

Java GAMUT (JGAMUT) User Manual

Michael Birch

August, 2015

Contents

1	Installation and Starting the Program	1
2	The Interface	3
3	Using JGAMUT	5
3.1	Overview	5
3.2	The Input File	5
3.3	The Intermediate File	7
3.4	The Output Averages File	9
3.4.1	Gamma-by-Gamma Approach Output	9
3.4.2	GAMUT Method Output	9
3.5	Runtime χ^2 analysis in the GAMUT Algorithms	10
3.6	The Calculation Options	10
4	Mathematical Descriptions of the Algorithms	13
4.1	Level Clustering Algorithm	13
4.2	Linear Energy Shifts Calculation	15
4.3	“Gamma-by-Gamma” Approach	15
4.4	GAMUT Algorithm for Gamma-Ray Energies	16
4.5	GAMUT Algorithm for Gamma-Ray Intensities	17

Introduction

Java GAMUT (JGAMUT) is a revitalization of the original GAMUT code by R. B. Firestone [1]. The purpose of this software is to assist evaluators in creating the “ADOPTED LEVELS, GAMMAS” dataset in an ENSDF evaluation. The program provides two ways of taking input gamma-ray energy and intensity measurements and producing adopted values. The first is a “gamma-by-gamma” approach in which, for each gamma-ray, all measurements of that gamma’s energy and intensity are averaged to produce the adopted values. The second method uses the algorithms of the original GAMUT code, which use global (over all input datasets) fitting routines to produce adopted energies and intensities. Details concerning the methods used in this code are given in [chapter 4](#).

Note that this code is designed for use with gamma-ray data only! Levels for which no gamma-rays are known (for instance because they were measured in a particle transfer reaction) will not be used in any way by this program.

[1] Firestone, R. B. (1991). GAMUT: A Computer Code for gamma-ray Energy and Intensity Analysis.
- Report Number: LBL-26024

Chapter 1

Installation and Starting the Program

Since this code was developed in Java, a Java Runtime Environment (JRE) is required. In particular, this software was developed for use with the Oracle Java and may not be compatible with other Java Virtual Machines (e.g. OpenJDK). The latest Oracle JRE can be downloaded from the [Oracle website](http://www.oracle.com/technetwork/java/javase-downloads-1344955.html). Ubuntu users can instead follow the tutorial found at <http://www.webupd8.org/2012/09/install-oracle-java-8-in-ubuntu-via-ppa.html>.

Once the appropriate JRE is installed, the executable jar file “JGAMUT.jar” located in the “dist” folder of the main JGAMUT directory will start the program. On some systems, simply double-click it to start. On other systems, it may be necessary to launch the program from the command line using the command `java -jar JGAMUT.jar`.

Chapter 2

The Interface

The interface is shown in [Figure 2.1](#) and will be discussed in some detail here.

JGAMUT handles four different files: the input file, the intermediate file, the output averages file and the generated ADOPTED LEVELS, GAMMAS file. The full file path (including file name and extension) must be given for each of these files in the file name input boxes (numbered 1 to 4 in [Figure 2.1](#)). The “Browse...” buttons next to each of the file input text boxes can be used to select a file using a graphic interface instead of typing in the path manually. See sections [3.2](#), [3.3](#) and [3.4](#) for descriptions of the contents of these files.

JGAMUT has several options for the calculations which can be enabled or disabled by checking or unchecking the checkboxes. See [section 3.6](#) for details concerning these options.

JGAMUT performs actions on the files given in the file name input boxes when the buttons along the bottom of the window (the “Action Buttons” in [Figure 2.1](#)) are clicked.

The “Create Intermediate File” button generates the intermediate file and writes it to the location given in file name input box 2. If that file already exists then it is overwritten, so be careful! The intermediate file is generated using the input file specified in file name input box 1, for more details see [section 3.3](#).

The “Calculate Linear Energy Shifts” button will compute gamma-ray energy shifts of all the datasets with respect to a dataset chosen by the user using the information in the intermediate file. These shifts include a constant term as well as a term linear in energy (see [section 4.2](#) for details). The shifts can be used to correct systematic differences between energy calibration of different experiments before using the gamma-by-gamma averaging approach. They may also be useful before running the GAMUT algorithms in the case that an energy dependant shift is needed since the GAMUT energy algorithm only accounts for a constant shift between datasets. Note that the linear shift fitting will only be done on gamma-rays for which the standard dataset has an energy measurement, but the fitted constant shift will be applied to all the gamma-rays in each dataset (the linear term can only be applied to gamma-rays for which there is a measurement of the energy in the standard dataset though).

The “Remove Dataset from Intermediate file” button will allow the user to choose a dataset in the intermediate file to not be considered in any of the averaging methods. This is done by placing a hash (`#`) symbol in front of every instance of that dataset in the intermediate file. The hash symbol is the “comment” character used to mark a dataset or measurement as ignored in adopted value calculations, see [section 3.3](#) for more details.

The “Create Averages File (Gamma-by-Gamma)” button will use the gamma-by-gamma approach to generate adopted values from the information contained in the intermediate file. The output is written to the averages file specified in file name input box 3.

