

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>22 Enter Ascii Model Selection</b>	<b>341</b>
22.1 The Bayesian Calculations	343
22.1.1 The Direct Probability With No Amplitude Marginalization	344
22.1.2 The Direct Probability With Amplitude Marginalization	346
22.1.2.1 Marginalizing the Amplitudes	347
22.1.2.2 Marginalizing The Noise Standard Deviation	352

22.2	Outputs Form The Enter Ascii Model Package	353
<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>

**H Outlier Detection**

**Bibliography**



# 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
22.1	The Enter Ascii Model Selection Package Interface . . . . .	342

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 8

# Bayes Analyze

The Bayes Analyze package analyzes one dimensional and one dimensional arrayed free inductions decays (Fid) using Bayesian probability theory [1]. The package is an implementation of the procedures described in [4, 5, 6, 2, 8]. These articles and books describe in detail the types of calculations that must be done using probability theory to detect the presence of resonances, estimate the associated parameters, and determine the number of resonances. For a tutorial on parameter estimation see Chapter 4 or reference [3]. For a tutorial on model selection see reference [11] and for a full exposition of the Bayesian Probability theory see references [33, 32]. Last, for a comparison of Bayesian and Fourier techniques for frequency estimation see reference [35, 47].

The interface to the Bayes Analyze package is shown in Fig. 8.1. The package specific widgets in the top portion of this interface have been shifted to the right. Consequently, the submit job and the server widgets are not shown. To use the Bayes Analyze package, you must do the following:

**Select** the Bayes Analyze package from the Package menu.

**Load** a spectroscopic Fid using the Files menu.

**Check** the various package specific widgets to make sure the default settings are correct. When the Bayes Analysis interface starts, the widgets on the interface are initialized using a set of defaults. These defaults automatically indicate that there is a first point problem, that a maximum of 10 resonances can be found on each run, that a correlated phase model is assumed, that all Fids should be analyzed jointly, that the entire Fid should be analyzed and finally that only singlets should be estimated.

**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 8.1: Bayes Analyze Interface

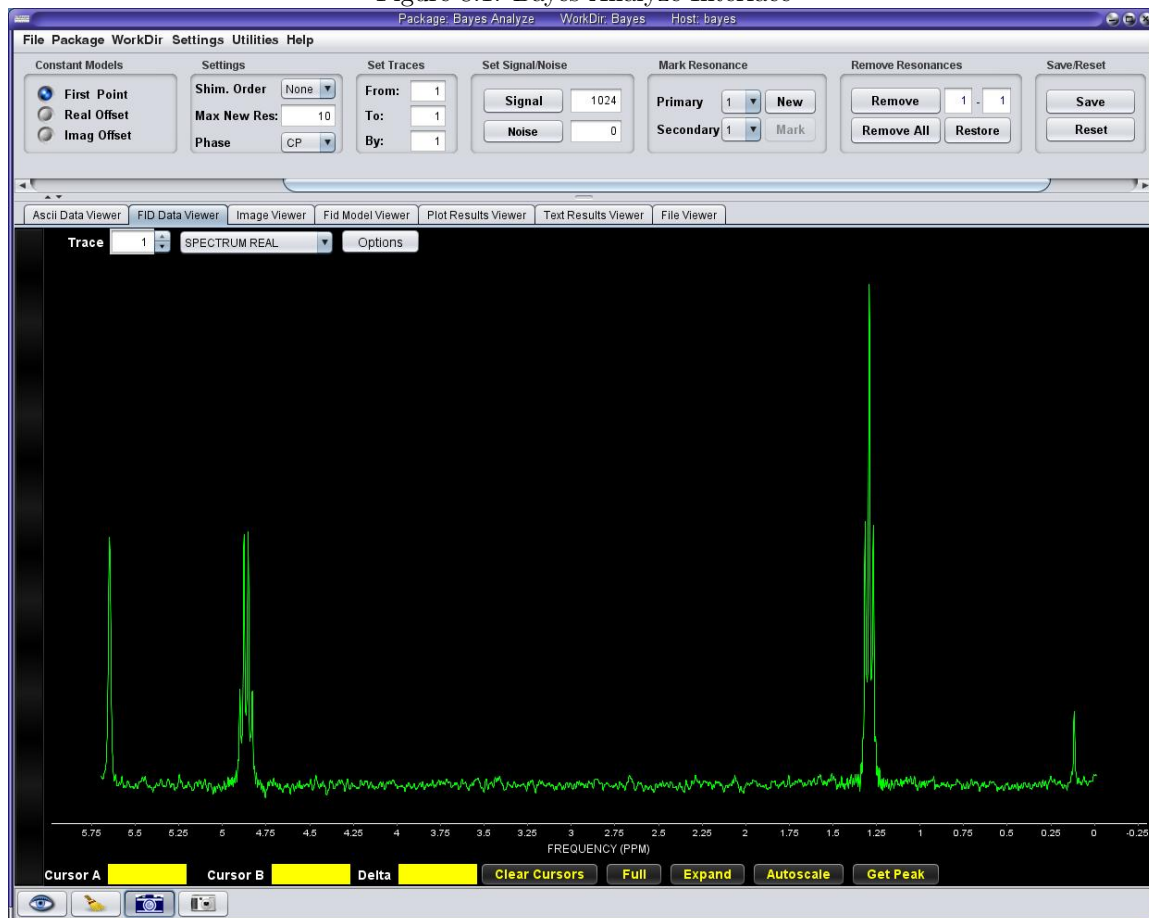


Figure 8.1: This window is used to control the execution of Bayes Analyze. On this window you may load Fids using the Files Menu, set all necessary Bayes Analyze parameters, mark multiplets, run the analysis, etc. Finally, after running the analysis, by switching to the Fid Model Viewer you can generate and view models of the loaded Fid.

Unlike, some of the other packages, Bayes Analyze has a number of package specific widgets. The reason for this is simply that Bayes Analyze is a searching routine and does not make use of Markov chain Monte Carlo. Additionally, Bayes Analyze has a number of different ways that it can model data and defining these models increases the complexity of the interface. Here we give a brief descriptions of the package specific widgets. We will divide this description into sections one section for each grouping of buttons in the interface. There are seven such groupings: Constant Models, Settings, Set Traces, Set Signal/Noise, Mark Resonances, Remove Resonances and Save/Reset. We begin by describing the constant models:

**Constant Models** NMR FID data often contain constant offsets. This is especially true when the time domain FID data are not acquired using phase cycling. Regardless of the cause of the constants, in order to analyze the FID data these offsets must be accounted for in the model. Here is a brief description of the offsets implemented in Bayes Analyze:

**First Point** is a model of the first complex data value in the Fid. This model is included by default, because the first point in an Fid is often misplaced and on some spectrometers, intentionally misplaced. The first point model allows Bayes Analyze to take into account this possibly misplaced data value.

**Real Offset** indicates that there is a constant offset in the real channel. The default is off, and checking this box causes the real offset model to be included.

**Imag Offset** indicates that there is a constant offset in the imaginary channel. The default is off, and checking this box causes the imaginary offset model to be included.

**Settings** allow the users to set some optional features of the model and the search:

**Shim. Order** allows the user to specify the shimming model used in Bayes Analyze. Loosely, the shimming model expands the line-shape of each resonance in a super position of Lorentzian's. The shimming order allows the user to set the number of Lorentzian's used. The number of Lorentzian's is forced to be odd and the valid entries are none, 1, 3, 5 and 7, where 1 is no shimming model, i.e., the lineshape is pure Lorentzian, 3 allows for lineshapes that have a single bulge on one or both sides of the line, and 5 and 7 allows for very badly shimmed lines.

**Max New Res** is the maximum number of resonances that can be added to the resonance model the next time Bayes Analyze is run. The default value is 10, but any number greater than zero can be entered.

**Phase** allows the user to specify the phase of the resonance model. There are two choices, one in which the phase of each resonance is linearly related to the phase of all other resonances. This phase model is called "CP" for correlated phase. In the correlated phase model, there are two phase parameters, a zero and first order phase. The other phase model is called "UP" for uncorrelated phase. In the uncorrelated phase model, each resonance has its own phase parameter. When marking resonances, the "Phase" widgets indicates the phase model of the marked resonances, i.e., when the phase model is uncorrelated a marked resonance has its own phase. Similarly, when the phase model is correlated phase, marked resonances have shared zero and first order phase parameters. Both correlated and uncorrelated phase resonance can be marked in the same analysis.



**Set Traces** is an area consisting of three buttons that allows the user to specify how the arrayed Fid data is to be processed.

**From** tells Bayes Analyze to start analyzing the Fid data using Fid number “From”. For example, if there are 10 Fids and you set “From” to 3, then the first Fid to be process is number 3.

**To** tells Bayes Analyze that the last Fid number to analyze is “To”. If “To” is less than the total number of Fid data, then Fids numbers greater than “To” are not analyzed. Continuing the example started in the previous item, if “To” is set to 7, then Fido’s 8 through 10 would not be processed; but Fido’s 3, 4, 5, 6 and 7 would be.

**By** tells Bayes Analyze to process the Fids in blocks of Fids of size “By”. If “By” is not an integer divisor of the total number of traces in the Fid, then the last block of traces processed by Bayes Analyze will have fewer than “By” Fids. So if “From” is 3, “To” is 7 and if “By” is 2, then Bayes Analyze will process Fido’s 3, and 4 together in the first analysis. These outputs files will have a 0003 suffix. It will then process 5 and 6 together and these outputs will be written into output files having a suffix of 0005. Finally, Fid 7 will be processed separately and the outputs from this analysis will have output files suffixed with 0007.

**Set Signal/Noise** is an area used to define both signal and noise regions in the Fid. A signal region is used by Bayes Analyze to speed up the analysis by essentially telling Bayes Analyze to only analyze the signal region. A Noise region is used by the interface to estimate the noise standard deviation in the data and there by increase Bayes Analyze’s sensitive to small resonances.

**Signal** tells Bayes Analyze to analyze the Fids using the data from the first point up to the data value entered in the “Signal” box. You can set this value by simply typing in the text entry box or you can place a single cursor on the Fid data and then activate the Signal button. Activating the signal button will set the text field to the number of the current data value.

**Noise** is used to set a noise region. The noise region is defined to that part of the Fid starting with “Noise” up to the end of the Fid. This region is used by the interface to calculate the estimated standard deviation of the noise. These values are written in to a file named bayes.noise located in Bayes/WorkDir/BayesAnalyzeFiles/bayes.noise where WorkDir must be replaced by the name of your current WorkDir. Bayes Analyze uses these noise standard deviations to assign a Gaussian likelihood rather than a Student’s  $t$ -distribution. You can set the noise region by simply typing in the noise entry box, or you can position a single cursor at the location you wish the noise region to start. Activating the noise button will set the entry box.

**Mark Resonance** is an area used to mark resonances. These resonances may be singlets, spin a half multiplets, or multiplets of multiplets.

**Primary** allows you to set the order of the primary multiplet. The primary multiplet is the multiplet having the largest  $J$ -coupling constant. The current value of the primary spinner is use to set the primary multiplet order.

**Secondary** allows you to set the order of a secondary multiplet. A secondary multiplet is the multiplet, in a multiplet of multiplets, that has the smallest  $J$ -coupling constant. For example, a triplet of doublets might have a primary coupling constant of 10 Hertz, and the secondary coupling constant could be from 1 to something less than 10 Hertz.

**New** tells the interface you are about to mark resonance. Based on the Primary and Secondary settings, the interface will display a message that indicates what must be marked and in what order. For example, if the primary coupling is 2, the secondary is 3 and you hit the “New” button:

Mark center. Mark primary J-coupling. Mark secondary J-coupling.

will be displayed and you must first mark the center of the multiplet. Second, you must mark the primary  $J$ -coupling constant. The primary coupling constant is the largest coupling constant in the multiplet. Third, you must mark the secondary  $J$ -coupling constant. The secondary  $J$ -coupling this coupling constant is the smallest coupling constants in the multiplet. In order for marking a multiple of multiplets to be successful, you must mark all three of these items in the specified order.

**Mark** tells the interface to use the current cursor position(s) for the item currently being marked. When marking coupling constants a double cursor must be used. However, a double or single cursor can be used to mark the center of a multiplet. If a double cursor is used, the multiplet center is assumed to be the average value of the two cursors. This is very useful when marking even numbered multiplets because there is no resonance at the center of things like doublets, and quartets.

**Remove Resonances** is an area used to remove resonances. You can remove individual resonances or you can remove all resonances from the current model.

**Remove** will remove the resonance number displayed in the entry box from the current model.

**Remove All** will remove *all* of the resonances in the current model.

**Restore** will restore a previously delete resonance provided the interface still has the resonance from the current model in memory.

**Save/Reset** is an area used copy an analysis back to the original Fid directory, or it can reset the analysis back to its default settings.

**Save** will copy the Bayes Analyze files from the current analysis back to the Fid directory that is current loaded. If this Fid directory contains Bayes Analyze files they are overwritten by the new analysis.

**Reset** will reset all of the parameters to their default values.

## 8.1 Bayes Model

After Bayes Analyze has been run, the output from the analysis can be used to generate a model of the time domain Fid data. This model is displayed in a special Fid Model Viewer, shown in Fig. 8.2, the bottom three quarters of the figure. To generate a Fid Model one first activated the Fid Model Viewer, and then use the “Build BA Model of FID #” button to build the model. When this button

Figure 8.2: Bayes Analyze Fid Model Viewer

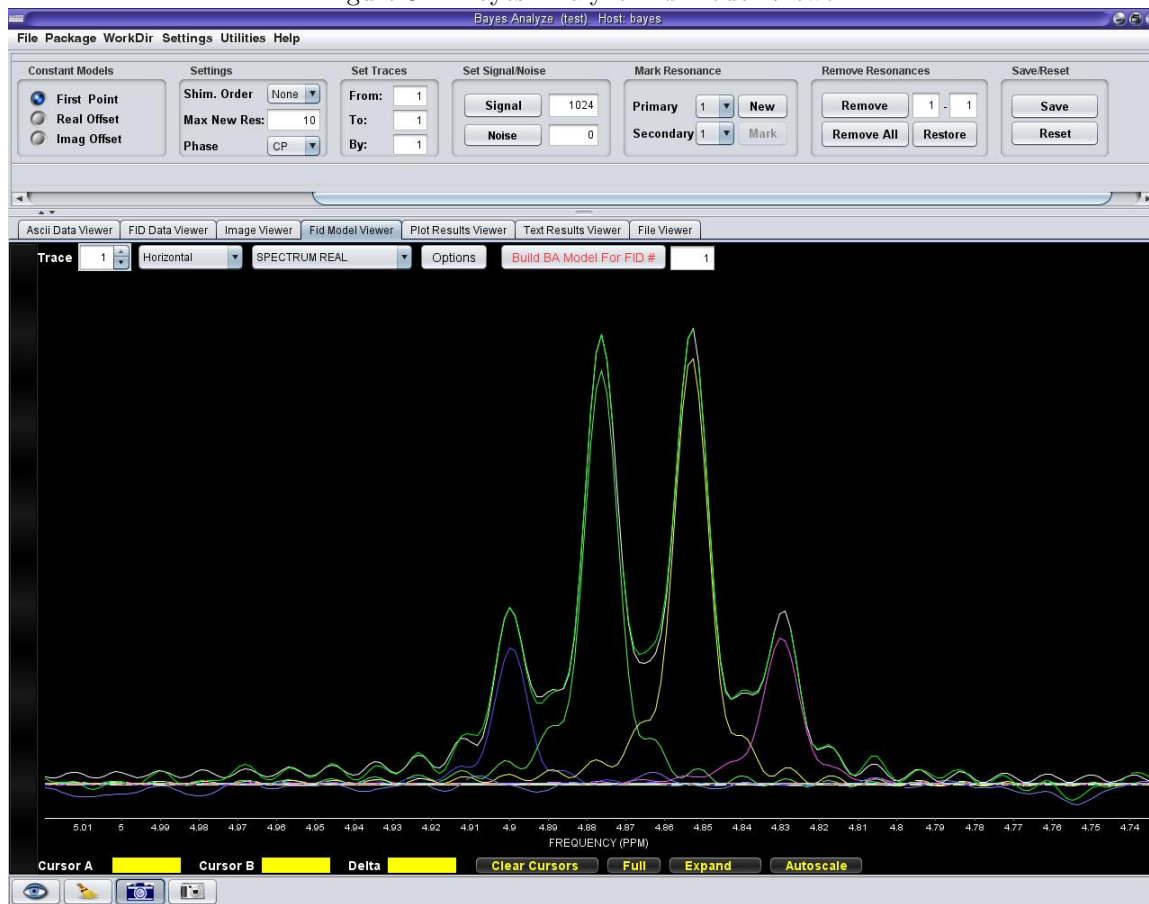


Figure 8.2: After the Bayes Analysis program has been run, the Fid Model Viewer can be used to view the results of the analysis in a graphically. Here we show the region around the quartet. This result was generated from an automatic analysis of this data. In the region shown there are a total of 4 model resonances and we have overlaid the spectrum of the time domain Fid model. This overlaid data matches the spectrum so well that the data and the model are almost indistinguishable. The other four resonances on this data are the four model components in this region. Note that because this is a time domain model, it even models the Gibbs ringing artifacts, sometimes referred to as sinc wiggles.

is activated the appropriate Fid data, and the estimated Fid parameters are sent to the server. The server unpacks the data and runs a job to build the model. During this time, the interface waits for the job to complete. When completed, the job is fetched, the model Fid is Fourier Transformed, and displayed using the Fid Model Viewer. The displayed plot in Fig. 8.2 is called a Horizontal display or plot. Here we shown the spectrum of the Ethyl Ether Fid model in the region of the quartet. This result was generated from an automatic run of Bayes Analyze using this Ethyl Ether data. In the region shown there are a total of 4 model resonances. The Horizontal plot overlays the spectrum of the time domain Fid data and the spectrum of the time domain model. This overlaid spectrum matches the spectrum so well that the data and the model are almost indistinguishable. If you look closely in the valleys between resonances, you can see a few small differences. In addition to displaying the spectrum of the data and the model, the spectrum of the individual resonances in the model are also displayed. These are the other resonances on this plot. Note that because this is a time domain model, it even models the Gibbs ringing artifacts, sometimes referred to as sinc wiggles.

