40) into Eq. (21.34) and using the definition of the  $H_{jl}$  functions, Eq. (21.37) one can rewrite the posterior probability for the model as:

$$\begin{aligned} P(U_j | DI) &\propto \int \left[ \prod_{l=1}^{v_{jk}} P(\omega_{jkl} | I) \right] \prod_{k=1}^n \left[ \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k+u_j}{2}} \gamma^{u_j} \frac{g_{j11} \cdots g_{ju_j u_j}}{\lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}}} \right. \\ &\times \exp \left\{ -\frac{1}{2\sigma_k^2} \left( d_k^2 - 2 \sum_{\ell=1}^{u_j} A_{jk\ell} h_{jkl} + \sum_{v=1}^{u_j} A_{jv\ell}^2 \right) \right\} \Big] \\ &\times dA_{jk_1} \cdots dA_{jk_{u_j}} d\sigma_1 \cdots d\sigma_n d\Omega_j \end{aligned} \quad (21.44)$$

where

$$h_{jkl} = \sum_{i=1}^{N_j} d_{jk}(t_i) H_{j\ell}(t_i). \quad (21.45)$$

Completing the square in Eq. (21.44), one obtains:

$$\begin{aligned} P(U_j | DI) &\propto \int \left[ \prod_{l=1}^{v_{jk}} P(\omega_{jkl} | I) \right] \prod_{k=1}^n \left[ \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k+u_j}{2}} \gamma^{u_j} \frac{g_{j11} \cdots g_{ju_j u_j}}{\lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}}} \right. \\ &\times \exp \left\{ -\frac{1}{2\sigma_k^2} \left[ d_k^2 - h_{jkl}^2 + \sum_{\ell=1}^{u_j} (A_{jk\ell} - h_{jkl})^2 \right] \right\} \Big] \\ &\times dA_{jk_1} \cdots dA_{jk_{u_j}} d\sigma_1 \cdots d\sigma_n d\Omega_j. \end{aligned} \quad (21.46)$$

Designating the integral over the amplitudes as  $I_j$  and isolating it, Eq. (21.46) can be written as:

$$\begin{aligned}
 P(U_j|DI) &\propto \int \prod_{k=1}^n \left[ \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k+u_j}{2}} \gamma^{u_j} \frac{g_{j11} \cdots g_{ju_j u_j}}{\lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}}} \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \right. \\
 &\quad \left. \times \exp \left\{ \frac{1}{2\sigma_k^2} [d_k^2 - h_{jk\ell}^2] \right\} \times I_j d\sigma_1 \cdots d\sigma_n d\Omega_j \right]
 \end{aligned} \tag{21.47}$$

where the amplitude integral is given by:

$$I_j \equiv \prod_{k=1}^n \int \exp \left\{ -\frac{1}{2\sigma_k^2} \sum_{\ell=1}^{u_j} (A_{jk\ell} - h_{jk\ell})^2 \right\} dA_{jk1} \cdots dA_{jk u_j}. \tag{21.48}$$

Introducing a change of variables

$$z_{jk\ell} = \frac{1}{\sqrt{2}\sigma_k} (A_{jk\ell} - h_{jk\ell}) \tag{21.49}$$

with volume element given by

$$\sqrt{2}\sigma_k dz_{jk\ell} = dA_{jk\ell} \tag{21.50}$$

the integral  $I_j$  becomes

$$\begin{aligned}
 I_j &\equiv \int_{-\infty}^{\infty} \exp \left\{ -\sum_{\ell=1}^{u_j} z_{jk\ell}^2 \right\} \sqrt{2}\sigma_k dz_{jk1} \cdots \sqrt{2}\sigma_k dz_{jk u_j} \\
 &= (\sqrt{2}\sigma_k)^{u_j} \int_{-\infty}^{\infty} \exp \left\{ -\sum_{\ell=1}^{z_j} z_{jk\ell}^2 \right\} dz_{jk1} \cdots dz_{jk u_j} \\
 &= (\sqrt{2}\sigma_k)^{u_j} (\sqrt{\pi})^{u_j} \\
 &= (\sqrt{2}\sigma_k \sqrt{\pi})^{u_j} \\
 &= (2\pi\sigma_k^2)^{\frac{u_j}{2}}.
 \end{aligned} \tag{21.51}$$