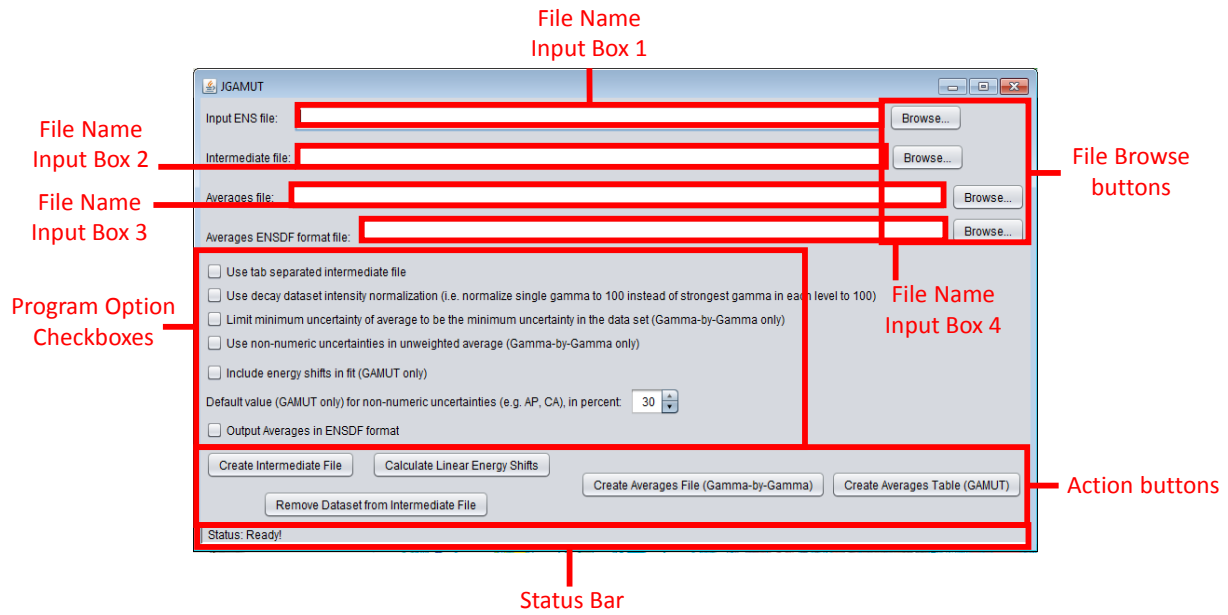


Figure 2.1: The main interface of JGAMUT

The “Create Averages Table (GAMUT)” button uses the GAMUT algorithms to generate adopted values from the information in the intermediate file and writes the output to the same averages file. Once again, be careful because the file specified by file name input box 3 is overwritten each time one of the middle or right buttons is pressed.

For the mathematical detail of what algorithms these buttons perform, see [chapter 4](#).

The status bar at the bottom of the interface informs the user what computation is being done. This is important since some calculations can be lengthy and the status bar changing helps to assure the user that the code is not frozen.

Chapter 3

Using JGAMUT

3.1 Overview

Here is a brief overview of the basic steps involved in using JGAMUT. For more details, see the other sections of this chapter.

1. Choose an [input ENSDF file](#) by clicking the first ‘Browse...’ button and selecting the file in the window which appears. JGAMUT will automatically fill in the remaining textboxes based on the file the user selected, however the user can edit these defaults. Alternatively, the user can select an [intermediate file](#) directly using the second ‘Browse...’ button.
2. If an intermediate file does not already exist for the chosen input, then press the ‘Create Intermediate File’ button to generate one. **Note: this process may take a few minutes to complete!**
3. Check the [intermediate file](#) for errors made in the [level matching process](#) and edit the file appropriately.
4. Click the ‘Create Averages File (Gamma-by-Gamma)’ button to [perform averaging](#) on the intermediate file and output the results. **Note: uncertainties are reported with two digits all the way up to 99. It is up to the evaluator to round the uncertainties as he or she feels is reasonable.**
5. Click the ‘Create Averages Table (GAMUT)’ button to perform the [GAMUT algorithms](#) on the intermediate file and output the results. **Note: uncertainties are reported with two digits all the way up to 99. It is up to the evaluator to round the uncertainties as he or she feels is reasonable.**