## 8.2 The Bayes Analyze Model Equation

To apply probability theory as logic, one always computes the probability for some hypothesis given what one knows. In NMR one knows a great deal about the types of signals that can appear in the detector. For free induction decays, the signals are sinusoids that decay in time. In NMR if a sample loses half of its magnetization in one second, then in the second second it will lose half of its magnetization. Thus, the decay of the signal will be exponential. In the frequency domain, exponentials decay have a Lorentzian lineshape. In the time domain, the equation that relates an exponentially decaying sinusoidal signal to the data is given by

$$d_{Ri} = \sum_{j=1}^m B_j \cos(\omega_j t_i + \theta_j) e^{-\alpha_j t_i} + n_R(t_i) \quad (8.1)$$

for the real channel, where  $d_{Ri}$  is the  $i$ th data value sampled at time  $t_i$ ,  $\omega_j$  is the resonance frequency for the  $j$ th resonance,  $\alpha_j$  is its associated exponential decay rate constants,  $B_j$  is the resonance amplitude,  $\theta_j$  is the phase of the resonance,  $m$  is the number of resonances, and  $n_R(t_i)$  is the error or noise in the data. The imaginary channel is  $90^\circ$  phase shifted, and so the sinusoids are given by:

$$d_{Ii} = - \sum_{j=1}^m B_j \sin(\omega_j t_i + \theta_j) e^{-\alpha_j t_i} + n_I(t_i) \quad (8.2)$$

where  $n_I(t_i)$  are the errors or noise in the imaginary channel. Note, that the frequency  $\omega_j$  does not have a factor of  $2\pi$  in it and internally the program that implements this package uses dimensionless times: consequently,  $t_i \in \{0, 1, 2, \dots, N-1\}$ , that is to say,  $t_i$  is a simple integer. In these units the frequencies and decay rate constants have units of radians. In this package when the user marks a frequency in Hertz, or PPM, the package converts the frequency into radians. This conversion to Hertz is given by:

$$\omega_j = \frac{(f_j + f_r)2\pi\Delta t}{N} \quad (8.3)$$

where  $\omega_j$  is the internal frequency in radians,  $f_j$  is the frequency in Hertz,  $f_r$  is the reference frequency in Hertz,  $\Delta t$  is the time interval between data values in seconds, and  $N$  is the total

number of complex data values acquired in the free induction decay. Similarly, the decay rates  $\alpha_j$  are also in radians. The conversion of the decay rates to these internal units is given by:

$$\alpha_j = \frac{\beta_j \pi \Delta t}{N} \quad (8.4)$$

where  $\beta_j$  is the full width at half-maximum of the resonance in the absorption spectrum. Note that there are two differences between the frequency and decay rate conversions: we don't reference decay rates and there is no factor of 2.

Equations (8.1 and 8.2) are the conventions used by Varian/Agilent and Siemens spectrometers. Bruker spectrometers use different conventions. The different conventions come about because the data are very high frequency and must be mixed down to a set of difference frequencies that can be properly digitized. In the process of mixing down, the reference sines and cosines may have either positive or negative frequencies, and the two channel could correspond to either a cosine or sine mixing. Consequently, there are four different sets of model equations corresponding to different sign conventions. These different sign conventions can effectively aliases either the positive or negative frequencies in the spectrum depending on which convention is used. Additionally, some NMR spectrometers use only a single digitizer and collect the real and imaginary channels at different times. This time difference is not accounted for in Eqs. (8.1 and 8.2). Varian/Agilent and Siemens use a cosine and minus sine combination as shown in Eqs. (8.1 and 8.2). Bruker, uses a different convention and because of this Bruker spectroscopic data cannot be correctly processed by the Bayes Analyze package. At least at the present time the "File/Load Spectroscopic Fid" menu does not load Bruker spectroscopic data. If one wished to analyze such data, an Ascii Model could be written using the proper sine and cosine conventions and then run using either the Enter Ascii Model package or the Enter Ascii Model Selection package. However, because these packages read Ascii data, the Bruker data would have to be converted to an appropriate Ascii file.

Model Eqs. (8.1 and 8.2) leave out a great deal of what is often known about the NMR signal. For example, these equations ignore constant offsets, baseline artifacts, multiplets, multiplets of multiplets, linear phase errors in the frequency domain, and non-Lorentzian lineshapes. In spite of this, a program that implemented this complex model would still be useful, because this model would allow for the estimation of frequencies, amplitudes, phases, decay rate constants, and the number of resonances in most NMR free induction decays.

The actual model signal used in the Bayes Analyze package is more sophisticated than Eqs.(8.1 and 8.2). For example, the way Eqs. (8.1 and 8.2) were written implies that the phase of the sinusoids is different for each resonance and this is not necessarily true. Indeed the model used by the Bayes Analyze package allows the user to specify whether the phases of the resonances are the same ("correlated") or different ("uncorrelated"). If the resonances are correlated the model includes a common "zero order" phase and a linear frequency-dependent "first order" phase shift. These two parameters are in exact analogue of the left and right phase parameters that VnmrJ uses. Additionally, the model includes spin  $\frac{1}{2}$  "AX" multiplets and multiplets of multiplets. If these modifications were the only additional prior information built into the model, the model would be considerably more sophisticated than Eqs. (8.1 and 8.2) would indicate. However, the actual model in use also takes into account two other types of information.

The first type of additional information is that there may be a constant offset in the data. No amplifier is perfect and, as a result, NMR free induction decays often have constant offsets. This is particularly true if phase cycling is not used. If these offsets are not accounted for in the model, they would introduce artifacts into the analysis. Three different types of constant models are allowed:

a first point model, a constant in the real channel and a constant in the imaginary channel. The first point model is just what its name implies: it is a model of the first time-domain point in the data. In essence this model assumes that the first time-domain data value is not correct. The types of things that could cause the first point to be off are the presence of “probe or filter ringing” or an extremely rapidly decaying sinusoid in the data. Including the first point model is the default and you must uncheck the “Constant Models/First Point” to turn this model off. The second and third constant models are a constant offsets in the real and imaginary channel. These offset models are different from the first point. The first point model is a constant that extends only one point in time, while the offset models extends all the way through the acquisition. The use of these constant models is selected by the user by toggling the “Constant Models/Real Offset.” These three constant models may be used either separately or in any combination.

The second type of additional information that was incorporated into the model equation was that the lineshapes are not pure Lorentzian. One of the problems that had to be faced early in the development of this package was what decay model to use. In the development of the Bayes Analyze package various models of the decay were postulated and probability theory was used to test these models against the Lorentzian model. It was found that no simple models of the physics were better than the Lorentzian model. To give just a few examples of the model tested, we tested Gaussian, Lorentzian plus Gaussian, linear drifts in the  $H_0$  field (which imply a  $J_0$  Bessel function corrections to the decay), Taylor expansions of the lineshape, multi-exponential decay, and a host of others. With one exception, none of these models worked better than Lorentzian.

The model finally implemented in the Bayes Analyze program takes into account one simple observation: badly shimmed resonances in the frequency domain look like a superposition of several frequencies. Indeed this is precisely what a shimming artifacts is. If the magnetic field is not homogeneous then different parts of the sample have different resonance frequencies. The model implemented in Bayes Analyze treats each resonance in the frequency domain as superposition of Lorentzian's. The Lorentzian's are equally spaced and, for reasons that will become apparent, the order of the expansion is always odd. The amplitude of each Lorentzian is chosen so that the lineshape of the modeled resonances match the common lineshape of the spectrum. The total relative amplitude of the Lorentzian's sums to one, so that the amplitude of the modeled resonance retains its normal meaning. The default model used in Bayes Analyze is no shimming. That is to say the lineshapes are not expanded in a super position of Lorentzian's. To select the shimming model, one must activate the “Settings/Shim Order” and select the expansion order that is to be used. Bayes Analyze does not automatically select this setting.

We are now in a position to state the model that describes a resonance in the Bayes Analyze package. Note that this is just the resonance portion of the model, it does not include the constants. In the Bayes Analyze package, a resonance is any group of related frequencies that are treated as a single unit by the package. Thus, a singlet, a multiplet, and a multiplet of multiplets are all examples of resonances even though the Fourier transform of some of these resonances could have multiple peaks. With this terminology out of the way, the model for a resonance in the Bayes Analyze package is given by

$$M_{Rl}(t) = \sum_{j=1}^p \sum_{k=1}^s \sum_{h=1}^o P_j S_k R_h \cos(\omega_{h_j k l} [t + t_0] + \phi_l) e^{-\alpha_l t} \quad (8.5)$$

where  $M_{Rl}(t)$  is the model of the  $l$ th resonance in the real channel, and

$$M_{Rl}(t) = - \sum_{j=1}^p \sum_{k=1}^s \sum_{h=1}^o P_j S_k R_h \sin(\omega_{h j k l} [t + t_0] + \phi_l) e^{-\alpha_l t} \quad (8.6)$$

is the model of the  $l$ th resonance in the imaginary channel, where  $\alpha_l$  is the decay rate constant of the resonance, the frequencies,  $\omega_{h j k l}$ , the three relative amplitudes,  $P_j$ ,  $S_k$ ,  $R_h$ , and the multiplet orders,  $p$ ,  $s$  and  $o$ , are defined shortly. The time delay,  $t_0$ , is a first order phase correction and is only present in correlated resonances. As noted earlier, correlated resonances are ones that can be phased using a common zero and first order phase. The zero order phase of the  $l$ th resonance is represented by  $\phi_l$  and it also is only present for correlated resonances. In NMR the phase of a resonance can be common to two or more correlated resonances. When the phase is specific to a single resonance, the model is called an uncorrelated resonance model. When the resonance model is uncorrelated, the model can be rewritten as a sum of sinusoids containing a real and imaginary amplitude with no constant phase and this rewritten model is actually used in Bayes Analyze for uncorrelated resonances. However, for the current discussion we do not introduce this additional complication, we are going to discuss the Bayes Analyze resonance model as if it is given by Eq. (8.5) and Eq. (8.6) respectively. Given this resonance model, the data and the resonance models are then related by:

$$d_{Ri}(t) = C_{Ri} + F_{Ri} \delta(t-1) + \sum_{l=1}^m B_{li} M_{Rl}(t) + n_{Ri}(t) \quad (8.7)$$

for the real channel, and

$$d_{Ii}(t) = C_{Ii} + F_{Ii} \delta(t-1) + \sum_{l=1}^m B_{li} M_{Rl}(t) + n_{Ii}(t) \quad (8.8)$$

for the imaginary channel, where  $C_{Ri}$  and  $C_{Ii}$  are real and imaginary constant offsets in the  $i$ th data set,  $F_{Ri}$  and  $F_{Ii}$  are the magnitude of the first point in the real and imaginary channel in the  $i$ th data set,  $\delta(t-1)$  is a Kronecker delta function,  $B_{li}$  is the time domain amplitude of the  $l$ th resonance in the  $i$ th data set, and  $n_{Ri}(t)$  and  $n_{Ii}(t)$  represents the noise in the real and imaginary channels in the  $i$ th data set. The time domain amplitude,  $B_{li}$ , is proportional to the integral over the resonance peak in the frequency domain in the  $i$ th data set. Consequently, when we use the word amplitude, for all practical purposes it is synonymous with integral. For multiplets the amplitude  $B_{li}$  is the total amplitude of the multiplet in a given data set. In the frequency domain this is the total integral over all peaks in the multiplet. The order of the primary multiplet (defined to be the multiplet with the largest coupling constant) has been designated as  $p$ . Similarly,  $s$  is the order of the secondary multiplet. The expansion order of the shimming model is  $o$ . The normalized relative amplitude of the  $j$ th peak frequency in a primary spin one half multiplet is designated by  $P_j$ . Similarly,  $S_k$  is the normalized relative amplitude for the  $k$ th peak frequency in the secondary multiplet. Up to the normalization constant,  $P_j$  and  $S_k$  are given by Pascal's triangle, Table 8.1. The normalized relative amplitudes of each Lorentzian used in the lineshape expansion are designated as  $R_h$ . Finally,  $\omega_{h j k l}$  is the frequency of the various peaks in the spectrum of a spin one-half multiple of multiplets where the Lorentzian expansion of the lineshape is, for all practical purposes, a third multiplet. The frequencies  $\omega_{h j k l}$  are given by

$$\omega_{h j k l} = \omega_l - \frac{J_o(o+1-2h)}{2} - \frac{J_{pl}(p+1-2j)}{2} - \frac{J_{sl}(s+1-2k)}{2} \quad (8.9)$$

Table 8.1: Multiplet Relative Amplitudes

Multiplet Name	Relative Amplitudes											
Singlet	1											
Doublet	1	1										
Triplet	1	2	1									
Quartet	1	3	3	1								
Pentet	1	4	6	4	1							
Heptet	1	5	10	10	5	1						
Septet	1	6	15	20	15	6	1					
Octet	1	7	21	35	35	21	7	1				
Nonatet	1	8	28	56	70	56	28	8	1			
Decatet	1	9	36	84	126	126	84	36	9	1		
Undecatet	1	10	45	120	210	252	210	120	45	10	1	
Dodecatet	1	11	55	165	330	462	462	330	165	55	11	1

Table 8.1: The relative amplitudes of the peaks in a multiplet are given by Pascal's Triangle. Here we lists the names and these relative amplitudes of the primary and secondary multiplet orders. Note that before the model is actually run the Bayes\_Analyze program normalizes this table so that the total area of a multiplet is one. Consequently, the amplitude reported for a multiplet is the total amplitude for all of the peaks within the multiplet.

where the location of the center of the  $l$ th resonance is given by  $\omega_l$  and is the resonance frequency. For a simple triplet,  $\omega_l$  would be the frequency of the center of the triplet and for a doublet  $\omega_l$  would be the average of the two doublet resonance,  $J_o$  is the spacing of the frequency components in the Lorentzian expansion of the lineshape. Note that this spacing is common to all resonances and is not specific to a given resonance. This ensures that the lineshape for all resonance is the same. The  $J$ -coupling constant for the primary multiplet is designated as  $J_{pl}$ , and  $J_{sl}$  is  $J$ -coupling constant for the secondary multiplet. Note that if  $o = p = s = 1$ , Eqs. (8.7,8.8) reduce to the single exponentially decaying sinusoidal model, Eqs. (8.1 and 8.2). Also note that so long as the shimming order is one,  $s = 1$ , this model is just the exponentially decaying sinusoidal model of multiplets of multiplets described earlier. When the shimming order is not one then this model is of multiplets of multiplets where the lineshape is determined by a superposition of  $o$  Lorentzian's of relative amplitudes  $R_h$ .

The actual resonance model used in the Bayes Analyze package is a superposition of multiple resonances plus whatever constant models have been selected by the user. The relative amplitudes  $R_h$  are common to every resonance in every data set. The resonance frequencies and decay rate constants are common for every data set. Correlated resonances have a common phase  $\phi$ , while uncorrelated resonances each have a unique phase in each data set. This unique phase is actually implemented as a cosine and sine amplitude in the programs; not as an amplitude plus a phase as illustrated in Eqs. (8.1 and 8.2). The first order phase correction  $t_0$  appears only in correlated resonances. Finally, the amplitudes  $B_{li}$  is the amplitude of the  $l$ th resonance in the  $i$ th data set.

For notational simplicity, we are going to consolidate the sum in Eq. (8.7) and Eq. (8.8) into a



single summation:

$$d_{Ri}(t) = \sum_{l=1}^{m'} B'_{li} M'_{Rl}(t) + n_{Ri}(t) \quad (8.10)$$

$$d_{Ii}(t) = \sum_{l=1}^{m'} B'_{li} M'_{Il}(t) + n_{Ii}(t) \quad (8.11)$$

where constant terms were made a part of the sum, the amplitudes have been redefined to include the constant offsets and the first point problem, the functions  $M'_{Rl}$  and  $M'_{Il}$  also include the constant and first point models, and the model count,  $m'$ , is a count of the total models including the constants and first point models.

The Bayes Analyze program searches for the most probable value of these parameters. The search is done in the logarithm of the joint posterior probability for the parameters independent of the amplitudes of the resonances. The search algorithm is a modified version of a Levenberg-Marquardt algorithm. When the shimming model is activated, the shimming delta  $J_o$  and the relative shimming amplitudes  $R_h$  are common to all resonance models. Because, these parameters are the same for every resonance model, the program is looking for effects that are common to all resonances, that is why we tend to call this model a shimming artifact model, not a lineshape model. A true lineshape model would allow each resonance to have a different lineshape, as sometimes happens, and this is not what the shimming model does. The shimming model is multiplied by the exponentially decaying part of the signal. Thus the width of the Lorentzian's that make up the lineshape expansion in the frequency domain are different for every resonance. When all resonance have comparable linewidth, the shape of the line in the discrete Fourier transform is highly effected by the shimming model. When some lines are narrow and some are broad, the narrow line are primarily effected by the shimming model and appear highly non-Lorentzian. Broad lines, however, are minimally affected by this model and appear highly Lorentzian. In either case the estimated decay rate constant is decreased, not increases. The lines do decay faster, but it is interference that causes the decay, not an increase in the decay rate constant.