Inserting  $I_j$  back into Eq. (21.47), the posterior probability for model  $U_j$  is given by

$$\begin{aligned}
 P(U_j|DI) &\propto \int \prod_{k=1}^n \left[ \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k}{2}} \gamma^{u_j} \frac{g_{j11} \cdots g_{ju_j u_j}}{\lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}}} \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \right. \\
 &\quad \left. \times \exp \left\{ -\frac{1}{2\sigma_k^2} [d_k^2 - h_{jk\ell}^2] \right\} d\sigma_1 \cdots d\sigma_n d\Omega_j \right].
 \end{aligned} \tag{21.52}$$

The reason for the unusual functional form for the prior probability for the amplitudes, Eq. (21.32), is that all of the pesky little values of  $2\pi$  drop out and leave one with a much cleaner functional form for the posterior probability for the model.



### 21.1.2.2 Marginalizing The Noise Standard Deviation

To evaluate the integrals over the  $\sigma_k$ , the integrand must be rearranged to isolate the integral:

$$P(U_j|DI) \propto \int \prod_{k=1}^n \left[ \gamma^{u_j} \frac{g_{j11} \cdots g_{ju_j u_j}}{\lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}}} \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \right] \times \int_0^\infty \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k}{2}} \exp \left\{ -\frac{1}{2\sigma_k^2} [d_k^2 - h_{jk\ell}^2] \right\} d\sigma_1 \cdots d\sigma_n d\Omega_j. \quad (21.53)$$

All of the integrals over the  $\sigma_k$  have the same functional form, so we can evaluate one of these integrals and then use the results to evaluate the remaining integrals. Evaluating the integral over  $\sigma_k$  one has

$$I_{\sigma_k} \equiv \int_0^\infty \frac{1}{\sigma_k} (2\pi\sigma_k^2)^{-\frac{N_k}{2}} \exp \left\{ -\frac{1}{2\sigma_k^2} [d_k^2 - h_{jk\ell}^2] \right\} d\sigma_k \quad (21.54)$$

and introducing the notation

$$Q_{jk} \equiv \frac{1}{2} [d_k^2 - h_{jk\ell}^2] \quad (21.55)$$

and  $I_{\sigma_k}$  becomes

$$I_{\sigma_k} \equiv (2\pi)^{-\frac{N_k}{2}} \int_0^\infty \sigma_k^{-(N_k+1)} \exp \left\{ -\frac{Q_{jk}}{\sigma_k^2} \right\} d\sigma_k \quad (21.56)$$

This integral can be transformed into a Gamma function. A gamma function is defined as:

$$\Gamma(z) = \int_0^\infty t^{z-1} \exp\{-t\} dz \quad (21.57)$$

so a simple change of variables and a little algebra and the integral  $I_{\sigma_k}$  is then given by

$$I_{\sigma_k} \equiv \left[ \frac{1}{2} \Gamma \left( \frac{N_k}{2} \right) (2\pi Q_{jk})^{-\frac{N_k}{2}} \right] \quad (21.58)$$

Inserting Eq. (21.58) into Eq. (21.53) one obtains

$$P(U_j|DI) \propto \int \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \prod_{k=1}^n \left\{ \left[ \gamma^{u_j} \frac{g_{j11} \cdots g_{ju_j u_j}}{\lambda_1^{-\frac{1}{2}} \cdots \lambda_{u_j}^{-\frac{1}{2}}} \right] \left[ \Gamma \left( \frac{N_k}{2} \right) (2\pi Q_{jk})^{-\frac{N_k}{2}} \right] \right\} d\Omega_j. \quad (21.59)$$

as the posterior probability for the model where we dropped the factor of one half, because it is exactly the same for all of the  $P(U_j|DI)$  and so cancels when this distribution is normalized.