3.2 The Input File

The input file should consist of all ENSDF datasets relating to the nuclide of interest, with a blank line (end record) separating one dataset from the next. A minimal example is shown below for the case of ^{76}Se (note that this example is **fictional** and does not actually represent the current knowledge of the actual ^{76}Se nucleus):

```
76SE      ADOPTED LEVELS, GAMMAS
76SE PN
```

```
95NDS      199503
6
```

```

76SE L 0.0 0+ STABLE A
76SE L 559.101 5 2+ 12.3 PS 2 A
76SE G 559.099 5 100 E2
76SEB G BE2W=44.0 8
76SE L 1122.280 8 0+ 11 PS 5
76SE G 563.171 7 100 E2
76SEB G BE2W=47 22
76SE G 1122.3 3 E0 0.023 5
76SE L 1216.150 6 2+ 3.4 PS 2 B
76SE G 657.041 5 100.0 22E2+M1(+E0) +5.2 2
76SEB G BE2W=43 3
76SE G 1216.149 25 58.0 22 E2
76SEB G BE2W=1.20 9
76SE L 1300.0 0+
76SE L X
76SE L 100.0+X
76SE G 100.1 10 100
76SE L X+150.0
76SE G 50.5 5 36 5
76SE G 149.8 15 100 1

76SE 76GE 2B- DECAY (1.84E+21 Y) 2013AG11,2013AG02,2004KL01
76SE N 1
76SE L 0 0+

76SE 76AS B- DECAY (26.254 H) 1998DE26,1982WE14,1980KA3695NDS 199503
76AS P 0.0 2- 26.254 H 11 2960.6 9
76SE N 0.4067 29 1.0 1.0
76SE PN 3
76SE L 0.0 0+
76SE L 559.100 9 2+ 23 PS 15
76SE2 L G=+0.40 12
76SE G 559.086 10 100 E2 0.00206 BC
76SE L 1122.257 12 0+
76SE G 563.147 10 2.69 6 E2 0.00192 B
76SE L 1216.153 10 2+
76SE G 657.042 10 14.0 3 E2+M1(+E0) +5.2 2 BC
76SE G 1216.200 20 7.60 12 E2 BC
76SE L 1300.0 0+

76SE 76BR EC DECAY (16.2 H) 1969DZ01,1974NA17,2004SH1795NDS 199503
76BR P 0.0 1- 16.2 H 2 4963 9
76SE N 0.741 6 1.0 1.0 1.0
76SE PN 3
76SE G 604.4 3 0.20 2
76SE G 2310.7 10 0.10 1
76SE G 4432.0 20 0.08 5
76SE DG E$4440 3 IN 1969DZ01, 1975VYZX; 4437.7 10 IN 2004Sh17
76SE L 0.0 0+
76SE L 559.12 4 2+

```

76SE	G	559.11	5	100	E2				C
76SE	L	1122.40	6	0+					
76SE	G	563.22	5	3.90	20	E2			C
76SE	G	1122.3	3			E0		0.00082	
76SE	L	1216.15	4	2+					
76SE	G	657.00	5	21.4	11E2+M1(+E0)	+5.2	2		C
76SE	G	1216.10	5	12.0	6				C
76SE	L	1300.0		0+					
76SE	L	X							
76SE	L	100.1+X							
76SE	G	100.5	10	100					
76SE	L	X+151.0							
76SE	G	50.8	5	37	6				
76SE	G	150.7	15	100	1				
76SE		74GE(A,2NG)				1982MA45,1984Z001,1980WE0795NDS			199503
76SE	PN								5
76SE	L	0.0		0+					A
76SE	L	558.7	4	2+					A
76SE	G	558.8	5	100	E2				C
76SE	L	1215.7	4	2+		3.5 PS	14		B
76SE	G	656.8	5	15.4	11 E2+M1	+4.7	+11-20		C
76SE	G	1215.6	5	5.8	4 E2				C
76SE	L	1300.0		0+					
76SE	L	X							
76SE	L	102.6+X							
76SE	G	101.9	13	100					
76SE	L	148.6+X							
76SE	G	48.0	11	35	8				
76SE	G	147.7	18	100	1				

Note that if there are floating level structures not connected to the main level scheme (as with the “+X” levels in the above example), then each floating structure must use a different place holder and these must be consistent across all the datasets. For example, if all the X’s in “74GE(A,2NG)” were changed to Y’s then that floating structure would not be recognized as the same as the ones given in “76BR EC DECAY” and “ADOPTED LEVELS, GAMMAS”, thus causing a problem when attempting to average all the measurements of that floating structure. It is also important that ‘X’ mean the same number across all datasets. E.g. even if the 101.9(13) and 147.7(18) gamma-rays were not observed in the ‘74GE(A,2NG)’ dataset, and hence the level currently labelled ‘X’ was not observed, then the last level in that dataset would still need to be labelled “150+X” in order to be consistent with the ‘X’ which appears in “76BR EC DECAY”.

3.3 The Intermediate File

The intermediate file is a different representation of the information contained in the input file. The purpose of this file is to make it easier for both the evaluator and the program itself to work with the data. The structure of this file is as follows. The levels for the nuclide of interest are

listed in order of energy and indicated by a truncated (up to the J field) ENSDF level record. If an ADOPTED LEVELS, GAMMAS dataset was present in the input file then that record is used. In that case two astriks (**) will follow the level record, indicating to the user that the level appeared in the adopted dataset. If only a single dataset listed a level, then that one level record is given as the level indicator in the intermediate file. Otherwise the record is generated by the program and the quoted energy is simply the unweighted average of all the levels from the individual datasets which were the same. Following each level record is a table which lists all the gamma-rays which have been observed depopulating that level. The first column gives the name of the input dataset from which that gamma-ray was taken. The remaining columns (“Eg”, “Ig”, “Ig (norm.)”, “Init. Level”, “Final Level”, “Mu”, “MR”) give the gamma ray energy, intensity, normalized intensity (in either the ADOPTED LEVELS, GAMMAS convention or decay data convention), the initial level, the final level, the multipolarity and the mixing ratio, respectively. Each of these quantities is taken from the original dataset where the gamma ray came from, except the initial and final level energies, which must **exactly** match one of the level records in the intermediate file. There is a blank line between sets of gammas which are the same. The energy and intensity columns will also show relevant ENSDF flags, such as placement uncertain (‘?’) and intensity split (‘@’).