There are three sets of relative amplitudes:  $P_j$ ,  $S_j$ , and  $R_j$ . Each of these relative amplitude are normalized. The relative amplitude of a resonance is a weighted average over these three, so the total relative amplitude of a resonance is one. Consequently, the amplitude  $B$  is the total amplitude of all components that makes up a resonance. In the frequency domain this means that the amplitude  $B$  is the proportional to the integral over all peaks that make up a resonance. The Bayes Analyze package takes special steps to make sure that the total relative amplitude is one. For  $P_j$  and  $S_j$  this means normalizing the amplitudes from Pascal's triangle. However, for the relative shimming amplitudes, the  $R_j$ , the problem is more complex. Imposing the condition that the total relative shimming amplitude be one is accomplished by assigning the amplitude of the center component such that the total is one. However, for this model to have physical significance all of the  $R_j$  must be positive, including the center one. But imposing these two conditions, assigning the center amplitude and positively, is not enough to ensure a physically meaningful result, because these conditions allow the resonance frequency to be different from the center of a peak in the discrete Fourier transform. To avoid this problem, three additional conditions are imposed: the shimming expansion order is always odd, the center component is required to have the largest relative amplitude, and the relative shimming amplitudes must increase up to the center and then decrease beyond. With these conditions in place the search algorithm places the center of the shimming expansion in the center of

the line in the frequency domain. Consequently, the resonance frequency corresponds to the center of a resonance, and the total amplitude is proportional to the total integral over the lines in the frequency domain.

All of these conditions are imposed using prior probabilities. These prior are such that they have almost no effect on the result so long as the appropriate conditions are met, but when the conditions are about to be violated they prevent the parameters from moving to physically meaningless values. The various priors make sure that  $J$ -coupling constants and decay rate constants never go negative. They ensure that the total shimming amplitudes are ordered, positive, and sum to one. Last, they ensure that resonances marked by the user stay near where the user marked them. All of these conditions are informative and aid in making sure the output from Bayes Analyze is physically meaningful.

### 8.3 The Bayesian Calculations

The program that implements this frequency finding package is called Bayes Analyze. Unlike other packages in the Bayesian Analysis this program does not implement the calculations using Markov chain Monte Carlo; rather the program that implements this calculation uses a series of approximations and a Levenberg-Marquardt searching algorithm to locate the peak of the joint posterior probability for the parameters. At the maximum of the posterior probability, Bayes Analyze then uses a Taylor expansion about the peak to approximate the joint posterior probability for the parameters. Integrating over this approximation gives an estimate of the parameters and the uncertainty in those parameter estimates and it gives an estimate of the posterior probability for the model. Using Bayes' theorem [1] and the sum rule of probability theory, the joint posterior probability for the nonlinear parameters is given symbolically by

$$P(\Omega|D_R D_I I) = P(\Omega|I) \int P(\{B'\}\{\sigma\}|I) P(D_R D_I | \Omega \{B'\}\{\sigma\} I) d\{B'\} d\{\sigma\} \quad (8.12)$$

where we are using the notation  $\{\cdot\}$  to stand for all of the parameters of the dot kind. For example  $\{B'\}$  means all of the amplitudes in all of the data sets. The integrals are over all of the amplitudes and the standard deviations of the noise in all of the data sets. Assuming that the priors for the amplitudes are independent, the integrand can be factored to obtain

$$P(\Omega|D_R D_I I) = P(\Omega|I) \prod_{i=1}^n \int P(\{B'\}_i | I) P(\sigma_i | I) P(D_{Ri} D_{Ii} | \Omega \{B'\}_i \sigma_i I) d\{B'\}_i d\sigma_i \quad (8.13)$$

where  $P(\{B'\}_i | I)$  is the prior probability for the  $m'$  amplitudes in the  $i$ th data set,  $P(\sigma_i | I)$  is the prior probability for the standard deviation of the noise in the  $i$ th data set,  $n$  is the number of Fids being analyzed jointly, and  $P(D_{Ri} D_{Ii} | \Omega \{B'\}_i \sigma_i I)$  is the joint likelihood of the real and imaginary data in the  $i$ th data set. If we further assume that the real and imaginary measurements are independent, then likelihood can be factored into a product of two likelihoods, one for the real and one for the imaginary data:

$$P(\Omega|D_R D_I I) = P(\Omega|I) \prod_{i=1}^n \int P(\{B'\}_i | I) P(\sigma_i | I) P(D_{Ri} | \Omega \{B'\}_i \sigma_i I) P(D_{Ii} | \Omega \{B'\}_i \sigma_i I) d\{B'\}_i d\sigma_i \quad (8.14)$$

where  $P(D_{Ri}|\Omega\{B'\}_i\sigma_i I)$  and  $P(D_{Ii}|\Omega\{B'\}_i\sigma_i I)$  are the likelihoods of the real data and the imaginary data in the  $i$ th data set. Assuming logical independence, the prior probability for the amplitudes in data set  $i$ ,  $P(\{B'\}_i|I)$ , can be factored to obtain:

$$P(\{B'\}_i|I) = \prod_{j=1}^{m'} P(B'_{ji}|I). \quad (8.15)$$

The prior probability for each amplitude in including the first point and constant offset amplitudes, is represented by  $P(B'_{il}|I)$  and we will assign this prior probability using a Gaussian prior:

$$P(B'_{il}) = \left(\frac{2\pi\sigma_i^2}{\gamma^2}\right)^{-\frac{1}{2}} \exp\left\{-\frac{\gamma^2 B'_{li}{}^2}{2\sigma_i^2}\right\} \quad (8.16)$$

where  $\gamma$  indicates how strongly we believe the amplitude is small rather than large and we have included the factor of  $\sigma_i$  for computational convenience. The direct probability for the real data is represented by  $P(D_{Ri}|\Omega\{B'\}_i\sigma_i I)$  and  $P(D_{Ii}|\Omega\{B'\}_i\sigma_i I)$  is the direct probability for the imaginary data. We will assign these likelihoods for each data set using a Gaussian likelihood:

$$P(D_{Ri}|\Omega\{B'\}_i\sigma_i I) = (2\pi\sigma_i^2)^{-\frac{N}{2}} \exp\left\{-\sum_{k=1}^N \frac{[d_{Ri}(t_k) - \sum_{l=1}^{m'} B'_{li} M'_{Rl}(t_k)]^2}{2\sigma_i^2}\right\} \quad (8.17)$$

and similarly for the direct probability for the imaginary data:

$$P(D_{Ii}|\Omega\{B'\}_i\sigma_i I) = (2\pi\sigma_i^2)^{-\frac{N}{2}} \exp\left\{-\sum_{k=1}^N \frac{[d_{Ii}(t_k) - \sum_{l=1}^{m'} B'_{li} M'_{Il}(t_k)]^2}{2\sigma_i^2}\right\} \quad (8.18)$$

where it is assumed that the average noise power in the two channels is the same and that each data set contains exactly  $N$  data values.

To compute the posterior probability for one of the parameters,  $P(\omega_j|D_R D_I I)$ , two sets of integrals must be done: one over the amplitudes  $\{B'\}$  and one over all of the nonlinear parameters except the parameter being estimated. In the program that implements this calculation, the integral over the amplitudes is done exactly while the integrals over the nonlinear parameters are done approximately using a Lorentzian approximation. We briefly sketch how the integrals over the amplitudes are evaluated. If we assign a Jeffreys' prior probability [33] for the noise standard deviations, Eq. (8.16) for the prior probability for the amplitudes and assign Gaussian likelihood functions, then the joint posterior probability for the  $\Omega$  parameters independent of the amplitudes becomes:

$$\begin{aligned} P(\Omega|D_R D_I I) &= P(\Omega|I) \prod_{i=1}^n \int \frac{1}{\sigma_i} \int \left(\frac{2\pi\sigma_i^2}{\gamma^2}\right)^{-\frac{m'}{2}} \exp\left\{-\sum_{l=1}^{m'} \frac{\gamma^2 B'_{li}{}^2}{2\sigma_i^2}\right\} \\ &\times (2\pi\sigma_i^2)^{-\frac{N}{2}} \exp\left\{-\sum_{j=1}^N \frac{[d_{Ri}(t_j) - \sum_{l=1}^{m'} B'_{li} M'_{Rl}(t_j)]^2}{2\sigma_i^2}\right\} \\ &\times (2\pi\sigma_i^2)^{-\frac{N}{2}} \exp\left\{-\sum_{j=1}^N \frac{[d_{Ii}(t_j) - \sum_{l=1}^{m'} B'_{li} M'_{Il}(t_j)]^2}{2\sigma_i^2}\right\} d\{B'\}_i d\sigma_i \end{aligned} \quad (8.19)$$

where  $d_{Ri}(t_j)$  is the  $j$ th real data value of the  $i$ th data set and similarly  $d_{Ii}(t_j)$  is the corresponding imaginary data value. There are  $n$  sets of integrals over the amplitudes. Each set of amplitude integrals are  $m'$  dimensional Gaussian quadrature integrals having bounds ranging from minus to plus infinity. One evaluates these amplitude integrals by first simplifying the exponent of the integrand, and the joint posterior probability is reduced to:

$$P(\Omega|D_R D_I I) = P(\Omega|I) \prod_{i=1}^n \int \frac{1}{\sigma_i} \int \left( \frac{2\pi\sigma_i^2}{\gamma^2} \right)^{-\frac{m'}{2}} (2\pi\sigma_i^2)^{-N} \exp \left\{ -\frac{Q_i}{2\sigma_i^2} \right\} d\{B'\}_i d\sigma_i \quad (8.20)$$

where  $Q_i$  is defined to be

$$Q_i \equiv \sum_{l=1}^{m'} \gamma^2 B_{li}'^2 + \sum_{j=1}^N [d_{Ri}(t_j) - \sum_{l=1}^{m'} B_{li}' M_{Ri}'(t_j)]^2 + \sum_{j=1}^N [d_{Ii}(t_j) - \sum_{l=1}^{m'} B_{li}' M_{Ii}'(t_j)]^2. \quad (8.21)$$

To proceed,  $Q_i$  must be simplified:

$$\begin{aligned} Q_i &= d_{Ri} \cdot d_{Ri} + d_{Ii} \cdot d_{Ii} \\ &\quad - 2 \sum_{j=1}^{m'} B_{ji}' [d_{Ri} \cdot M_{Rj} + d_{Ii} \cdot M_{Ij}] \\ &\quad + \sum_{j=1}^{m'} \sum_{l=1}^{m'} B_{ji}' B_{li}' [M_{Rj} \cdot M_{Rl} + M_{Ij} \cdot M_{Il} + \gamma^2 \delta_{jl}] \end{aligned} \quad (8.22)$$

where “ $\cdot$ ” means sum over time and  $\delta_{jl}$  Kronecker delta function. If the quadrature data are thought of as a complex vector of length  $N$ , then the three lines making up  $Q_i$  are the total squared data value, the length of the projection of the data onto the model, and, if not for the prior probability for the amplitudes, the third term is the total squared length of the model projected onto itself. In this equation, the prior acts as a safeguard and prevents the total projection of the model onto the data from being equal to the model projected onto itself. When this calculation is implemented, the projection of the model onto the data should never be equal to the projection of the model onto itself, the only what this could happen is if the data had no noise. Regardless, the prior prevents this singular arithmetic and, consequently, the amplitudes that maximize the function  $Q_i$  are always real numbers, i.e., positive or negative, but never complex numbers.

Defining an interaction matrix  $g_{jk}$ :

$$g_{jl} \equiv M_{Rj} \cdot M_{Rl} + M_{Ij} \cdot M_{Il} + \gamma^2 \delta_{jl}, \quad (8.23)$$

$Q_i$  can be rewritten as:

$$\begin{aligned} Q_i &= d_{Ri} \cdot d_{Ri} + d_{Ii} \cdot d_{Ii} \\ &\quad - 2 \sum_{j=1}^{m'} B_{ji}' [d_{Ri} \cdot M_{Rj} + d_{Ii} \cdot M_{Ij}] \\ &\quad + \sum_{j=1}^{m'} \sum_{l=1}^{m'} B_{ji}' B_{li}' g_{jl}. \end{aligned} \quad (8.24)$$

These integrals are multivariate Gaussian integrals and any multivariate integral of this form may be done analytically. These integrals are exactly the same for each data set, and here we do these

integrals for an arbitrary data set and then use that result to generate the appropriate product. First, a change of variables is introduced:

$$B'_{k'i} = \sum_{j=1}^{m'} \frac{A_{ji}e_{jk}}{\sqrt{\lambda_j}}, \quad A_{ki} = \sqrt{\lambda_k} \sum_{j=1}^{m'} B'_{ji}e_{kj}, \quad (8.25)$$

with

$$dB'_{1i} \cdots dB'_{m'i} = \lambda_1^{-\frac{1}{2}} \cdots \lambda_{m'}^{-\frac{1}{2}} dA_{1i} \cdots dA_{m'i}. \quad (8.26)$$

Next we introduce a change of function:

$$R_k = \sum_{j=1}^{m'} \frac{M_{Rj}e_{kj}}{\sqrt{\lambda_k}}, \quad \text{with} \quad \sum_{j=1}^{m'} B'_{ji}M_{Rj} = \sum_{j=1}^{m'} A_{ji}R_j(t), \quad (8.27)$$

$$I_k = \sum_{j=1}^{m'} \frac{M_{Ij}e_{kj}}{\sqrt{\lambda_k}}, \quad \text{with} \quad \sum_{j=1}^{m'} B'_{ji}M_{Ij} = \sum_{j=1}^{m'} A_{ji}I_j(t), \quad (8.28)$$

where  $e_{jk}$  is the  $k$ th component of the  $j$ th eigenvector of the interaction matrix, Eq. (8.23), and  $\lambda_j$  is its  $j$ th eigenvalue. Note the  $g_{jk}$  matrix is not data set dependent, and consequently, these eigenvalues and the corresponding eigenvectors are not data set dependent. Using the property

$$\sum_{i=1}^N R_j(t_i)R_k(t_i) + I_j(t_i)I_k(t_i) = \delta_{jk}, \quad (8.29)$$

the joint posterior probability of the nonlinear  $\Omega$  parameters Eq. (20) becomes

$$\begin{aligned} P(\Omega|D_R D_I I) &= P(\Omega|I) \prod_{i=1}^n \int \frac{d\sigma_i}{\sigma_i} (2\pi\sigma_i^2)^{-2N} \left( \frac{2\pi\sigma_i^2}{\gamma^2} \right)^{-\frac{m'}{2}} \lambda_1^{-\frac{1}{2}} \cdots \lambda_{m'}^{-\frac{1}{2}} \\ &\times \int_{-\infty}^{\infty} dA_{1i} \cdots dA_{m'i} \exp \left\{ -\frac{Q'_i}{2\sigma_i^2} \right\}, \end{aligned} \quad (8.30)$$

where

$$Q'_i = d_{Ri} \cdot d_{Ri} + d_{Ii} \cdot d_{Ii} - 2 \sum_{j=1}^{m'} A_{ji} h_{ji} + \sum_{j=1}^{m'} A_{ji}^2 \quad (8.31)$$

and

$$h_{ji} \equiv d_{Ri} \cdot R_{ji} + d_{Ii} \cdot I_{ji}, \quad (8.32)$$

after completing the square and evaluating the  $m'$  integrals for each data set, one obtains

$$P(\Omega|D_R D_I I) = P(\Omega|I) \prod_{i=1}^n \int d\sigma_i \lambda_1^{-\frac{1}{2}} \cdots \lambda_{m'}^{-\frac{1}{2}} \sigma_i^{-2N-1} \gamma^{m'} \exp \left\{ -\frac{Q''_i}{2\sigma_i^2} \right\}, \quad (8.33)$$

where

$$Q''_i = d_{Ri} \cdot d_{Ri} + d_{Ii} \cdot d_{Ii} - \sum_{j=1}^{m'} h_{ji}^2. \quad (8.34)$$

Bayes Analyze uses this joint posterior probability in its calculation. However, it does not use this functional form of the calculation. This functional form is the result of doing the multivariate Gaussian integrals but it is not computationally convenient because of the presence of the eigenvalues and eigenvectors of the  $g_{kl}$  matrix. If you look at Eq. (8.31) and Eq. (8.34), it's clear that evaluating the integrals has just constrained the amplitudes  $A_{ji}$  to the values that maximized the joint posterior probability. If we go back to Eq. (8.24) and constraining  $Q_i$  to the amplitudes that maximize  $Q_i$ , then Eq. (8.24) and Eq. (8.34) must be identical. However, the maximum of Eq. (8.24) can be computed with a matrix inverse while Eq. (8.34) requires a full eigenvalue decomposition.