This equation is correct, but not very computationally convenient because it requires an eigenvalue decomposition every time one evaluates the posterior probability. Fortunately, there are a few modifications that can be done that improve computational efficiency significantly without making approximations. First, note the presence of the eigenvalues, they are simply the determinant of the  $g_{jlm}$  matrix and can be computed with a matrix inverse. Making this substitution one obtains

$$P(U_j|DI) \propto \int \left[ \prod_{l=1}^{v_j} P(\omega_{jl}|I) \right] \prod_{k=1}^n \left[ \gamma^{u_j} g_{j11} \cdots g_{ju_j u_j} |g_{jlm}|^{-\frac{1}{2}} \right] Q_{jk}^{-\frac{N_k}{2}} d\Omega_j. \quad (21.60)$$

Next the function  $Q_{jk}$  can be computed using a matrix inverse instead of an eigenvalue decomposition, thus supplying the determinate mentioned earlier as well as allowing evaluation of the  $Q_{jk}$  function. The integral over an amplitude in a Gaussian quadrature integral, just constrains the amplitudes to their maximum posterior probability values. Consequently, we can write

$$Q_{jk} \equiv \sum_{i=1}^{N_j} \left[ d_k(t_i) - \sum_{\ell=1}^{u_j} \hat{B}_{j\ell} G_{j\ell}(t_i, \Omega_j) \right]^2 \quad (21.61)$$

and its mathematically the same as doing the eigenvalue decomposition. The amplitudes in this equation, the  $\hat{B}_{j\ell}$ , are given by the solution to

$$\sum_{k=1}^{\nu} g_{jk\ell} \hat{B}_{j\ell} = T_{j\ell}. \quad (21.62)$$

The right-hand side of this equation is the projection of the data onto the model components and is given by

$$T_{j\ell} = \sum_{i=1}^{N_j} d_j(t_i) G_{j\ell}(t_i). \quad (21.63)$$

See [2], and [11] for more on how the integrals over the amplitudes are evaluated.

Equation 21.60 is the joint posterior probability for the nonlinear parameters and is targeted by the Markov chain Monte Carlo simulations used to evaluate the remaining integrals. These simulations only vary the nonlinear parameters, the amplitudes simply do not appear in the posterior probability. However, the amplitudes are output from the simulation. The output amplitudes are given by Eq. (21.62). Because these amplitudes are estimated for each value of the nonlinear parameters, there is as many samples from the distributions of the amplitudes as there is for each of the nonlinear parameters. Consequently, the model that use marginalization do output density functions for the amplitudes.

## 21.2 Outputs Form The Enter Ascii Model Package

The Text outputs files from the Enter Ascii Model Selection packages consist of: “Bayes.prob.model,” “BayesModelAscii.mcmc.values,” “Bayes.params,” “Console.log,” “Bayes.accepted” and a condensed file “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 each parameter in the currently loaded Fortran/C model. These output probability density functions are named

`ModelFileName.ParamName`

where `ModelFileName` is the name of the currently loaded model. For example, if you have a model named `MyFunnyExp` model, and it has a decay rate named `FunnyRate` the output file containing the posterior probability for `FunnyRate` would be named:

`MyFunnyExp.FunnyRate.`

This naming convention also applies to derived parameters. So, if in addition to generating samples for `FunnyRate`, you also generated samples from a derived inverse decay rate, which was called `FunnyDecayTime` then there would also be an output file named

`MyFunnyExp.FunnyDecayTime`

containing the posterior probability for the decay time. For more on writing Ascii models in either Fortran or C, see Appendix [E](#).

# 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 für 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  $^{13}\text{C}$  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  $^2\text{H}$  and  $^{13}\text{C}$  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  $^{13}\text{C}$  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  $^{13}\text{C}$  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.