The columns of the intermediate file are separated by a double space (i.e. two space characters; ‘ ’) and cells are padded with additional spaces to ensure all the columns align. The reason for this format is for maximum user readability, however such a format is not as easy for a user to produce from scratch using data which is not in the ENSDF format. For this reason, there is an option to instead separate the columns of the intermediate file by a single tab character rather than spaces. Simply check the first check box below the file name input boxes to enable this feature. See [section 3.6](#) for more details.

The intermediate file is generated from the input datasets by matching levels between datasets. This level matching procedure (see [section 4.1](#) for details of the algorithm) is not perfectly accurate (although it is quite good!) and therefore the evaluator must be sure that the intermediate file is correct. **Since the intermediate file is used for all adopted value calculations it is very important for the intermediate file to be correct.** For example, occasionally the program fails to recognize when two levels are the same and hence includes more levels than actually exist in the nucleus. Such errors are corrected by the evaluator by editing the intermediate file. **When editing a space-separated intermediate file, make sure that any changes you make do not cause the columns of the table to become misaligned! For each character added, remove a space from the same row.** The previous warning does not apply to tab-separated intermediate files, where the columns are simply separated by a single tab character and alignment is not important.

The evaluator can also edit the intermediate file to modify the input data. For instance, an underestimated uncertainty could be increased by simply changing the number in the intermediate file. Data can also be excluded from the adopted value calculations by adding a hash (‘#’) symbol.

For example, consider the excerpt from an edited space-separated intermediate file below (note some columns of the file are not shown):

```
76SE  L      0.0      0+   **

76SE  L 559.101    5 2+   **
Data set                                Eg          Ig (orig.)  Ig (norm.)  Init. Level
ADOPTED LEVELS, GAMMAS                 559.099 5      100          100          559.101 2+
```

74GE(A,2NG)	558.8 5	100	#100	559.101 2+
#76BR EC DECAY (1.31 S)	559 ?	100	100	559.101 2+
76SE(P,P'G),(A,A'G)	#559.00 10		100	559.101 2+
76AS B- DECAY (26.254 H)	559.086 10	100	100	559.101 2+

In the above example, the intensity from the dataset “74GE(A,2NG)” will not be considered in the gamma-by-gamma average, no quantities from the dataset “76BR EC DECAY (1.31 S)” will be considered in the adopted values calculation (by either method) and the energy from the data set “76SE(P,P'G),(A,A'G)” will also not be considered. Note that if the energy of a gamma-ray is excluded then its intensity will also **not** be used in the GAMUT algorithm, however it will be used in the gamma-by-gamma approach. The converse does not hold, if there is a hash symbol in the “Ig (orig.)” column then this has no effect on the energy column and the gamma ray energy will still contribute to the GAMUT adopted value calculation. Also note that it is the normalized intensity (“Ig (norm.)”) which is used by the gamma-by-gamma approach to determine the adopted intensity, while the original intensity is used in the GAMUT algorithm. Hash symbols in the normalized intensity column have no effect on the GAMUT algorithms, but cause that measurement to not be considered in the gamma-by-gamma intensity average for that gamma-ray.

3.4 The Output Averages File

The output files for the gamma-by-gamma approach and GAMUT algorithms differ. Each is explained separately in the subsections below.

3.4.1 Gamma-by-Gamma Approach Output

The gamma-by-gamma output is structured as follows. As with the intermediate file, levels are indicated by the beginning (up to the J field) of an ENSDF level record. Below each level is a list of the gamma-rays depopulating that level. The first column gives the energy, while the second column gives the intensity. These numbers are obtained by the averaging routine detailed in [section 4.3](#). If an averaging procedure different than the weighted average is used to obtain the result then the name of the method is annotated next to the result in asterisks. An excerpt example gamma-by-gamma output file is shown below

```
-----
76SE  L 1216.150   6 2+
-----
657.041  4                100.0 18
1216.150 22                59 5  *UnwtAve*
```

In the above example the level with E=1216.150 has two gamma rays. Both energies were obtained from a weighted average of the inputs, however the second gamma intensity was determined by an unweighted average.

3.4.2 GAMUT Method Output

The GAMUT output file is formatted similarly to the intermediate file, with a few notable differences. Firstly, the level energies in the level records which appeared in the intermediate

file are modified to become the result of the GAMUT fit (since level energies are what is fitted and the gamma-ray energies are determined by level energy differences). Secondly, the GAMUT adopted gamma-ray (with the computed energy and intensity) is given first for each group of gammas. The energy and intensity of GAMUT gamma-rays are calculated using the algorithms detailed in [section 4.4](#) and [section 4.5](#). Third, the “Mu” and “MR” columns are no longer present and instead there are “Chi² (Eg)” and “Chi² (Ig)” columns, which give that measurement’s contribution to the overall χ^2 . These numbers can be used to help spot outliers in the data. A summary of the χ^2 ’s by dataset is given at the top of the file to help indicate when an experiment may be less reliable than others. The summary also includes the dataset intensity multipliers (α_j in [section 4.5](#)) and energy shifts (Δ_j in [section 4.4](#)). Note that the χ^2 ’s given in the summary are reduced, i.e. divided by the number of measurements in the indicated dataset. Finally, gamma-rays which were modified in the calculation (see the next section) are annotated with asterisks in this file and levels which did not contribute to the calculation (because they have no gamma-rays associated with them) are also annotated as ignored.