To locate the maximum of Eq. (8.24), one takes the derivative of Eq. (8.24) with respect to an arbitrary amplitude  $B'_{ki}$  and sets the resulting derivative to zero

$$\begin{aligned} \frac{dQ_i}{dB'_{ki}} &= -2 \sum_{j=1}^{m'} \frac{d}{dB'_{ki}} \hat{B}'_{ji} [d_{Ri} \cdot M_{Rj} + d_{Ii} \cdot M_{Ij}] + \sum_{j=1}^{m'} \sum_{l=1}^{m'} \frac{d}{dB'_{ki}} \hat{B}'_{ji} \hat{B}'_{li} g_{jl} \\ &= -2 [d_{Ri} \cdot M_{Rk} + d_{Ii} \cdot M_{Ik}] + \sum_{l=1}^{m'} \hat{B}'_{li} g_{kl} + \sum_{j=1}^{m'} \hat{B}'_{ji} g_{jk} \end{aligned} \quad (8.35)$$

but the matrix  $g_{jk}$  is a symmetric matrix so the two sums are identical and the maximum is given by the solution of:

$$0 = -2 [d_{Ri} \cdot M_{Rk} + d_{Ii} \cdot M_{Ik}] + 2 \sum_{l=1}^{m'} \hat{B}'_{li} g_{kl} \quad (8.36)$$

where we have designated the amplitudes that maximize the joint posterior probability by  $\hat{B}'_{li}$ . The  $\hat{B}'_{li}$  are given by

$$\hat{B}'_{li} = \sum_{v=1}^{m'} (g_{kl})_{vl}^{-1} T_{vi} \quad (8.37)$$

where  $T_{vi}$  is given by

$$T_{vi} = d_{Ri} \cdot M_{Rv} + d_{Ii} \cdot M_{Iv} \quad (8.38)$$

and  $(g_{kl})_{vl}^{-1}$  is the  $l$ th element of the  $v$ th column of the inverse of the  $g_{kl}$  matrix. If we were to substitute  $\hat{B}'_{ui}$  into Eq. (8.24), then  $Q_i$  becomes

$$Q_i = d_{Ri} \cdot d_{Ri} + d_{Ii} \cdot d_{Ii} - \sum_{u=1}^{m'} \hat{B}'_{ui} T_{ui} \quad (8.39)$$

where the quadratic term exactly cancels one of the linear terms.

## 8.4 Levenberg-Marquardt And Newton-Raphson

The search algorithm used by Bayes Analyze is a modified Levenberg-Marquardt algorithm that combines features of Levenberg-Marquardt and Newton-Raphson. Levenberg-Marquardt algorithms are themselves modified versions of the Newton-Raphson algorithm, where an approximation has been used in calculating the second derivatives of  $\chi^2$ . Newton-Raphson is a technique used in

finding the roots of a function. For example to find a root of a function  $f(x)$ , one computes the first derivative of  $f(x)$  with respect to  $x$  at some value of  $x$  which we will call  $x_i$ :

$$\frac{d}{dx_i} f(x_i) \equiv \dot{f}(x_i). \quad (8.40)$$

The Newton-Raphson estimate of the correction to  $x_i$  is then given by

$$\delta_i = -\frac{f(x_i)}{\dot{f}(x_i)} \quad (8.41)$$

and one updates the value of  $x_i$ :

$$x_{i+1} = x_i + \delta_i. \quad (8.42)$$

One then takes this new value of  $x_{i+1}$  and uses it as an initial guess in Eq. (8.40). The convergence of Newton-Raphson is quadratic, so the iteration procedure will usually converge to a root very quickly. The above procedure is discussed [50] and we refer you that discussion for more on Newton-Raphson and how to extend it to multiple variables.

Levenberg-Marquardt algorithms minimize  $\chi^2$ ,

$$\chi^2 \equiv \sum_{i=1}^N \frac{[d_i - y(t_i, \Omega)]^2}{\sigma^2} \quad (8.43)$$

they do not find roots; where in this equation,  $d_i$ , represents a single data item sampled at time  $t_i$ ,  $y(t_i, \Omega)$  is a model as a function of a set of  $\Omega$  parameters and  $\sigma$  is the standard deviation of the noise in the current data set. To minimize  $\chi^2$  one must find the values of the  $\Omega$  parameters for which the first derivative is zero. Consequently, in a Levenberg-Marquardt algorithm it is first derivative of  $\chi^2$  that plays the part of  $f(x)$  in a Newton-Raphson algorithm. The basic ideal is essentially the same as the Newton-Raphson algorithm. One has a function,  $\chi^2$ , that one wishes to maximize. Taking the first partial derivative of  $\chi^2$  with respect  $\omega_j$ , one obtains:

$$\beta_j \equiv -\frac{1}{2} \frac{\partial}{\partial \omega_j} \chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^N [d_i - y(t_i, \Omega)] \frac{\partial y(t_i, \Omega)}{\partial \omega_j} \quad (8.44)$$

where this is a definition of  $\beta_j$ . Next one defines the second derivatives:

$$\alpha_{jk} \equiv \frac{1}{2} \frac{\partial^2}{\partial \omega_j \partial \omega_k} \chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^N \left[ \frac{\partial y(t_i, \Omega)}{\partial \omega_j} \frac{\partial y(t_i, \Omega)}{\partial \omega_k} - \left\{ [d_i - y(t_i, \Omega)] \frac{\partial^2 y(t_i, \Omega)}{\partial \omega_j \partial \omega_k} \right\} \right] \quad (8.45)$$

which is a symmetric matrix of second derivatives. So far nothing in this calculation distinguishes a Levenberg-Marquardt algorithm from a Newton-Raphson algorithm. There are two distinguishing characteristic. First, in a Levenberg-Marquardt algorithm the term in curly brackets is assumed small when compared to the other terms in this equation, consequently,

$$\alpha_{jk} \approx \frac{1}{\sigma^2} \sum_{i=1}^N \left[ \frac{\partial y(t_i, \Omega)}{\partial \omega_j} \frac{\partial y(t_i, \Omega)}{\partial \omega_k} \right]. \quad (8.46)$$

To motivate this approximation note that the term,  $d_i - y(t_i, \Omega)$ , is essentially the residual at time  $t_i$  and this residual should alternate between positive and negative values. Consequently, the sum over time should on average be small. The approximation made in Eq. (8.46) has profound implications from a computational stand point. Note that the approximation does not make use of any second derivatives; rather it approximates the second derivative by a product of two first derivatives. Computationally, this means that the program that implements this calculation does not have to be programmed with second derivative information and this can be a major simplification in the calculations.

The second distinguishing characteristic in a Levenberg-Marquardt algorithm is  $\alpha_{jk}$  is modified by multiplying the diagonal of this matrix by a control parameter:

$$\alpha'_{jk} = \alpha_{jk}(1 + \lambda\delta_{jk}) \quad (8.47)$$

where  $\delta_{jk}$  is the Kronecker delta function, and  $\lambda$  controls the algorithm by making it possible to switch back and forth between a steepest descents algorithm and a Newton-Raphson algorithm. When  $\lambda$  is zero the algorithm is a Newton-Raphson algorithm and when  $\lambda$  is large the algorithm simply follows the gradients, i.e., steepest descents. The corrections to the  $\Omega$  parameters are then given by

$$\sum_{k=1}^{\nu} \alpha'_{jk} \delta_k = \beta_j \quad (8.48)$$

where  $\delta_k$  is the correction for the  $k$ th parameter. This simple set of linear algebraic equations can be solved for the  $\delta_k$  and by iterating the procedure the global minimum of  $\chi^2$  can be located. As noted, the parameter  $\lambda$  is used to control the algorithm. If  $\lambda$  is large the  $\alpha'_{jk}$  matrix is diagonally dominate and the Levenberg-Marquardt algorithm just follows the gradients. However, if  $\lambda$  is small the algorithm follows the full curvature of the space, i.e., Newton-Raphson. In Bayes Analyze, if a step increases the joint posterior probability,  $\lambda$  is decreased by about 10%, and smaller  $\lambda$  means the convergence is speeds up. However, if a step decreases the posterior probability, the value of  $\lambda$  is increased by 10% and the corrections to the  $\Omega$  parameters are recomputed. Using these new corrections, the joint posterior probability is computed and if the posterior probability increases, the new values of the  $\Omega$  parameters are keep and the iteration cycle is repeated until the first derivatives are zero.

Bayes Analyze searches in the logarithm of the joint posterior probability for the  $\Omega$  parameters, it does not search in  $\chi^2$ , consequently Bayes Analyze is not a Levenberg-Marquardt algorithm. None the less, the above discussion can be directly applied to the Bayesian calculations done because the logarithm of the joint posterior probability has components that can be written in the form of a  $\chi^2$ . Because integrating out the amplitudes just constrains the amplitude to the value that maximizes the joint posterior probability, the  $Q_i$  function can also be written as:

$$Q_i \equiv \sum_{l=1}^{m'} \gamma^2 \hat{B}'_{li}{}^2 + \sum_{j=1}^N [d_{Ri}(t_j) - \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ri}(t_j)]^2 + \sum_{j=1}^N [d_{Ii}(t_j) - \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ii}(t_j)]^2 \quad (8.49)$$

where this is just Eq. (8.20) with the amplitudes replaced by there peak joint posterior probability estimates. As explained, a Levenberg-Marquardt algorithm is essentially a Newton-Raphson method specialized to search in  $\chi^2$ . The reason we rewrote the statistic this last time, Eq. (8.49), was to express  $Q_i$  in a  $\chi^2$  like form. Now  $Q_i$  is not strictly  $\chi^2$  and the statistic used in the search isn't either.



None the less we are going to define a  $\chi^2$  term and then discuss how Bayes Analyze implements it's search using this definition:

$$\chi^2 \equiv \log P(\Omega|I) + \sum_{i=1}^n \frac{Q_i}{\sigma_i^2} \quad (8.50)$$

where  $P(\Omega|I)$  is the prior probability for the nonlinear parameters and  $Q_i$  is given by Eq. (8.49). This definition of  $\chi^2$  is essentially the logarithm of the joint posterior probability for the  $\Omega$  parameters independent of the amplitudes where a some constants have been dropped.

To implement a Levenberg-Marquardt algorithm, one must define both an  $\beta_j$  and an  $\alpha_{jk}$ , in the Bayes Analyze program these are defined as:

$$\begin{aligned} \beta_j &\equiv -\frac{1}{2} \frac{\partial}{\partial \omega_j} \chi^2 \\ &= -\frac{1}{2} \frac{\partial}{\partial \omega_j} \log P(\Omega|I) - \frac{1}{2} \gamma^2 \sum_{l=1}^{m'} \frac{\partial}{\partial \omega_j} \hat{B}'_{li}{}^2 \\ &\quad + \sum_{i=1}^n \frac{1}{\sigma_i^2} \sum_{j=1}^N \left[ d_{Ri}(t_j) - \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ri}(t_j) \right] \left[ \frac{\partial}{\partial \omega_j} \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ri}(t_j) \right] \\ &\quad + \sum_{i=1}^n \frac{1}{\sigma_i^2} \sum_{j=1}^N \left[ d_{Ii}(t_j) - \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ii}(t_j) \right] \left[ \frac{\partial}{\partial \omega_j} \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ii}(t_j) \right] \end{aligned} \quad (8.51)$$

and  $\alpha_{jk}$  is defined using the same approximation used in a Levenberg-Marquardt algorithm:

$$\begin{aligned} \alpha_{jk} &\equiv \frac{1}{2} \frac{\partial^2}{\partial \omega_j \partial \omega_k} \chi^2 \\ &\approx \frac{1}{2} \frac{\partial^2}{\partial \omega_j \partial \omega_k} \log P(\Omega|I) + \frac{1}{2} \gamma^2 \sum_{l=1}^{m'} \frac{\partial^2}{\partial \omega_j \partial \omega_k} \hat{B}'_{li}{}^2 \\ &\quad + \sum_{i=1}^n \frac{1}{\sigma_i^2} \sum_{\nu=1}^N \left[ \frac{\partial}{\partial \omega_j} \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ri}(t_\nu) \right] \left[ \frac{\partial}{\partial \omega_k} \sum_{l=1}^{m'} \hat{B}'_{li} M'_{Ii}(t_\nu) \right]. \end{aligned} \quad (8.52)$$

With this definition of  $\beta_j$  and  $\alpha_{jk}$  it is now possible to implement a Levenberg-Marquardt algorithm. This algorithm has a number of interesting properties not present in a standard Levenberg-Marquardt algorithm. For example, by including the prior probabilities for the nonlinear parameters,  $P(\Omega|I)$ , one can effectively bound a parameter to positive values and we do this using a positive prior, Eq. (3.7). Second, by taking into account the presence of multiple data sets with differing standard deviations, this statistic allows many different kinds of data to be combined. Finally, by marginalizing out the amplitudes, the Levenberg-Marquardt algorithm does not search on the amplitudes. This turns out to be very significant when multiple resonances and data sets are present. In NMR experiments, one often takes multiple data sets. These data sets measure the same quantity, but with some parameter varied, often a delay time or a gradient. The amplitudes of the

resonances from these multiple data sets are then used in additional analysis, such as inversion recovery experiments. However, the number of these data sets can be large, in the hundreds or even thousands. If we had not marginalized out the amplitudes, a 10 resonance problem, looking for common frequencies in one thousand free induction decays would have 10 frequencies, 10 decay rate constants, perhaps one or two phase parameters, and 10,000 or 20,000 amplitudes depending on how the problem is parameterized. For all practical purposes, the problem would be intractable because all of the frequencies, decay rate constants and amplitudes are correlated, i.e., one must search on all of these parameters simultaneously. However, by marginalizing out the amplitude the resulting search only involves 20 or so parameters. Consequently, marginalization is a key feature in jointly analyzing arrayed experiments.

Of course, implementing this calculation is a tricky problem. We have not, for example, talked about how one computes the derivatives of the amplitudes, nor how to estimate the noise standard deviations. Its enough to know that the derivatives of the amplitudes and the standard deviations can be computed at each step in the calculation. Consequently, when the search completes, one has the peak value of the joint posterior probability, the zero first derivatives, a very good approximation to the second derivatives, and estimates of the standard deviation of the noise in each data set. Together these can be used to write an Gaussian approximation to the joint posterior probability for the parameters. The approximate joint posterior probability for the parameters can be used to estimate the parameters using mean  $\pm$  standard deviation estimates. Additionally, the posterior probability for the model can be approximated as the integral over the approximate joint posterior probability for the parameters and this approximate joint posterior probability for the model can be used to determine when the algorithm should stop. So a general description of the Bayes Analyze algorithm is:

1. Using the Fourier transform of the residuals [7], compute the odds ratio [5] that any given peak in the spectrum is a resonance.
2. If no peak in the odds ratio rises above the noise threshold, exit the analysis. Otherwise, using the single most probable resonance frequency from the signal detection algorithm, postulate a resonance model containing one more resonance.
3. Using the Levenberg-Marquardt search algorithm optimize the parameters in this model.
4. Using the peak, the covariance matrix and the estimated noise standard deviations, approximate the joint posterior probability for the parameters, then using this approximation estimate the parameters and the posterior probability for the model.
5. Output these parameter estimates to the various Bayes Analyze output files.
6. If the posterior probability for the model decreases or the maximum number of resonances has been added to the model, finish up the analysis and exit Bayes Analyze.
7. If the posterior probability for the model increased, go back to the first item and repeat this procedure until either no additional candidate resonances can be found or the posterior probability for the model decreases.

## 8.5 Outputs From The Bayes Analyze Package

Unlike the other packages, the Bayes Analyze package is not a single program, rather it is 4 different programs that are run one right after the other and a fifth that is run on demand. The main program is called Bayes Analyze and it is the frequency finding program. There are three programs, called Bayes Summary 1, 2 and 3, that take the outputs from Bayes Analyze and reformat them into three summary reports. Finally, there is a modeling program, called Bayes Model, that takes the output from Bayes Analyze and creates a model of the input Fid. The Bayes Analyze program does not use Markov chain Monte Carlo to do its calculations. Consequently, the outputs from the Bayes Analyze package are completely different from the outputs from MCMC packages. Thus to view these outputs, a different set of output widgets are needed. The interface has two different output areas on the text results viewer, one for MCMC packages and one for the Bayes Analyze package.

The Bayes Analyze output files and the summary reports are numbered with a suffix. For example the main Bayes Analyze output file is named “bayes.output.nnnn” where the “nnnn” is related to the From, To, and By fields described earlier, see page 158. When Bayes Analyze process a set of Fids, it begins with Fid number “From” and analyzes “By” number of Fids. The output from this analysis is written into a series of files labeled “name.From” where “name” is one Bayes Analyze output files and “From” is the “From” Fid number. After Bayes Analyze finishes this analysis, it increments the “From” Fid number by “By” Fid number and then checks to see if it has reached Fid “To”. If not it repeats the analysis using this new “From” Fid number, again writing out files named “name.From.” This loop continues until all of the Fids have been processed. Some of these outputs are available through the interface, and some only by looking at them within the WorkDir. Here is a brief description of the more important input and output files used by the Bayes Analyze package:

**bayes.params** file is written by the interface and serves as the input parameter file to the Bayes Analyze program. It contains various parameter settings and the initial model to be processed.

The parameter file is divided into three general sections, a header, the global parameters, and the resonance parameters. Each of these three sections is describe in detail in Subsections 8.5.1.1, 8.5.1.2 and 8.5.1.3.

**bayes.params.nnnn** are files written by Bayes Analyze for each individual sub-analysis. They are updated version of bayes.params, updated in the sense of containing the maximum joint posterior probability estimates of the frequencies, decay rate constants, etc. Additionally, the “From,” “To,” and “By” numbers are updated to reflect this particular sub-analysis. For example, if the current analysis is running Fids 5 through 10, then a parameter file, bayes.params.0005, is written and used by the interface to mark the location of the various resonances. After Bayes Analyze completes the interface reads the bayes.params.nnnn file and the updated model is then available for further processing within the interface. Note that which bayes.params.nnnn file is read by the interface depends on which Fid is to be modeled. That is to say if you are planning to model Fid 25 from an array, the interface will figure out which bayes.params.nnnn file contains the parameters for Fid 25 and then read those parameters. Again, each of these three sections is describe in detail in Subsections 8.5.1.1, 8.5.1.2 and 8.5.1.3.

**bayes.output.nnnn** are the main output files from Bayes Analyze and they contain detailed information about each model Bayes Analyze processed. The description of this file may be found in Subsection 8.5.3.

**bayes.probabilities.nnnn** contain a one line entry for each model processed by Bayes Analyze. These lines contain the base 10 logarithm of the posterior probability for each model processed by Bayes Analyze. Note the file is cumulative and results from multiple runs of Bayes Analyze may be viewed. For a complete description of this file, see Subsection 8.5.4.

**bayes.log.nnnn** is a detailed log of each step taken in the searching algorithm for each model processed by Bayes Analyze, see Subsection 8.5.5 for a complete description of this file.

**Bayes.accepted** is used by the interface to show the current status of an analysis. For example, if you are running Bayes Analyze when the “Status” button is activated, the interface will display the current contents of the Bayes.accepted file. This file is a recent addition to Bayes Analyze and is a simple copy of the bayes.status.nnnn file. It was added to Bayes Analyze so the new interface could determine the status of the current analysis. For a complete description of this file see Subsection 8.5.6.

**bayes.status.nnnn** contains the status of analysis nnnn. Usually by the time this file can be viewed, it simply indicates the analysis is completed. As noted above this file is also written into Bayes.accepted and the Bayes.accepted file may view while Bayes Analyze is running. For a complete description of this file see Subsection 8.5.6.

**bayes.model.nnnn** is used by Bayes Model to produce a maximum joint posterior probability model of an Fid. As one of its outputs Bayes Analyze writes the parameters that maximize the joint posterior probability into the model file. The model file, unlike the bayes.params.nnnn files, contain both the frequency and amplitude estimates. For a complete description of this file see Subsection 8.5.7.

**bayes.summary1.nnnn** The summary files, bayes.summary1.nnnn, bayes.summary2.nnnn and bayes.summary3.nnnn, are produced by Bayes Analyze package whenever you run the analysis. The summary1 or “Best Model” report contains parameters from the most probable model. For a complete description of the summary1 report see Subsection 8.5.8.

**bayes.summary2.nnnn** The summary2 or “Best Summary” report is a highly condensed version of the summary1 report. However, because each line displays the information for a single resonance, the report is available only for nonarrayed free induction decays. For a complete description of the summary2 report see Subsection 8.5.9.

**bayes.summary3.nnnn** Finally, the summary3 or “Best Regions” report displays the intensity (normalized and total), contained within a series of regions. These regions are defined by using the “Options/Set Regions” button. Use of this report and how to set the regions is described further in Subsection 8.5.10.

### 8.5.1 The “bayes.params.nnnn” Files

Bayes Analyze reads the parameter file and uses the information in the parameter file to direct its operations. The parameter file is used to communicate between the interface and the Bayes Analyze program. The name of the parameter files may be either “bayes.params” or “bayes.params.nnnn.” If the number is not present the file was output from the interface and is a Bayes Analyze input file. If the number is present the file was output from Bayes Analyze package and serves as input to the

interface. The parameter file consists of three general sections, a header, the global parameters, and the resonance model. We describe each of these sections in the following.

### 8.5.1.1 The Bayes Analyze File Header

The Bayes Analyze package writes a file header into many of its output files. This header describes the analysis, and serves as a record of the analysis. An example header is shown in Fig. 8.3. In the Bayes Analyze package, blank and comments lines (lines starting with “!”) are ignored and may appear anywhere within the parameter or model files. As you can see, the first three lines in the file header are comments. Comments are ignored by Bayes Analyze and Bayes Model. Also, file names appearing in Fig. 8.3 are referenced from the directory where Bayes Analyze is to be run. In the interface when Bayes Analyze is run, the program runs in the current WorkDir so the file and directory names appearing on Lines 05–08 are referenced from that working directory. Here is a description of the parameter file header:

**Line 01** is a comment and indicates the name of the Fid loaded into the current WorkDir.

**Line 02** is a comment containing the name of the input file type, in this case a “Parameter File.”

**Line 03** is a comment containing name of the user who created the file and the date and time.

**Line 04** indicates the number of parameters that are to follow, in this example there are 32. Each line contains a description followed by the value of the parameter. Some of the parameters are text strings and some of them are numbers. Text strings start in column 22 and extend to the right. Numerical values may appear anywhere from column 22 through 32. Bayes Analyze recognizes parameters by the labels that appear in column 5, these labels must appear as shown, including capitalization. The ordering of the parameters is generally not critical, although keeping them in the order shown is advisable. The numbers on the left and the column headings were added for reference and do not appear in the file.

**Line 05** is the file version number and, as Fig. 8.3 implies. This parameter must follow the configuration parameters. The programs have the ability to read old file versions. We have included this parameter because output files from a Bayesian Analysis may be saved – sometimes for long periods of time and we did not want to lose the ability to read and model older versions of these files.

**Line 06** is the fully qualified name of the input Fid file. As noted earlier, when Bayes Analyze is run it is run in the current WorkDir so the fid file must fully qualified from WorkDir. Consequently, when the parameter file is written by the interface this field will always be fid/fid, however any fully qualified name is allowed.

**Line 07** is the fully qualified name of the input procpa file.

**Line 08** is the name of the current analysis directory. This is the fully qualified name of the directory in which the output from Bayes Analyze is to be written. When the parameter file is written by the interface, the analysis directory is “BayesAnalyzeFiles”.

**Line 09** is the name of the directory where Bayes Model is to create the model. This is Bayes.model.fid when the interface generates this file, but it need not be that way and may be set to any value if Bayes Analyze is run manually.

Figure 8.3: The Bayes Analyze File Header

```

          1          2          3          4          5          6
123456789012345678901234567890123456789012345678901234567890123

01! File: "~larry/bayesian/Bayes.test.data/BayesAnalyze/MTz_C13_irradR_20C.fid"
02! Bayesian Analysis Input Parameter File
03! Created 07-May-2001 15:14:21 by Larry

04 31 Configuration Parameters

05 File Version = 3.000
06 Fid File = fid/fid
07 Procpa File = fid/procpa
08 Analysis Dir = BayesAnalyzeFile
09 Model Dir = Bayes.model.fid
10 Model Dir Org = DATA
11 Units = PPM
12 Activate Shims = YES
13 Activate Delay = YES
14 Data Type = VNMR
15 Noise = NO
16 Output = FULL
17 Default Model = (CP)
18 First Fid = 1
19 Last Fid = 1
20 No Fids = 1
21 Total Points = 1024
22 Complex Points = 512
23 Noise Start = 0
24 Model Points = 1024
25 Model Fid = 1
26 Prior Odds = 0.00000
27 Sampling Time = 0.29900
28 Spec Freq = 300.00000
29 User Reference = -2.84349
30 True Reference = -2.84349
31 Total Models = 3
32 Shim Order = 7
33 Max Freqs = 10
34 Max Candidates = 1
35 Default Lb = 1.08755

```

Figure 8.3: The Bayes Analyze File Header is present in most output files from the Bayes Analyze package. This header contains all of the parameters used to run Bayes Analyze. Among other things, it contains the name of the Fid file, the procpa and the output directory where the model is to be written. It also specifies the directory organization, the units and many other parameters needed to run Bayes Analyze. See text for a more extensive description of the file header.

Figure 8.4: The bayes.noise File

```
1 3.22299585E+01
2 3.19859327E+01
3 3.22769104E+01
4 3.18850559E+01
5 3.21451727E+01
6 3.19790225E+01
7 3.20911816E+01
8 3.19975486E+01
9 3.19348404E+01
```

Figure 8.4: The file bayes.noise is a two column Ascii file generated by the interface when a noise region is set. The first column is the number of the Fid and the second is the root mean-square data value in the noise region, and is an estimate of the standard deviation of the noise in the noise region. Note that the bayes.noise contains one entry for each trace in the Fid, i.e., the “From,” “To” and “By” fields are not used when generating the bayes.noise file.

**Line 10** specifies the organization of the directories. The new interface sets this to “DATA” meaning that the output directory is a standard Fid directory. However, this can also be set to “VNMR,” in which case the Bayes Analyze expects the structure of the analysis directory to be in VnmrJ format.

**Line 11** specifies the units to use in the analysis. The valid entries are “PPM,” or “HERTZ”.

**Line 12** specifies whether the lineshape model is to be used. The valid values are “YES” or “NO” and these mean yes the lineshape is to be expanded, or no it is not. If lineshape, or shimming model is activated then the shimming expansion order, shown on Line 32, must be specified as greater than one.

**Line 13** is used to activate the first order phase correction. Note that the new interface does not allow you to modify this parameter, it is automatically activated by the interface.

**Line 14** indicates the type of input data to be processed. The new interface sets this to “VNMR” indicating that the input file is a VNMR Fid.

**Line 15** indicates whether an estimate of the root mean-square estimate of noise standard deviation is available. The valid values for this indicator are “YES” or “NO.” These indicate yes a noise measurement is available or no a noise measurement is not available. If a noise measurement is available the Bayes Analyze program will attempt to read the “bayes.noise” file. An example of this file is shown in Fig. 8.4. This noise file is two columns, the first being the number of the free induction decay and the second being the estimated root mean-square noise value. This file must be present if the noise used indicator is “YES.” If used, there must be one entry in this file for every free induction decay in the input data. This file is built by the menus using the Bayes Noise program.

**Line 16** indicates the format of output written to standard out. This entry may be be “FULL,” “BRIEF” or “NONE.” This option is set to “BRIEF” by the new interface.

Table 8.2: Bayes Analyze Models

Model Name	Description	Resonance
Real Constant	A constant in the real channel	no
Imaginary Constant	A constant in the imaginary channel	no
DC Offset In Both	The same constant in both channels	no
First Point Problem	Model the data as if the first point were in error	no
(CP)	The Correlated Phase model	yes
(UP)	The Uncorrelated Phase model	yes

Table 8.2: The Bayes Analyze package currently supports the 6 models listed here. The first column is the model name. The second column is a longer description of the model. Finally, the third column indicates whether the model is a resonance model or not.

**Line 17** is used to set the resonance model used in the automatic mode of Bayes Analyze. Table 8.2 is a list of all the model names currently in use in the Bayesian Analysis package. There are two resonance models in this table, these are the Correlated Phase model, “(CP)” and the Uncorrelated Phase model, “(UP)”.

**Lines 18** through 20 are the first, last, and number of Fids respectively. Bayes Analyze will begin processing the data starting with the first Fid. It processes number of Fids at a time looking for common resonances. After processing a given block of free induction decays, the program will increment first Fid by number of Fids and then proceed to process this block of Fids. This process continues until all indicated Fids have been processed. As noted in Section 8.5.5 the number of the first Fid processed in a cycle is appended to all of the output files from Bayes Analyze.

**Lines 21** is the total and number of data values to be processed by Bayes Analyze. Note this is the sum of the total real data values plus the number of complex data values, i.e., 2 times the total complex data values.

**Lines 22** are the total number of complex data values to be processed by Bayes Analyze. The complex number of data values is redundant in that it is just half the total number of data values.

**Line 23** is the starting position of the noise region. This field is not used by Bayes Analyze. This field was placed in the parameter file so that it could be saved when an analysis is saved.

**Line 24** is the number of model points (real plus complex) in the output modeled Fid. When the model program, Bayes Model, creates a model of a Fid it creates an arrayed Fid in the model experiment. The first element of this array is a copy of your data, if number of model points is less than “np” only part of your data is copied to the output Fid. If it is greater than “np”, the data are copied and then zero padded. The default is for this value to be equal to “np”.

**Line 25** is the number of the free induction decay to model. If the user does not specify the free induction decay to model it defaults to the free induction decay that was last displayed in the experiment. If no free induction decay has been displayed it defaults to the first free induction decay processed by the user.



- Line 26** the prior odds is automatically set to zero by the interface and should not be changed.
- Line 27** is the sampling time of the data being processed. This field is the same as the VNMR parameter “at” unless you are using the signal region, then its the sampling time for the data being processed. If you are processing 25% of the free induction decay then this field is 25% of “at” etc.
- Line 28** is the spectrometer frequency (the “reffrq” parameter) and is used in the conversion to and from ppm. Internally the Bayes Analyze program uses radian as the frequency units. When either Bayes Analyze or Bayes Model reads the input parameter file the programs determine what units are being used and makes the appropriate conversions to internal units.
- Line 29** is the reference frequency use by the Bayes Analyze. This parameter has no exact correspondence to any VNMR parameter. If the number on line 29 is called “ref,” then it is computed as:
- $$\text{“ref”} = -([\text{“rfp”} - \text{“rfl”}] + [\text{“sw”}/2.0]) \quad (8.53)$$
- where “rfp,” “rfl” and “sw” are the reference point, the reference line, and the sweep width respectively. For ppm, “ref” is then divided by the reference frequency “reffrq.”
- Line 30** is the true reference. The meaning of the field is exactly the same as the user reference.
- Line 31** is a count of the total model components in the input model to be analyzed in slave mode. A model component might be a singlet, a multiplet, or a constant.
- Line 32** is the order of the shimming expansion being done on the lineshape. This field is ignored unless the activate shims indicator is “YES”. An example of setting the activate shims indicator is shown on Line 11 in Fig. 8.3. When this indicator is set, the expansion order must be 3, 5, or 7.
- Line 33** is the maximum number of new resonances that may be incorporated into the model on this run of Bayes Analyze.
- Line 34** is the maximum number of resonances that may be incorporated from one run of the signal detection algorithm. This number is set to 1 by the interface and should not be changed.
- Line 35** is current value of “lb.”

#### 8.5.1.2 The Global Parameters

Immediately following the configuration parameters are the global parameters. An example of the global parameter section is shown in Fig. 8.5. The numbers on the left as well as the column indicators on the top have been added for reference purposes. In the example shown there are 10 global parameters, the maximum. When these parameters appear in the parameter file, they must appear in the order shown.

**Line 01** is a count of the number of global parameters.

**Line 03** is the constant phase used by Bayes Analyze.

Figure 8.5: Bayes Analyze Global Parameters

```

0          1          2          3          4          5
1234567890123456789012345678901234567890123456789012
01  10 Global Model Parameters
02
03  Center Phase = 219.17041042
04  Time Delay   = 0.00830983
05  Shim Delta   = 36.923702254267 2.50000000000000
06  R Minus 3 = 0.0121083700262
07  R Minus 2 = 0.0123818187190
08  R Minus 1 = 0.2878006370120
09  R Center   = 0.2881405188680
10  R Plus 1   = 0.1340506287029
11  R Plus 2   = 0.1334510175742
12  R Plus 3   = 0.1320670090977

```

Figure 8.5: The global parameter immediately follow the header. There are at most 10 global parameters. Here we have exhibited all 10 of these. These parameters include the two phasing parameters, the spacing of the Lorentzian's in the shimming model (Shim Delta), and the relative amplitudes of the Lorentzian's in the expansion.

**Line 04** is the first order phase parameters used for correlated resonances. VNMR allows spectra to have both a constant and a linear frequency dependent phase. These are specified by the left and right phases. These parameters are natural for spectra but are clumsy in the time domain. The Bayesian Analysis package expresses this phase variation as a constant phase with a time delay.

**Lines 05** through 12 concern the expansion of the lineshape in Lorentzian's. These parameters are not present unless the shimming model is turned on and Bayes Analyze has been run previously. Line 05 is the shim delta: it is the separation frequency between the Lorentzian's in the shimming model expansion of the lineshape. It is the exact analogue of the  $J$ -coupling constants in multiplets. There are two entries present on line 05: its current value and its initial value. Its initial value is set to "lb"/(expansion order).

**Lines 06** through 12 are the relative amplitudes assigned to the Lorentzian's that make up the lineshape expansion. If the shimming order is set to three, then there are three Lorentzian's for each peak in the spectrum. The total area under these peaks is required to add up to one, so there are only two independent amplitudes; the third is redundant. By convention we take the redundant amplitude to be the center amplitude. In Fig. 8.5 there are 7 relative amplitudes shown, i.e., a seventh order expansion of the lineshape is being done. Each relative amplitude is labeled "R" plus or minus something. For a seventh order expansion the largest relative amplitude is the center relative amplitude, Line 04. Line 06 is labeled "R Minus 3" and this means the third and smallest relative amplitude whose frequency is given by  $\omega - 3\delta_{\text{shim}}$  where  $\omega$  is the resonance frequency. Similarly, "R Plus 2" means the relative amplitude of the shimming frequency given by  $\omega + 2\delta_{\text{shim}}$ .

Figure 8.6: The Third Section Of The Parameter File

```

01! # Name
02 1 Real Constant
03
04! # Name
05 2 Imaginary Constant
06
07! # Name          Order      Frequency      Init Value
08 3 (CP) Triplet  (3,1) 0      1.2913619868959 1.2900933000000
09
10                      Decay Rate      Init Value
11                      1.0405457695713 1.0875500000000
12
13                      Primary J      Init Value
14                      6.9823931646449 6.6850000000000
15
16                      Secondary J      Init Value
17                      0.0000000000000 0.0000000000000

```

Figure 8.6: The third section of the parameter file contains the model to be used in the analysis. The first part of this model is the constant models. Here there are two constant models, Line 02, and Line 05. Lines 08 through line 17 is an example of a triplet resonance model. In the actual parameter file all the parameters for a resonance are on a single line, so lines 10 through 17 should be part of line 08 and each field should occupy exactly 15 positions with an additional space separator. However, I broke that line apart so the entire contents of the the line would be visible.