3.5 Runtime χ^2 analysis in the GAMUT Algorithms

The GAMUT algorithms use global fitting procedures to determine adopted energies and intensities. Once a fit is complete GAMUT automatically checks for outlying measurements, those which differ by four or more standard deviations from the fitted value (three or more in the case of energies). When such an outlier is found, the user is alerted in the form of a message box. In addition, another window appears which has the part of the intermediate file table that includes the outlying gamma-ray, with the outlier highlighted in yellow to help the evaluator locate it easily. The additional window provides the context of that gamma-ray which is necessary for the evaluator to decide how to handle it. The message box asks the user to increase the uncertainty of the discrepant quantity (energy or intensity) in order to make it consistent with the fit. If the evaluator responds yes, then the change is made and the fit is repeated. If the evaluator responds no then a second prompt will appear asking if the evaluator wishes to enter some other uncertainty for the measurement (perhaps the evaluator felt that the suggested uncertainty change was too severe). If the evaluator responds yes to this prompt then a new window appears where a custom uncertainty can be given and once the new uncertainty is input then the change is made and the fit is repeated. This process will continue until there are no more outliers or the evaluator responds no to the second prompt. Note that an especially discrepant measurement can be simply excluded from the calculation by adding a hash symbol (‘#’) in the intermediate file and re-running the GAMUT algorithms. An example message and corresponding highlighted intermediate file excerpt are shown in [Figure 3.1](#).

3.6 The Calculation Options

JGAMUT offers some options to the user for fine-tuning its behaviour when performing computations.

The first option is meant to allow the user to create his or her own intermediate files easily from data which is not in the ENSDF format. When checked, this option will cause the program to assume that a single tab character separates the columns of the intermediate file, instead of the space character separation. Using this option, one could create a tab-separated table (e.g. from a spreadsheet) which has the same format as the intermediate file and use that as input

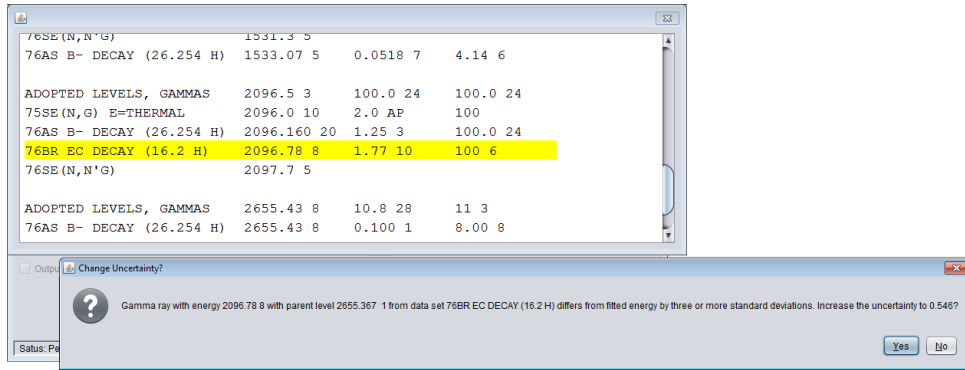


Figure 3.1: An example of a message asking the user to modify a discrepant gamma-ray.

to the algorithms offered by JGAMUT without having to translate the data into the ENSDF format first. This feature could be useful in, for instance, decay data evaluations. Sample input and output illustrating this feature is available in the “exampleFiles” folder in the main JGAMUT directory. These files are based on the decay of ^{57}Co . Also note that when this option is checked the “Create Intermediate File” button will also produce a tab-separated file from an ENSDF input.

The second option changes how the normalized gamma-ray intensities are calculated from the ADOPTED LEVELS, GAMMAS convention (i.e. strongest gamma is 100 for each level) to the convention used in decay datasets (i.e. the strongest level in the entire dataset is 100).

The third option applies to the gamma-by-gamma approach only. When checked, it will ensure that the uncertainty on a result will be no smaller than the uncertainty of any of the measurements used in the average, even if the uncertainty calculated by the averaging routine would have been smaller.

The fourth option only applies to the gamma-by-gamma unweighted average. Since uncertainties are not considered in an unweighted average it is easy to include measurements with non-numeric uncertainties in the calculation. When checked, this option includes all measurements (even those with non-numeric uncertainties) in the unweighted average. Note that measurements with non-numeric uncertainties are always excluded from the weighted average and NRM methods, even with this option checked.

The fifth option applies to the GAMUT energy fitting algorithm only. When checked, it includes the systematic energy shifts in the fitting. See [section 4.4](#) for details on what these energy shifts mean and accomplish.

The GAMUT algorithms assign a default uncertainty to measurements with a non-numeric uncertainty. Such measurements often have little impact on the over all fit because they will be weighted very low compared to measurements which were given a numeric uncertainty initially. The number given in the seventh option determines what percent uncertainty should be assigned by default to non-numeric uncertainties. Note that no uncertainty counts as non-numeric. Data points which have non-numeric uncertainties which the evaluator wishes to exclude entirely must be explicitly excluded using the hash (‘#’) symbol in the intermediate file.

The seventh option will create an ADOPTED LEVELS, GAMMAS dataset which contains the results of the adopted values calculations (i.e. the same numbers which appear in the output averages file, but in the proper ENSDF format). This dataset is saved to the file indicated by file name input box 4 (see [Figure 2.1](#)). If an ADOPTED LEVELS, GAMMAS dataset existed in

the input file then that dataset will be used as a template and only the gamma-ray energies and intensities (also the level energies in the GAMUT case) will be changed. Note that changing the numbers in an existing ADOPTED LEVELS, GAMMAS dataset requires level matching, which as with the intermediate file, can be sometimes not be done perfectly, so the evaluator must **always check the generated ADOPTED LEVELS, GAMMAS dataset to ensure it is correct.**

Chapter 4

Mathematical Descriptions of the Algorithms

4.1 Level Clustering Algorithm

The first step in creating the intermediate file is to determine which levels from different datasets are the same (i.e. different observations of the same physical level). This is done via a clustering algorithm called k -Medoids, together with a criterion used to determine the number of clusters. This algorithm is similar to the popular k -means algorithm, but allows for more general distance functions than the Euclidean distance. In the case of nuclear levels we use the following distance function:

$$\rho(L_1, L_2) = \delta(\pi_1, \pi_2) + \delta(J_1, J_2) + \frac{|E_1 - E_2|}{\sqrt{dE_1^2 + dE_2^2}} + (1 - \delta(s_1, s_2)), \quad (4.1)$$

where $\delta(\cdot, \cdot)$ is a Kronecker delta for Strings, π_i is the parity of level i , J_i is the spin of level i , E_i is the energy of level i , dE_i is the uncertainty on E_i and s_i is the name of the dataset level i was observed in (i.e. the source of L_i). The String Kronecker delta has the additional property that $\delta(\cdot, " ") = 0$ since blank strings provide no information. The first two terms constrain the spin and parities to be the same (if they are known) and the final term constrains that the two levels come from different experiments (since the same level cannot be observed twice in the same experiment). Note that the expression above is formally not a distance metric since $\rho(L_1, L_1) \neq 0$ due to the presence of the last term (each level is observed in the same experiment as itself), so the above expression only applies when $L_1 \neq L_2$ and we make ρ into a distance function by defining $\rho(L_1, L_2) = 0$ when $L_1 = L_2$. This is the distance function used to cluster the levels in the k -Medoids algorithm described below.

The k -Medoids algorithm proceeds as follows. The number of clusters, k , is given *a priori*. Let the group of elements to be grouped be denoted by Y and the distance metric on pairs of objects in Y be denoted by ρ . First, k elements of Y are chosen randomly to be the initial set of “medoids”, M (which will serve as the representative for their respective groups). The k groups are made by associating each element of Y with its closest medoid, i.e. let $y \in Y$ and $m \in M$ then y is grouped with m if $\rho(y, m) \leq \rho(y, \mu)$ for all $\mu \in M$. The cost of a given group configuration, G , is the sum of the distances of the members of each group from their medoid, i.e. $C(G) = \sum_{g \in G} \sum_{y \in g} \rho(y, m_g)$, where m_g is the medoid of group g . The set of medoids is updated at each step in the following way: for each $m \in M$ and each $y \in Y \setminus M$ if swapping m and y (i.e. making y a medoid and removing m from the set of medoids) produces a group configuration of lower cost then this swap is performed. The algorithm terminates when no swap lowers the cost, i.e. the current medoid set is locally optimal.

The correct number of groups to use in the clustering is determined from two indicators: the “silhouette” (P.Rousseeuw, *Journal of Computational and Applied Mathematics*, 20, p. 53 (1987)) and the internal to external dissimilarity ratio. First we define the following distance metrics derived from ρ

$$D(y, g) = \sum_{x \in g} \rho(x, y) \quad (4.2)$$

$$D(g_1, g_2) = \frac{1}{2} \sum_{x \in g_1} \sum_{y \in g_2} \rho(x, y), \quad (4.3)$$

where $y \in Y$, $g, g_1, g_2 \subseteq Y$. The silhouette is defined for an element $y \in Y$ and a grouping, G , of the elements of Y as

$$s(y, G) = \frac{b(y) - a(y)}{\max\{a(y), b(y)\}}, \quad (4.4)$$

where $a(y) = D(y, g_y)/|g_y|$, $g_y \in G$ is the group which contains y , $|g_y|$ denotes the number of elements in g_y , is the average distance of y from other members of its group and $b(y) = \min_{\{g \in G | g \neq g_y\}} D(y, g)/|g|$ is the average distance to members of the closest group to y . The silhouette for the group configuration is then

$$s(G) = \frac{1}{|Y|} \sum_{y \in Y} s(y, G). \quad (4.5)$$

Note that since G is determined by the k -Medoids algorithm above, s actually only depends on the number of groups used and we can safely write $s(k)$.