### 8.5.1.3 The Model Components

The parameter file consists of the header, the global parameters, and then the individual resonance models. Figure 8.6 is an example containing some constant models as well as a resonance model. In this example there are three model components. Each model component consists of a blank line, a comment line used as a header, here is a description of these lines:

**Line 01** is the comment line for the first model.

**Line 02** is the description of the model. In this case the model is a real constant model and constant models do not have parameters, so there is nothing else on this line except the model number and the model name.

**Line 03** is a blank line.

**Line 04** is the model header.

**Line 05** is the description of the imaginary constant model. The two fields are simply the model number and the model name.

**Line 06** is a blank line.

**Line 07** 10, 13 and 16 is the model header for the third model. This header is for a multiplet and when this header was generated it labels the most complicated model used by Bayes Analyze, a multiplets of multiplets. Also note, that lines 10, 13 and 16 should be part of line 07 with each field on the header taking 15 spaces with an additional space separator.

**Line 08** 11, 14 and 17 are the parameters associated with a multiplet. In the real parameter file lines 08, 11, 14 and 17 are on the same line with each item taking exactly 15 spaces and is space separated from its neighbor. Here we have separated them so that they have not dropped off the right-hand margin.

**3** is the model number.

**(CP)** is the resonance type, in this case the resonance is a Correlated Phase resonance. See Table 8.2 for a list of the valid model type.

**triplet** is the resonance name, i.e., single, doublet, triplet, etc.

**(3,1)** is the order of the primary and secondary multiplets. In this case a triplet of singlets.

**0** the zero following the multiplets order is the relative phase of the multiplet. The relative phase may have values 0, 90, 180, and 270. The new interface sets this phase to zero. Here zero means that the resonance is in phase with the other resonances.

**1.2913619868959** is the frequency of the multiplet in currently selected units, in this case PPM.

**1.2900933000000** is the initial value of this frequency prior to optimization.

**Line 09** is a blank line.

**Line 10** is header for the decay rate constant.

**Line 11** is the decay rate constant and its initial value.

**Line 12** is a blank line.

**Line 13** is header for the primary J-coupling constant.

**Line 14** is the primary J-coupling constant.

**Line 15** is a blank line.

**Line 16** is header for the secondary J-coupling constant.

**Line 17** is the secondary J-coupling constant. Note these entries for the secondary coupling constant are not present unless you have a multiplet of multiplets.

### 8.5.2 The “bayes.model.nnnn” Files

The bayes.model.nnnn file is a copy of the bayes.params file that contains the output amplitudes. These amplitudes follow each frequency component and are written in this form:

```
!                               Fid           Cos Amplitude   Sine Amplitude
                               1             13413.806578806  411.29599022536
```

when uncorrelated resonances are modeled and in this form:

!	Fid	Amplitude
	1	4537.7682653001
	2	4571.1890384819
	3	4450.1289475121
	4	4577.9827364511
	5	4537.7818733113

when correlated resonances are modeled. The Fid number on this line indicates that there were 5 different fid's or traces in the data set analyzed. When Bayes Model is activated, it finds the resonances and amplitudes for the trace being modeled and then generates an output Varian Fid that contains the original trace as trace 1. Trace 2 is the model built by Bayes Model and is a sum of the sinusoids in the Bayes Model file. Trace 3 is the residual, the difference between the data and the model Fid. As Fid data are complex all three of these three traces are complex time domain Fid data. Finally, if Bayes Analyze found 10 resonances then following each of these first 3 traces there will be 10 additional traces, one containing a model of each of the 10 sinusoids. The output Fid model is used by the interface whenever the Fid Model viewer is activated.

### 8.5.3 The “bayes.output.nnnn” File

The output file is the main output report from the Bayes Analyze program. The output file is name “bayes.output.nnnn” where, as explained in Section 8.5.5, the number nnnn is the number of the first free induction decay who's analysis is contained in this output file. This file contains the detailed output from each model processed by Bayes Analyze. Because it contains so much detail, it can be rather difficult to follow and should be printed only when this information is needed. When the detail is not needed, the summary report should be used. The summary report is essentially the section out of the output file pertaining to the model with maximum probability. For more information on the summary report and the difference between these two reports see Sections 8.5.8 and 8.5.9.

When Bayes Analyze runs in its automatic mode many different models may be analyzed. At the end of the analysis on each model, the optimized values of the model parameters are written to the output report. In this section we describe this report in detail. The output reports starts with a listing of the configuration parameters associated with the analysis. An example of these configuration parameters is given in Fig. 8.3 and an explanation of these configuration parameters is given in Section 8.5.1.1. There are two small differences between what is shown in Fig. 8.3 and what is printed in the output report. These difference consist of a header a the beginning of the report. This header merely indicates that the initial configuration parameters follow. The other difference is at the end of the configuration parameters. The output file contains a list of the resonances processed in the slave mode. Figure 8.7 is a small sample of what this listing might look like. There are two multiplets, one triplet and one quartet in the initial model. The notation used to indicate the triplet and quartet is the same as that explained in Section 8.5.1.3. This is followed by the resonance frequency and the primary and secondary coupling constants. These coupling constants are zero if they are not used.

One or more sets of detailed information follows the initial model. The detailed information consists of the output from the signal detection calculation, output from the optimization of the joint probability for the parameters, the output from the model selection calculation, and the normalized

Figure 8.7: Example Of An Initial Model In The Output File

```
Initial Model:
# Order      Frequency      Decay Rate      Primary J      Secondary J
1 (3,1) 0    1.2913982000000 0.8574675300000 6.9821668000000 0.0000000000000
2 (4,1) 0    4.8648472000000 0.8881744100000 6.9795215000000 0.0000000000000
```

Figure 8.7: This example shows two resonances, one triplet of singlets, the (3,1), and one quartet. The parameters, other than the amplitudes, follow. Here there are nontrivial values of the frequencies, the decay rate constants and the primary  $J$  coupling constant.

Figure 8.8: Base 10 Logarithm Of The Odds

```
Base 10 Log Evidence for The First Frequency is: 49.6
-----
```

Figure 8.8: After the signal detection algorithm runs, the base 10 logarithm of the odds in favor of a resonance are printed to the output file. Here the odds are  $10^{49.6} : 1$  in favor of a sinusoidal signal component.

probability for the models analyzed so far.

The signal detection output consists of only single line preceding the dashed line, see Fig. 8.8 for an example of this output. As the message indicates, this is a base 10 logarithm. In this particular case the evidence indicates that it is a bet of approximately  $10^{49.6} : 1$  that a resonance is present.

After the signal detection calculation is performed, and assuming a resonance was detected, Bayes Analyze builds a model having one new resonance and then optimize the parameters associated with this model. The resonance added corresponds to the frequencies having maximum signal detection probability. The optimization step locates the maximum of the joint posterior probability for the parameters. The second section of the output report contains the values of the parameters for which the joint posterior probability is maximum. An example of this output is shown in Fig. 8.9. Figure 8.9 consists of the global parameters, here the two phase parameters, and this is followed by the model parameters. In the parameter file this model information was on one very long line. In the output report each model component is on several lines and each component has one or more amplitudes associated with it. The individual model components all start with a header. This header begins “#Name” and is always present. There are five headers shown in Fig. 8.9, so there are five model components. Each model component consists of a model name, the estimated resonance parameters, and the estimated amplitudes. Both constants and resonances have amplitudes; however, only resonances have frequencies, decay rate constants and  $J$ -coupling constants. The model name, multiplet order, relative phase, and other parameters are the same as those described in Section 8.5.1 and we will not repeat that here. The resonance frequency is printed on the same line with the model name and is followed by up to three additional lines containing the decay rate, the primary  $J$ -coupling constant, and the secondary  $J$ -coupling constant. Each of these parameters has an associated uncertainty printed with it.

Following the model component name and the resonance parameter, if present, are the amplitudes associated with each free induction decay processed. If there were 10 free induction decays, there will be 10 amplitudes. The number of the free induction decay that the amplitude is associated with appears on each line. In addition to printing the amplitudes, the uncertainty in the estimated

Figure 8.9: A Small Sample Of The Output Report

```

Base 10 Log Evidence for The Next Frequency is: 49.6
-----
Number of Models In This Step 6
Center Phase      Uncertainty      Time Delay      Uncertainty
-5.7941E+01      7.32E-01          8.322E-03      6.28E-03

# Name
1 Real Constant
      Fid  Amplitude  Uncertainty
      1  5.106E+01  4.41E+01

# Name
2 Imaginary Constant
      Fid  Amplitude  Uncertainty
      1  1.16E+01  4.41E+01

# Name      Order      Type      Parameter      Uncertainty
3 (CP) Singlet      (1,1) 0      Freq  1.16733E-01  3.54E-04
      Decay  1.007E+00  3.25E-01
      Fid  Amplitude  Uncertainty
      1  7.997E+02  6.58E+01

# Name      Order      Type      Parameter      Uncertainty
4 (CP) Triplet      (3,1) 0      Freq  1.2913983E+00  6.07E-05
      Decay  8.577E-01  3.73E-02
      Jp    6.9822E+00  2.31E-02
      Fid  Amplitude  Uncertainty
      1  9.381E+03  9.97E+01

# Name      Order      Type      Parameter      Uncertainty
5 (CP) Quartet      (4,1) 0      Freq  4.8648472E+00  8.50E-05
      Decay  8.879E-01  5.96E-02
      Jp    6.9795E+00  3.26E-02
      Fid  Amplitude  Uncertainty
      1  6.345E+03  1.10E+02

Base 10 Log Of The Probability of 4 Resonances =-3.17799725E+03

              Probability For The Model
----- Model ---- Probability ---- Prob/I --- Prob/F ----- Ended -----
  1 Resonances  3.149126-312      431.0      431.1 Fri May 10 10:41:53 1996
  2 Resonances  3.274689E-50      626.3      707.6 Fri May 10 10:41:54 1996
  3 Resonances  1.000000E+00      764.3      768.3 Fri May 10 10:41:55 1996

Bayesian Analysis Ended Fri May 10 10:42:11 1996
And Took: 15 Sec. (0.25 Min.)

```

Figure 8.9: The top of the reports indicates the number of models and outputs the global parameters used in the analysis. The numbered entries are the individual models and the parameters estimates for those models. Finally the bottom shows the base 10 logarithm of the posterior probability for the models and how much the probability increased during the search.

Figure 8.10: Bayes Analyze Uncorrelated Output

#	Name	Order	Type	Parameter	Uncertainty		
3	(UP) Triplet	(3,1) 0	Freq	1.291424E+00	1.28E-04		
			Decay	8.688E-01	7.77E-02		
			Jp	6.9828E+00	4.84E-02		
			Fid	Cos Amplitude	Uncertainty	Sine Amplitude	Uncertainty
			1	4.826E+03	2.10E+02	-8.118E+03	2.10E+02

Figure 8.10: The output is different for uncorrelated resonance. In uncorrelated resonance models, there is a phase parameter for each resonance which, for computational reasons, is written as two amplitudes. If the sine amplitude is represented by  $S$  and the cosine amplitude by  $C$ , then the amplitude of the sinusoid  $A$  is related to  $S$  and  $C$  by  $A = \sqrt{C^2 + S^2}$ .

amplitude is also printed. These uncertainties are different for each Fid. The amplitudes shown in Fig.8.9 are all for correlated phase models components so there is only a single amplitude. However, uncorrelated phase models have two amplitudes associated with them: a cosine amplitude and a sine amplitude. Under these conditions the output amplitudes have two entries, see Fig. 8.10. If you compare the frequency, decay rate constant and  $J$ -coupling constants you will find them to be essentially identical. The amplitude has been replaced by the corresponding cosine and sine amplitudes along with an estimate of how uncertain one is of the true value. In the summary reports these two amplitudes are reported as an amplitude and a phase. For more on the summary reports see Sections 8.5.8 and 8.5.9.

After the output associated with the parameter estimates comes the output from the model selection calculation. This output is in two forms, one is a logarithm of the posterior probability for the model. In Fig. 8.9 this is given by the following line:

Base 10 Log Of The Probability of 4 Resonances =-3.17799725E+03

This is a logarithm of an unnormalized probability distribution; consequently, only differences between this number and the logarithm of the probability for some other model are meaningful. Roughly speaking a difference of one means that the new model is fitting one additional data value. Because the number of data can be large it is not unusual to see large difference in the probability for the models.

The second form of this output is the normalized probability for the model. This cumulative probability distribution consists of one entry for every model tested. Here are these lines reproduced from Fig. 8.9:

Probability For The Model						
Model	Probability	Prob/I	Prob/F	Ended		
1 Resonances	3.149126-312	431.0	431.1	Fri May 10 10:41:53	1996	
2 Resonances	3.274689E-50	626.3	707.6	Fri May 10 10:41:54	1996	
3 Resonances	1.000000E+00	764.3	768.3	Fri May 10 10:41:55	1996	

They indicate several things about the models that were processed. First, three models were tested on this run of Bayes Analyze. If Bayes Analyze is run multiple times, all of the models tested are written into the probabilities file and a blank line separates the output from successive runs. Each model written into the output file gets a short description. This description is under the



labeled, “Model,” in fig. 8.9. The description is usually just the number of resonance components in the model. The model description is followed by the normalized posterior probability for the model. These are labeled “Probability.” When this probability distribution is normalized, one is essentially using probability theory to answer the following question: Given the one, two, and three exponentially decaying resonances models which of these three models best accounts for the data. In order to normalized the probability distribution, we had to tell probability theory to pick the best model from this set and only from this set. This is not equivalent to saying that the three exponential model is the best, only that it is the best of the three tested. Indeed the statement that some model is the “best” is indeterminate until the alternatives against which it is being tested are given.

In this example, the one resonance model has a probability that is essentially zero, the two resonance model has a probability 272 orders of magnitude higher; but, it is still 50 orders of magnitude lower than the probability for the three resonance model. These numbers are computed using a Student-*t* distribution. The logarithm of the Student-*t* distributions is multiplied by essentially the number of data values. A change of one order of magnitude in these probabilities means, at least very approximately, that the model has fit one more data value – so a change of 272 means that the two resonance model fits roughly 272 more data points than the one resonance model, etc. This rule – one point equal one order of magnitude – is only approximate and should be used only as a guide in ones thinking. Usually the changes from one model to the next are so large that little thought need go into which to accept. However, sometimes these probabilities can be very close, and when that happens one must carefully evaluate the different models to determine which is appropriate.

After the probabilities comes the initial (Prob/I) and the final (Prob/F) base 10 logarithm of the joint posterior probability for the parameters in the model. The joint posterior probability for the parameters is the probability density function used in optimizing the model parameters. The initial and final probabilities listed here are the initial and final probabilities from that optimization.

Finally, the last entry is the date and time the processing occurred. When Bayes Analyze begins a step it gets the date and time from the system clock. The numbers displayed are those times. Their primary function is to supply the user with one more indication of which model one is dealing with. When the user tests many different models it is possible for this set of probabilities to become fairly long. These dates and times along with the initial and final search probabilities can help determine which models one is dealing with.

The list of probabilities printed in this section are taken from the probabilities files. Bayes Analyze will never delete the probabilities files. If they are present, it will simply append the results of any new calculations to the end of these files. We maintain this information across multiple runs precisely so that one may test different models. To reset this list of probabilities one must clear the experiment, load new data, or manually delete the probabilities files.

After Bayes Analyze completes a run, it will compute the total computer time used and print this number at the end of the output file. This number is purely informational.

#### 8.5.4 The “bayes.probabilities.nnnn” File

The probabilities file is perhaps one of the simplest and most important files in the Bayesian Analysis package. These files are named “bayes.probabilities.nnnn” where the number of the first Fid analyzed in this file. An example of the probabilities file is shown in Fig. 8.11. This file is particularly simple in its format. It consists of a description of the model, i.e., “A Constant,” “1 Resonance,” etc., the

Figure 8.11: The bayes.probabilities.nnnn File

Model Desc	Model Prob	Search/I	Search/F	Date-Time
1 Resonance	0.000	102.0	124.9	Mon May 20 08:54:44 1996
2 Resonances	40.823	169.4	174.9	Mon May 20 08:54:45 1996
3 Resonances	93.923	228.5	237.7	Mon May 20 08:54:45 1996
4 Resonances	165.324	298.0	318.8	Mon May 20 08:54:46 1996
5 Resonances	261.385	396.5	426.5	Mon May 20 08:54:47 1996
6 Resonances	443.048	536.2	622.7	Mon May 20 08:54:48 1996
7 Resonances	478.148	666.7	668.5	Mon May 20 08:54:49 1996
8 Resonances	515.355	713.7	716.9	Mon May 20 08:54:51 1996
9 Resonances	576.265	770.8	790.1	Mon May 20 08:54:52 1996
10 Resonances	576.079	797.5	798.3	Mon May 20 08:54:55 1996
1 Resonance	90.088	220.1	220.2	Mon May 20 08:55:09 1996
2 Resonances	156.099	286.7	296.1	Mon May 20 08:55:09 1996
3 Resonances	260.609	384.5	410.9	Mon May 20 08:55:10 1996
4 Resonances	450.966	531.9	614.9	Mon May 20 08:55:11 1996
5 Resonances	486.696	659.5	661.2	Mon May 20 08:55:12 1996
6 Resonances	524.166	706.5	709.6	Mon May 20 08:55:13 1996
7 Resonances	584.155	762.5	781.5	Mon May 20 08:55:14 1996
8 Resonances	583.867	788.8	789.5	Mon May 20 08:55:16 1996
2 Resonances	284.782	431.2	431.4	Mon May 20 08:55:30 1996
3 Resonances	546.734	624.3	707.2	Mon May 20 08:55:31 1996
4 Resonances	596.811	763.7	768.4	Mon May 20 08:55:32 1996
5 Resonances	596.499	775.7	776.4	Mon May 20 08:55:33 1996
2 Resonances	281.672	428.1	430.4	Mon May 20 08:56:02 1996
3 Resonances	542.561	624.6	705.9	Mon May 20 08:56:05 1996
4 Resonances	592.327	762.4	767.0	Mon May 20 08:56:08 1996
5 Resonances	592.005	774.4	775.1	Mon May 20 08:56:11 1996

Figure 8.11: Unlike the other files in the Bayesian Analysis package the bayes.probabilities.nnnn file accumulates across multiple runs of Bayes Analyze. Indeed the only way to get rid of this file is either clear the experiment, load a new Fid, or issue a Unix “rm” on this file. The probabilities file contains the probability for the various models that were tested by Bayes Analyze. The first line in this figure, the header, is here only to label the columns; it is not normally present in this file. The file consists of an indication of the type of model tested, the base 10 logarithm of the probability for that model, the initial and final search probabilities and the date and time the model was tested. For more on this example see the text.

base 10 logarithm of the probability for the model, the beginning and ending search probability (if applicable), and the date and time the model was analyzed.