Evidently, $-1 \leq s \leq 1$, where s close to 1 indicates the elements are clustered well. Unfortunately, when $k = |Y|$ then $s(G) = 1$ since each element is in its own group and therefore we cannot simply choose the k with the largest silhouette. Instead, we let the silhouette inform which k is most appropriate by choosing the one for which the silhouette is sufficiently high ($s(k) > 0.95$) and for which $s(k+1) < s(k)$.

This criterion alone is not enough to ensure the correct number of groups is chosen in all cases. Therefore, another criterion is also used to judge how many groups should be used. Define the variation ratio, v , for a grouping configuration as

$$v(G) = \frac{\sum_{g \in G} D(g, g)}{\sum_{g_1 \in G} \sum_{g_2 \neq g_1} D(g_1, g_2)}, \quad (4.6)$$

where the numerator represents the total within-group variation, while the denominator represents the total variation between the groups. Once again, since G is determined by the k -Medoids algorithm we really think of v as a function of the number of groups, k . Smaller values of $v(k)$ indicate better clustering, however $v(|Y|) = 0$, again since each element would be in its own group, so we cannot simply choose the k with the smallest variation ratio. Instead, it has been observed that sudden decreases in the value of v indicate that the correct number of clusters has been reached. A number of groups k is selected using the variation ratio criterion if

$$\ln(v(k)) - \ln(v(k-1)) < -3.$$

The number of groups is selected within the JGAMUT code when either of the two criteria are satisfied.

4.2 Linear Energy Shifts Calculation

Linear shifts in the gamma-ray energies are calculated with respect to a “standard” dataset, which is chosen by the user. These shifts are parametrized by the slopes, m_j and intercepts b_j as follows:

$$E_i^{(j)} + (m_j E_i^{(s)} + b_j) = E_i^{(s)}, \quad (4.7)$$

where $E_i^{(j)}$ is the energy of the i -th gamma-ray as measured in the j -th dataset and $E_i^{(s)}$ is the energy of the i -th gamma-ray as measured in the standard dataset. Define $\Delta E_{ij} = E_i^{(j)} - E_i^{(s)}$, then the set of linear equations given in Equation 4.7 can be rewritten as a single matrix equation,

$$\Delta \mathbf{E} = A \mathbf{x}, \quad (4.8)$$

where $\Delta \mathbf{E} = (\Delta E_{11}, \Delta E_{21}, \Delta E_{31}, \dots, \Delta E_{12}, \Delta E_{22}, \Delta E_{32}, \dots)^T$, $\mathbf{x} = (m_1, m_2, \dots, b_1, b_2, \dots)^T$ and

$$A = \begin{bmatrix} -E_1^{(s)} & 0 & 0 & \dots & -1 & 0 & 0 & \dots \\ -E_2^{(s)} & 0 & 0 & \dots & -1 & 0 & 0 & \dots \\ -E_3^{(s)} & 0 & 0 & \dots & -1 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \\ 0 & -E_1^{(s)} & 0 & \dots & 0 & -1 & 0 & \dots \\ 0 & -E_2^{(s)} & 0 & \dots & 0 & -1 & 0 & \dots \\ 0 & -E_3^{(s)} & 0 & \dots & 0 & -1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Solving this system by the standard weighted least-squares method we have

$$\mathbf{x} = (A^T W A)^{-1} A^T W \Delta \mathbf{E}, \quad (4.9)$$

where W is the diagonal matrix of weights with elements given by the reciprocal uncertainties, $\left((\Delta E_i^{(j)})^2 + (\Delta E_i^{(s)})^2 \right)^{-1}$. The uncertainties in the fitted slopes and intercepts are given by

$$\Delta x_k^2 = \sum_i \left(\frac{\partial x_k}{\partial (\Delta \mathbf{E})_i} \right)^2 \sigma_{(\Delta \mathbf{E})_i}^2 \quad (4.10)$$

$$= \sum_i W_{ii} (V A^T)_{ki}^2, \quad (4.11)$$

where $V = (A^T W A)^{-1}$.

4.3 “Gamma-by-Gamma” Approach

The gamma-by-gamma approach treats each set of gamma-ray energy measurements and gamma-ray intensity measurements as a separate dataset and computes the average of these datasets to obtain adopted values. The averaging algorithms used are the same as those implemented in the Visual Averaging Library (V.AveLib), a code also available through the IAEA. The way in which the averaging is done is as follows:

1. The weighted average is calculated. If the χ^2 is smaller than the critical χ^2 (at 99% confidence), then the result is accepted, otherwise go to 2.
2. The NRM (M.F. James, R.W. Mills, D.R. Weaver, *Nucl. Instr. and Meth. in Phys. Res.* **A313**, 277 (1992)) is used to compute the average. If the χ^2 is smaller than the critical χ^2 (at 99% confidence), then the result is accepted, otherwise go to 3.
3. The unweighted average is calculated and accepted since the weighted average and NRM failed to produce an acceptable result.