To compare several different models using probability theory, one computes the probability for each model. The model with highest probability is the best. When Bayes Analyze computes the probability for a model, that probability is appended to the probabilities file. Because information is appended to this file it accumulates across multiple runs of Bayes Analyze and so it may be used to compare many different models to see which model is best. To illustrate how one can use this file to do model selection see Fig. 8.11. This figure is an example of the probabilities file for four different runs of Bayes Analyze. Note that Bayes Analyze has separated these four different runs with a blank. In this figure, we have changed the normalization so that the logarithm of the probability for the one resonance model is zero. The probabilities for the models are stored as unnormalized logarithm, so any constant may be added or subtracted from every value in the file without changing the normalized probabilities.

The probabilities file illustrated in Fig 8.11 are from an analysis of a H1 Fid of ethyl ether. Part of the spectrum of the Fid is shown in Fig 8.1. Bayes Analyze was run on this data four different times using different initial models to see which model was the best description of this data. In the first of the four runs Bayes Analyze was allowed to find correlated resonances in its automatic mode. The maximum resonances was set to 20 resonances, many more than the 9 present. In that run the program found 9 resonances, three associated with the triplet, four with the quartet, and the two nuisance resonances. When Bayes Analyze tried to add a tenth resonance the probability for the model went down thus indicating that this last candidate resonance was probably noise. This run is represented in Fig. 8.11 by the first 10 lines. Notice that the probability for the model starts at zero and then rises steadily until it reaches a maximum of 576.265 and then it decreases slightly.

The first test indicated that a 9 resonance model was the best. However, we know that there is a triplet and a quartet in this data. In the next test, we marked the triplet, and let Bayes Analyze add resonances until the probability for the model decreased. Bayes Analyze added resonances until it reaches an 8 resonance model and then the logarithm of the model probability decreases; so Bayes Analyze has indicated that a 7 resonance model (1 triplet and 6 singlets) is the best model tested so far. The logarithm of the probability for 9 singlets was 576.265; while the logarithm of the probability for the 6 singlet plus one triplet is 584.155, an increase of 7.89 orders of magnitude. So probability theory prefers this model to the 9 singlet model, as it should.

In the third test, we marked the triplet and quartet and then allowed Bayes Analyze to add resonances in its automatic mode. Bayes Analyze tested three additional resonances and on the third resonance the logarithm of the probability for the model decreased. So the model preferred in this test was a triplet, a quartet and two singlets for a total of 9 peaks. The base 10 logarithm of the probability for this model is 596.811. This should be compared to the best model in the previous two test: 584.15 and 576.265. Note that this model is 12.656 orders of magnitude more probable than the best previous model.

Just for illustrative purposes we ran one additional test. We turned on third order shimming and allowed Bayes Analyze to run using the triplet and quartet model as a starting point. In this case Bayes Analyze added three resonances and on the third singlet the probability decreased. But the probability for this shimming model has decreased from 596.811 to 592.327, a decrease of 4.484 orders of magnitude – so the shimming model does not help. The shimming model did nothing to improve the fit and the prior probability for this model has decreased enough to rule this model out.

The bayes.probabilities.mmm files may be safely deleted anytime, even while Bayes Analyze is running. If a probabilities file is not found, Bayes Analyze creates it when it begins analyzing a

new model. There are three ways to reset the probabilities file. Here “reset” means removing the content of the file. Resetting the file may be desirable when you wish to compare a given set of models. First, loading a new free induction decay will clear all of the files out of the current WorkDir. Second, hitting the “Reset” button on the Bayes Analyze interface will remove all of the current Bayes Analyze output files from the current WorkDir, including the probabilities file. Finally, you can manually delete the file by navigating to Bayes/WorkDir/BayesAnalyzesFiles and deleting the bayes.probabilities.mnmm files.

### 8.5.5 The “bayes.log.mnmm” File

A record of what the search algorithm is doing is written to the bayes.log.mnmm file in the format shown in Fig. 8.12. Each search, with the possible exception of the first one, begins with a line

Figure 8.12: The bayes.log.mnmm File

```

Base 10 Log Evidence for The Next Resonance is: 0.3
----- 5 Resonance Model -----
Log Prob      Relax      L      Parameter      Parameter      Parameter
776.5         I 1.000  0.010   0.1167(f)     0.9950(d)     -57.9401(Ph)
              I 1.000  0.010   1.2914(f)     0.8575(d)     6.9822(Jp)
              I 1.000  0.010   0.0083(T0)    4.8648(f)     0.8882(d)
              I 1.000  0.010   6.9795(Jp)    5.6478(f)     3.2275(d)
              I 1.000  0.010   4.4656(f)     1.0876(d)
777.1         A 1.000  0.000   0.1167(f)     0.9822(d)     -57.8867(Ph)
              A 1.000  0.000   1.2914(f)     0.8571(d)     6.9821(Jp)
              A 1.000  0.000   0.0088(T0)    4.8648(f)     0.8869(d)
              A 1.000  0.000   6.9794(Jp)    5.6477(f)     3.2289(d)
              A 1.000  0.000   4.4648(f)     2.5345(d)
777.3         Z 1.000  0.000   0.1167(f)     0.9806(d)     -57.8711(Ph)
              Z 1.000  0.000   1.2914(f)     0.8569(d)     6.9821(Jp)
              Z 1.000  0.000   0.0090(T0)    4.8648(f)     0.8868(d)
              Z 1.000  0.000   6.9795(Jp)    5.6477(f)     3.2284(d)
              Z 1.000  0.000   4.4645(f)     3.5147(d)

Base 10 Log Of The Probability of 5 Resonances =-3.17473576E+03

                Probability For The Model
----- Model ---- Probability ---- Prob/I --- Prob/F ----- Ended -----
  A Constant    4.358829-431      12.7           Wed May 1 13:14:23 1996
  2 Resonances  5.610831-313       123.8          431.4 Wed May 1 13:14:25 1996
  3 Resonances  5.051696E-51        624.3          707.2 Wed May 1 13:14:25 1996
  4 Resonances  6.026667E-01        763.7          768.4 Wed May 1 13:14:26 1996
  5 Resonances  2.940744E-01        775.7          776.4 Wed May 1 13:14:28 1996

Bayesian Analysis Ended Fri May 3 14:23:44 1996
And Took: 1 Sec. (0.02 Min.)

```

Figure 8.5.5: The “bayes.log.mnmm” file is a history of what Bayes Analyze did during an analysis. This file generally starts with the base 10 logarithm of the odds ratio for another resonance in the data. Bayes Analyzes uses the frequency associated with that logarithm of the evidence to postulate a new model. It optimizes that model. Each step shown here is a single step in the Levenberg-Marquardt algorithm. The Bayes Analyze program and accepts or rejects the model depending on this posterior probability.

containing the logarithm of the evidence in favor of another resonance and ends with a line containing the logarithm of the posterior probability for the current model. Both these log probabilities are illustrated in Fig. 8.12, the evidence in favor of a resonance the first line in Fig. 8.12 and the logarithm of the posterior probability for the model is the line beginning “Bayes 10 Log Of The Probability of”. We indicated that the first search may not contain the the logarithm of the evidence in favor of another resonance, that’s because the first search may be done on a user defined model and in this case the evidence calculation is not performed.

If the evidence in favor of another resonance is greater than zero, i.e., odds greater than 1:1, Bayes Analyze will create a model containing the frequency with the highest evidence and perform a search for the peak value of the joint posterior for the parameters in this model. As Bayes Analyze is searching, it write out a set of lines showing the current status of the search. In Fig. 8.12 these search status lines are essentially the center of the figure. The Levenberg-Marquardt search algorithm has several parameters associated with it, these are: the logarithm of the joint probability for the parameters, a relaxation parameter and a Levenberg-Marquardt parameter. These parameters are identified by the headings: “Log Prob”, “Relax”, and ‘L’ shown in Fig. 8.12. The single character in front of the relaxation parameter indicates whether the search step is the **I**nitial step, a step in which the posterior probability increased and the step was **A**ccepted, or whether or not the posterior probability decreased and the step was therefore **R**ejected. Finally, **Z**, indicate the covariance matrix is being computed.

The remaining three fields in Fig. 8.12, under the “Parameter” label, are the parameters that appear in the model. There is no particular order to these parameters except to note that resonances are listed in increasing resonance order. However, the resonances are ordered before a search begins – not while it is going on, so it is possible for the resonance frequencies to become disordered while the search is in progress. Each parameter is listed to four decimal places, although, the number of decimal places varies depending on the size of the parameter. The parameter is followed by a short description in parentheses – See Table 8.3 for a list of these descriptions and their meaning. This short description is only used in the log file.

After Bayes Analyze locates the maximum of the joint posterior probability for the parameters, the approximate covariances matrix is computed with the Levenberg-Marquardt parameter set to zero. This covariance matrix is used to compute a Gaussian approximation of the joint posterior probability for the parameters. Finally, by integrating this joint posterior probability for the parameters, one is able to compute an approximation to the posterior probability for the model. The base 10 logarithm of the posterior probability for the model is printed at the end of each search. Here is an example of this line:

```
Base 10 Log Of The Probability of 5 Resonances ==-3.17473576E+03
```

Because it is the logarithm of the posterior probability for the model that is printed, it is only differences between between the logarithm of the model probabilities that are meaningful. For example if the previous model probability was -128 and the current model has probability of -100, then this new model is  $10^{28}$  times more probably than the previous model.

After printing out the base 10 logarithm of the probability for the model, the output files are updated and the calculations, starting with the signal detection calculation, are repeated. The log file contains a record of what occurs as Bayes Analyze optimizes and test each model. These calculations continue until one of three things happen: the logarithm of the posterior probability for the model decreases, the signal detection routine failed to find a candidate resonance, or the maximum number of new resonances has been reached. After completing the calculation on a given

Table 8.3: Bayes Analyze Short Descriptions

Short Desc.	Full Description	Units
(Ph)	The center phase	Degrees
(T0)	The starting time of the fid	Points
(f)	The Resonance frequency	Hertz or PPM
(d)	The decay rate constant	Hertz
(Jp)	The primary $J$ -coupling constant	Hertz
(Js)	The secondary $J$ coupling constant	Hertz
(Sd)	Lorentzian spacing in lineshape expansion	Hertz
(1)	The first relative shim amplitude	No units
(2)	The second relative shim amplitude	No units
⋮		
(7)	The seventh shim amplitude	No units

Table 8.3: Bayes Analyze prints the parameters along with a short description of them to the bayes.log.mnmn file. These descriptions are shown here along with an explanation and the units used.

model, Bayes Analyze prints the the normalized posterior probability for the models to the log file. This part of the printout is shown at the bottom of Fig. 8.12 and begins with header reading “Probability For The Model” This list of model probabilities is cumulative and accumulates across multiple runs of the Bayes Analyze program. Different runs are separated by a blank line.

In Fig. 8.12 the lines that make up the normalized probability for the model,

```

                Probability For The Model
----- Model ----- Probability ----- Prob/I --- Prob/F ----- Ended -----
  A Constant    4.358829-431      12.7           Wed May  1 13:14:23 1996
  2 Resonances  5.610831-313      123.8          431.4 Wed May  1 13:14:25 1996
  3 Resonances  5.051696E-51      624.3          707.2 Wed May  1 13:14:25 1996
  4 Resonances  6.026667E-01      763.7          768.4 Wed May  1 13:14:26 1996
  5 Resonances  2.940744E-01      775.7          776.4 Wed May  1 13:14:28 1996

```

consists of a description of the model, the normalized probability for the model, the logarithm of the joint posterior probability for the parameters at the beginning of the search, under the heading “Prob/I,” followed by the logarithm of the joint posterior probability for the parameters at the completion of the search under the heading “Prob/F” and finally the time this particular model was run. If constant models are present they are shown at the top of the entries for a given run of Bayes Analyze.

After this table of model probabilities is printed, Bayes Analyze determines the amount of computer time it took to run this job and prints this information along with the data and time the job started and ended. After printing this last entry, Bayes Analyze terminates.

Figure 8.13: The bayes.status.nnnn File

```

Bayesian Analysis started: Tue May 14 13:11:01 1996
Status last updated: Fri May 17 10:27:21 1996
Current Fids: 15 Through 25
Status: Analysis completed

```

The “bayes.status.nnnn” file indicates the current status of a run. This file is also written using the name Bayes.accepted and is displayed when the “Get Job” button is pressed if the job has not completed.

### 8.5.6 The “bayes.status.nnnn” and “bayes.accepted.nnnn” Files

The status file is a three or four line file that is written by Bayes Analyze and Bayes Model. The accepted file is just the current version of the status file. Both Bayes Model and Bayes Analyze will write and or update the status and the accepted files. Bayes Model places the number of the last Fid modeled in the status and accepted files, while Bayes Analyze continuously continuously updates both the status and accepted files with the current status of the Bayes Analyze run. An example of this file is shown in Fig. 8.13. The first line of the file is the time the analysis started. The second is the last time this file was updated. Optionally the free induction decays being processed is given on the third line. This line is present when multiple Fids are being processed in blocks. The last line indicates what action happen at the time of the last update. In the example shown, this line indicates that the run had completed. The status file is also used to communicate setup problems, and program errors. A complete listing of all of the messages that may be written to the status file is given in subsection 8.6. In this section, informational messages are numbered 1-8; setup messages are numbered 9-29; and errors messages are numbered 30-49. The informational messages are written into the status file for no other purpose than to allow you to determine the current status of a run. The status display contains the status file as part of that display. The informational messages might be as simple as indicating that the input data is being read (message number 1) or they might indicate that the analysis is completed as message 7 indicates. Message numbers 9-29 are setup errors. These indicate that the Bayes Analyze detected a problem in the setup. They include such things as: not finding the input free induction decay (message 15), exceeding to the capabilities of the programs (message 27), and a number of others. All of these messages require one to correct one or more problems in the parameter file. The remaining messages, messages 30-49, are program errors in one form or another. These message consists of the name of the routine in which the error occurred and a message that indicates you should contact *me*. In almost all of these cases these message could not occur without something being seriously wrong with the program. Indeed most of these messages can occur only while the program is being tested. However, there are two messages, message number 37 and 38, that are not terribly serious, and when they occur are probably related to how the analysis was setup.

Message 37 occurs when a matrix inversion routine detects a singular matrix. But singular matrices can only occur when one or more of the model functions are nearly identical and even then this is rare because the prior probabilities used in the calculations are suppose to prevent this. You should check to make sure that the analysis is makes physical sense. For example, did you mark a peak and then Bayes Analyze tried to put a second peak at nearly the same location. This type of problem can occur whenever you mark a peak too far from the “true” location of the resonance.

Figure 8.14: The bayes.model.nnnn File

```

! # Name Order Frequency Init Value
  3 (CP) Singlet (1,1) 0 24.305874354538 24.172984999931

!
      Fid Amplitude
      1 76499.464296592
      2 75736.815735717
      3 75008.870579233
      4 74424.926696927
      5 73761.523349848
      6 73730.846586329
      7 72942.021502301
      8 72735.215054879
      9 72103.722022270
     10 72022.172495999

```

Figure 8.14: The bayes.model.nnnn file is a modified version of the bayes.params file. In the bayes.model.nnnn file each model is followed by the amplitudes, one amplitude for each data set. Otherwise, the format of the Bayes Model and Bayes Params file are the same.

It can also occur when you mark a peak thinking it is a single line and it is really multiple lines. The solution to this type of problem is to remove the marked peaks and allow Bayes Analyze to put the peak in by itself. There are other reasons that this problem can occur, such as when all of the constant models are used and Bayes Analyze attempts to model a baseline artifact. Baseline artifacts are modeled with vary rapid decaying sinusoids, and so are very nearly the same as the first point models – so they can be come redundant. This can cause the matrix inversion to fail. The pattern should be clear: this problem occurs when several model components are nearly identical. The solution is to remove the redundant models and rerun the analysis. If the problem persists and you can make no progress on solving the problem, contact [me](#). The other error, message number 38, is essentially the same as the previous one. In this case the matrix inversion routine failed but did not detect the failure. One or more of the model functions are redundant and that redundancy is causing either convergence or roundoff problems in the matrix inversion routine. The solution is again to remove the redundant vectors and try the analysis again. The remaining errors in Subsection 8.6 are true program errors and should never occur. Indeed if they do occur, you should save the WorkDir in which the problem occurred (making sure you get the Fid and not a pointer to it) and then contact [me](#).

### 8.5.7 The “bayes.model.nnnn” File

The model file is a modified version of the parameter file. In addition to all of the information contained in the parameter file, the model file also contains the estimated amplitudes of each model component. An example of the amplitude information is shown in Fig. 8.14. The resonance shown is the same triplet shown in Fig. 8.6. The amplitude section has a comment header, followed by the estimated amplitude of the resonance. This line is repeated for each free induction decay processed. Note that this is the output for a correlated phase model. Some models, uncorrelated phase resonances, have multiple amplitudes per resonance. Fig. 8.15 is an example of the output amplitude for an uncorrelated resonance. The uncorrelated resonance has both a cosine and a sine



Figure 8.15: The bayes.model.nnnn File Uncorrelated Resonances

```
!
                                Fid      Cos Amplitude  Sine Amplitude
                                1      4848.0043462323 -8028.116298822
```

Figure 8.15: Uncorrelated resonance models have two amplitudes, effective an amplitude and a phase. Otherwise the format of the an uncorrelated resonances is exactly the same as for a correlated resonance model.

Figure 8.16: Bayes Analyze Summary Header

```
Summary1: /home/akgrp/glb/vnmrsys/exp3/bayes.output.0001
Date/time: Fri May 10 10:42:13 1996
Directory: /home/akgrp/glb/vnmrsys/exp3
File:      "/home/akgrp/glb/data/p13.fid"
Text File: This is test data
```

Figure 8.16: The summary report header contains a few items particular the the analysis that you are running. In particular it contains the name of the output file being summarized, the data and time, the analysis directory and the text file from the Fid. For more on this report, see the text.

amplitude. Together these specify the amplitude and phase of the resonance.

### 8.5.8 The “bayes.summary1.nnnn” File

The “bayes.summary1.nnnn” file is produced when the **Best** button is activated. This button calls a program name Bayes Summary1. In the process of analyzing a Fid, Bayes Analyze develops a hierarchy of models. These models usually begin with one resonance, followed by two, etc., until either the signal detection algorithm fails to find evidence for an additional signal or the probability for the model goes down. The output file written by Bayes Analyze contains all of this information and can get fairly long. The purpose of the summary report is to take the information in the output files and to print out only that part of it that is most important to the analysis. In particular it extracts that part of the output file that corresponds to the model having maximum probability. In this section we describe the output from Bayes Summary1 and point out where the information has been reformatted.

First, the preceding discussion indicates that the summary report is built out of the output file; this is not totally correct. The summary report is built out of pieces of the output, model, probabilities, text, and procar files. When Bayes Summary1 is run the first thing the program does is to determine how many output files there are in the experiment. Each of these output files are summarized one at time – so if there are 10 output files, 10 different summary reports are written.

Much of the summary report is identical to other outputs and we will not repeat those descriptions here. We will simply reference you to the correct descriptions. The summary output starts with a header that is unique to the summary report. The header is shown in Fig. 8.16. This header is made up of the name of the output file that was summarized, the date the summary was done, the directory in which the output files were located, the name of the Fid that was processed, and last the contents of the text file. In building this header Bayes Summary1 had to process both the text file and the procar file. However, this program also gathers information from the model file, the

Figure 8.17: The Summary2 (Best Summary)

```

Summary2: ./bayes.output.0001
Date/time: Fri Sep 6 08:57:32 1996
Directory: .
File: "/home/jagrp/glb/vnmrsys/data/ethyl.ether"
Text File: This is test data

      Freq      Error      Decay      Error  Amplitude  Error  Type Resonance
      ppm      ppm      Hertz      Hertz  arbitrary
0.116750  1.10E-04  0.604000  0.110000  648.0000  27.00000 (CP) Singlet
1.291439  7.90E-05  0.792000  0.023000  9033.000  47.00000 (CP) Triplet (See 1)
4.864975  2.50E-05  0.772000  0.026000  5926.000  51.00000 (CP) Quartet (See 2)
5.648083  6.30E-05  2.991000  0.057000  3999.000  50.00000 (CP) Singlet

No.      Jp      Err      Sp      Err
      Hertz  Hertz  Hertz  Hertz
  1  6.978000  0.008100
  2  6.972500  0.009600

Bayesian Analysis Ended Thu Dec 1 15:35:19 1994
And Took: 0 Sec. (0.00 Min.)

```

Figure 8.17: The output from the Summary2 program is condensed listing of the frequencies, decay rate constants, amplitudes, and the uncertainties in these parameters.

output file, and the probabilities file. These files are used in building the main body of the report. This header is followed by the configuration parameters and these have already been explained – see Section 8.5.1 for a description of these parameters.

The remaining part of the summary report is essentially identical to the output report and we will not repeat it here. There is one difference that is worth noting, in the summary report uncorrelated resonances have their amplitudes reported as an amplitude and a phase, this is different from the output report where there are listed as a cosine and sine amplitude. This difference is mostly for convenience, sometimes the amplitude of an uncorrelated resonances is needed.

### 8.5.9 The “bayes.summary2.nnnn” File

The “bayes.summary2.nnnn” file is produced when the **Freq** button is activated the Bayes Analyze interface. This button calls a program name Bayes Summary2. Like the first summary report, this report is only on the section of the output file most important to the user. Because of its extremely compressed format this report can only be run on non-arrayed free induction decays. The report is essentially a column listing of the estimated frequencies, decay rate constants, the amplitudes and the uncertainties in these parameters. The report can be run on either correlated or uncorrelated phase models but when it is run on uncorrelated phase models the sine and cosine amplitudes are reported as the square root of the sum of the squares of these amplitudes. This quantity is the amplitude of the resonance.

An example of this report is shown in Fig. 8.17. Note that the first few lines of this report are essentially the same as in the Summary1 report. The main difference is in the detailed output. In this report the detailed output is essentially a column listing of the most probable model in the output

report. In this column listing the frequency, decay rate constant, amplitude, and the estimated uncertainty in our knowledge of these parameters are given. Additionally, the type of resonance and the resonance name are given. When the resonance is a multiplet the coupling constants are given at the end of the report. When no multiplets are present this section of the report does not appear. When uncorrelated resonances models are encountered the sine and cosine amplitudes are reported in the form of a total amplitude for the resonances. The phase of these resonances are not reported. If this information is needed it is reported in the Summary1 report. Last, the resonances are reported in the units that were in use at the time the analysis was run.

### 8.5.10 The “bayes.summary3.nnnn” File

The summary3 file and report are generated whenever a regions file is present in the current Working Directory. The regions/summary3 report takes as its input a “bayes.regions” input file. This file is produced automatically whenever the regions button is activated. It contains a series of lines, each line containing a region of the spectrum to be summarized. These regions are written into the file as low-high frequency pairs. As a user you set them by displaying the spectrum, and activating the “Options/set regions” button.

The regions/summary3 report contains the standard header, the regions information, followed by the probabilities information. Figure 8.18 shows a copy of the regions report. First, the actual regions are shown. For this report, these regions were specified by “Options/Set Regions” and when they are defined the regions report is automatically run when Bayes Analyze is run. Each region contains the total intensity in that region. This information is displayed in two forms, as a percent of intensity and as the total intensity. Additionally, each of these regions contain an estimate of the uncertainty in the intensity estimates.

## 8.6 Bayes Analyze Error Messages

This appendix contains a complete listing of all of the messages that Bayes Analyze and Bayes Model write to the status file. In the following list we have numbered the messages, the message is in bold, and this is followed by an explanation of the message. These messages are divided into three general categories: informational, setup errors, and program errors. The categories are not organized in any particular order. There are some additional comments about each category at the end of each category.

When Bayes Analyze is running in either its automatic or slave mode the status file is updated with an information message whenever the program changes from one general function to another. The informational messages follow:

1. **Reading input data** – The input free induction decays are being read by Bayes Analyze.
2. **Completed Baseline Artifact Model** – The baseline model has been processed. Bayes Analyze will attempt to model a baseline artifact whenever the baseline routines are activated in the parameter file. These routines can only be activated manually at present.
3. **Probability For A Constant Completed** – Computing the probability for the constant model is now completed.

Figure 8.18: The Summary3 Report

```

----- Regions -----
#      From      To
1      2.47      2.73
2      2.99      3.18
3      3.18      3.33
4      3.33      3.46
5      3.46      3.62
6      3.66      3.86
7      3.93      4.28
8      5.21      5.64

Fid:  1 of  1 Total Intensity In The Region
Region 1      Err      Region 2      Err      Region 3      Err
3.3930E+03  2.55E+02  3.9720E+03  3.00E+02  3.8620E+03  3.37E+02

Region 4      Err      Region 5      Err      Region 6      Err
4.8270E+03  2.98E+02  5.5800E+03  3.67E+02  4.2380E+03  3.52E+02

Region 7      Err      Region 8      Err
6.8470E+03  3.38E+02  5.9030E+03  3.39E+02

Fid:  1 of  1 Percent of the Intensity In The Region
Region 1      Region 2      Region 3      Region 4      Region 5      Region 6
8.79 0.66  10.28 0.78  10.00 0.87  12.50 0.77  14.45 0.95  10.97 0.91

Region 7      Region 8
17.73 0.88  15.28 0.88

```

Figure 8.18: The Regions/Summary3 Report is output by the Bayes Analyze package whenever regions are defined. Briefly, this report is the total resonance intensity in each of the defined rations.

4. **Performing  $n$  Resonance Search** – Bayes Analyze is optimizing the parameters for an  $n$  resonance model where  $n$  is the number of resonances components in the model.
5. **Completing  $n$  Resonance Model** – The maximum of the joint posterior probability for the  $n$  resonance model has been located and Bayes Analyze is now writing the output files and computing the probability for the model.
6. **Completed  $n$  Resonance** – Bayes Analyze has finished the  $n$  resonance model but has not yet started the  $n + 1$  model.
7. **Analysis completed** – Bayes Analyze has completed running in both the slave and automatic mode. You may now print, model, or modify the analysis in anyway deemed appropriate.
8. **Fid Number  $n$  Modeled** – The analysis has been completed.

The proceeding messages are informational and do not indicate any problem. They are written to the “`bayes.status.nnnn`” file to indicate the current status of the analysis. However, the following messages indicate some type of a problem. If you receive one of these messages you must take action to correct the problem.

9. **Configuration error - Did not recognize (text)** – There is an error in the configuration parameters in the parameter file. The word “(text)” is replaced by the text that was not recognized.
10. **Memory Allocation Failed – Rerun Later** – Both Bayes Analyze and Bayes Model dynamically allocates virtual memory for the data and for the models. If there is not enough memory available you will receive this message. If you have sufficient virtual memory, you can try again later. If you do not have enough virtual memory to run the program you can increase the swap space and then rerun the analysis. Alternately, you can try analyzing fewer free induction decays. Along the same line, you can try reducing the amount of data analyzed by setting a signal region. Reducing either of these reduces the amount of virtual memory requested.
11. **Fid File Not Found – Correct Model File** – The input Fid file was not found. Bayes Model is looking for the Fid specified in the model file. Either that Fid has been deleted, the model file is wrong, or Bayes Model is being run on a machine that does not have access to the Fid.
12. **Initial Model In Error - Fix** – When Bayes Analyze read the input parameter file it detected a problem in the file. If you created the file manually, you need to correct the problem. Before Bayes Analyze proceeds to process the input model, Bayes Analyze will call the routine that computes the prior probability for all of the parameters. If this model is valid, in the sense that all of the parameters are physically reasonable, the routine that evaluates the prior probability will return a zero status. If the parameters are not physically reasonable, the routine will set a nonzero status and this is what happened. The types of problems that can cause this are: negative decay rate constants, the secondary coupling constant is greater than the primary coupling constant, one of the coupling constants or shimming amplitudes is negative. The number of things that can be wrong here is extensive, but it will be some type of logical problem. Getting an input field in the wrong place could also cause this problem.

13. **The input file is in an old format Vers('nnn')**– The input file either contains an bad file version or this is a file version that is no longer recognized by the Bayesian Analysis package.
14. **ASCII Fid File Error** – Either the input ASCII Fid file was not found, or it contained too few total data values.
15. **The Input Fid File Was Not Found** – The open for the input Fid file failed. The input file name is probably incorrect. This error can occur in either the Bayes Model or Bayes Analyze.
16. **The ASCII Parm File Was Not Found** – Either the file name in the parameters file is in error, or the input ASCII parameter file is not present. Either way Bayes Analyze could not find the file.
17. **The ASCII Parm File Is In Error** – The ASCII input parameter file is in error. Check the file and make the appropriate corrections. See Section 8.5.1 for more on this file.
18. **Input Procpa File Not Found – See log file** – The open failed on the procpa file. Either the name in the parameter or model file is incorrect, or the file has been deleted. The log file contains the name of the file the program was trying to read.
19. **Input Fid File Was Not Found – See log File** – The open failed for the Varian Fid file. Either the name in the parameter or model file is incorrect, or the file has been deleted. The log file contains the name of the file the program was trying to read.
20. **Adding Freq Would Cause Max Freq To Be Exceeded** – The requested analysis has exceeded the capabilities set at the time Bayes Analyze was compiled. You can determine the limits by running the command “bayes analyze limits” This will cause Bayes Analyze to print its hard coded limits.
21. **Adding Freq Would Cause MaxRes To Be Exceeded** – Same as above.
22. **Adding Freq Would Cause Rmax To Be Exceeded** – Same as above.
23. **Adding Freq Would Cause MGMAX To Be Exceeded** Same as above.
24. **Adding Resonance Would Cause RMAX To Be Exceeded** – Same as above.
25. **No of Fids exceeded max – see log file** – The requested number of Fids exceeded the maximum allowed for this version of Bayes Analyze. Either reduce the number of Fids requested or contact Varian for a version of the programs with larger limits. The log file contains the current limits.
26. **No of Fids exceeded actual – see log file** – The requested number of Fids exceeded the actual number in the array.
27. **Total data values exceeded max – see log file** – The requested number of data values exceeds the maximum allowed for this version of Bayes Analyze. You can determine the maximum allowed for the parameters by running the Unix command “bayes analyze limits” This will cause Bayes Analyze to print its hard coded limits.

28. **Total data values requested exceeded actual – see log file** – The number of data values that are suppose to be used is greater than the actual number of data values in the Fid.
29. **Requested Fid Is Not In Model File** – The input parameter file does not contain the parameters for the Fid you wanted to model.

Except for messages 37 and 38, the following list of errors are true program errors. If you receive one of them it means that Bayes Analyze has been pretty thoroughly messed up and you should contact Varian immediately. To illustrate what is meant by a “true” program error, consider message number 39. The routine SETMODEL has the function of taking a particular model and generating the time domain model. The model to generate is requested by the calling routine. The SETMODEL routine consists of a series of IF statements that determine the type of model to be processed and then calls the appropriate model dependent code. Message number 39 can only be obtained if these IF statements fail to recognize the model, the only way the model could be unrecognized is if we forgot to modify this routine when a new model was added; so this message represents a very significant problem in Bayes Analyze that should never be seen outside of testing. With the exception of messages 37 and 38 the other messages have similar implications. Here is the list of program errors:

30. **CONSTGIJ – Program Error Contact Varian**
31. **DEVZERO – Program Error Contact Varian**
32. **OUTMODEL – Program Error Contact Varian**
33. **PMSG – Program Error Contact Varian**
34. **PUTMODEL – Program Error Contact Varian**
35. **SETGIJ – Program Error Contact Varian**
36. **TRANS – Program Error Contact Varian**
37. **ALOGP – Possible Program Error Investigate Further**
38. **ALOGP2 – Possible Program Error Investigate Further**
39. **SETMODEL – Program Error Contact Varian**
40. **SCCETRANS – Program Error Contact Varian**
41. **SCUETRANS – Program Error Contact Varian**
42. **GETMODEL – Program Error Contact Varian**
43. **APPEND\_TEXT – Program Error Contact Varian**
44. **READHDR – Program Error Contact Varian**
45. **WRITEDATA – Program Error Contact Varian**
46. **WRITEMODEL – Program Error Contact Varian**

47. **WRITERESID** – Program Error Contact Varian

48. **COPY\_PROCPAR** – Program Error Contact Varian

49. **LIST\_MODEL** – Program Error Contact Varian

The two exceptions that we were discussing are messages numbers 37 and 38. These two messages are program errors, but they may be caused by the model. For example if the estimated frequencies become redundant, then the covariance matrix will be singular and these errors are issued. Message number 37 indicates that this matrix is singular. When this matrix becomes singular it is almost always because one or more of the resonances have converged to the same resonance frequency. This behavior may occur under several circumstances: first it is possible that a peak was marked and this peak was so far from the value indicated by probability theory that Bayes Analyze added one or more additional resonances near the same location. These resonances are nearly identical and this can cause the error. The easiest thing to try under these conditions is to remove the offending resonances and allow Bayes Analyze to run in its automatic model. There is one other condition where this problem has been known to occur. NMR lines never decay in a perfectly exponential manner because the magnetic fields are never perfectly uniform. When the signal-to-noise is very high Bayes Analyze may try to fit so many resonances at nearly the same resonance frequency that the matrix it is trying to invert becomes singular. Unfortunately, in this case, there is very little that can be done. You can try turning on the shimming models and rerunning the analysis, but if the problem re-occurs then you may not be able to analyze this data.

So far we have been discussing error number 37, but error number 38 is almost the same and it is caused by essentially the same conditions. The previous error occurs when the matrix inversion routine detects that the matrix is singular. Here the inversion routine did not detect failure and it returned a zero status. After the inverse is computed, Bayes Analyze computes the projection of the data onto the model. What it found was that the projection of the data onto the model was greater than the projection of the data onto itself. This is simply not possible, and when it happens it means that the matrix inverse failed.



# 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.