4.4 GAMUT Algorithm for Gamma-Ray Energies

Measured gamma-ray energies are related to the level energies of the nucleus according to the following set of linear equations

$$E_i^{(j)} = \sum_{k=1}^m g_{ik} \xi_k - \Delta_j, \quad (4.12)$$

where $E_i^{(j)}$ is the j -th measurement of the i -th transition, ξ_k is the k -th level energy, Δ_j is a systematic energy shift for the j -th experiment and g_{ik} is the placement coefficient given by

$$g_{ik} = \begin{cases} 1 & \text{if the } i\text{-th transition deexcites the } k\text{-th level} \\ -1 & \text{if the } i\text{-th transition populates the } k\text{-th level} \\ 0 & \text{otherwise} \end{cases}$$

These linear equations can be rewritten into a single matrix equation

$$\mathbf{E} = G\xi, \quad (4.13)$$

where \mathbf{E} is the vector of gamma-ray energy measurements organized by dataset,

$\mathbf{E} = \left(E_1^{(1)}, E_2^{(1)}, E_3^{(1)}, \dots, E_1^{(2)}, E_2^{(2)}, E_3^{(2)}, \dots, E_1^{(3)}, E_2^{(3)}, E_3^{(3)}, \dots \right)^T$, G is the placement matrix augmented with additional columns for the energy shifts and ξ is the vector of level energies and energy shifts, $\xi = (\xi_1, \xi_2, \xi_3, \dots, \Delta_1, \Delta_2, \Delta_3, \dots)^T$.

As stated, this problem has infinitely many solutions since the zero level has not been specified. We can alleviate this problem by adding an extra equation, $\xi_1 = 0$, which amounts to adding an extra row to G and an extra element to \mathbf{E} . With this modification, we can then solve Equation 4.13 using the standard weighted least-squares method

$$\xi = (G^T W G)^{-1} G^T W \mathbf{E}, \quad (4.14)$$

where W is a diagonal weight matrix with coefficients given by the squared reciprocal uncertainty of the measurement, $W_{ii} = 1/\Delta E_i^2$. The added element in \mathbf{E} for the zero level constraint also requires a weight, which technically should be infinite since $\xi_1 = 0$ is a *definition* and hence has zero uncertainty. However, we numerically weight the constraint by 100 times the sum of all the other weights. Also note that in the numerical implementation of this algorithm the Moore-Penrose pseudoinverse is calculated using singular value decomposition in place of $(G^T W G)^{-1}$. The pseudoinverse agrees with the actual inverse in the case that $G^T W G$ is invertible, but also gives a solution in the case that it is not invertible. The solution in that case is the minimum norm solution in the sense that $\|\xi\|$ will be smaller than all other solutions to $G^T W G \xi = G^T W \mathbf{E}$. A warning is given to the user in the case that $G^T W G$ is not invertible since the Moore-Penrose solution may not be the one desired by the evaluator.

The uncertainties in the fitted level energies and energy shifts are given by

$$\Delta\xi_k = \sigma\sqrt{V_{kk}}, \quad (4.15)$$

where $V = (G^T W G)^{-1}$ is the covariance matrix and

$$\sigma^2 = \frac{1}{n-m} (\mathbf{E}^T W \mathbf{E} - \mathbf{E}^T W G \xi)$$

for a data set with n transitions (unique gamma-rays) and m levels. The adopted gamma energies are given by the differences in fitted level energies,

$$E_i = \xi_p - \xi_f \quad (4.16)$$

and their uncertainties are given by

$$\Delta E_i = \sigma (V_{pp} + V_{ff} - 2V_{pf})^{1/2}, \quad (4.17)$$

where p and f are the indices of the parent (initial) and final levels, respectively. Notice that Equation 4.17 could also be written as $\sigma\sqrt{\sum_{\text{elements}} (\mathbf{G}_i \mathbf{G}_i^T) \circ V}$, where \mathbf{G}_i is the column vector with elements from the i -th row of G (not including the energy shift columns), \circ denotes the element-wise product of matrices and the sum is over all the elements of $(\mathbf{G}_i \mathbf{G}_i^T) \circ V$.

4.5 GAMUT Algorithm for Gamma-Ray Intensities

Intensity measurements between different experiments can be on different arbitrary scales. In this method, we assume that the scales are linearly related by factors α_j ,

$$I_{ij}\alpha_j = \bar{I}_i,$$

where I_{ij} is the intensity measured for the i -th gamma-ray by the j -th experiment and \bar{I}_i is the intensity of the i -th gamma-ray on the common scale. We choose $\alpha_1 = 1$ by definition (i.e. we place all gamma-ray intensities on the same scale as the first experiment) and fit the remaining scale factors together with the common intensities by minimizing the global χ^2 ,

$$\chi^2 = \sum_i \sum_j \frac{(I_{ij}\alpha_j - \bar{I}_i)^2}{(\alpha_j \Delta I_{ij})^2}, \quad (4.18)$$

with respect to α_j and \bar{I}_i . We define $\omega_{ij} = \Delta I_{ij}^{-2}$ and $\beta_j = \alpha_j^{-1}$, then we can write the global χ^2 more simply as

$$\chi^2 = \sum_i \sum_j \omega_{ij} (I_{ij} - \beta_j \bar{I}_i)^2. \quad (4.19)$$

The minimization is then performed numerically using the coordinate decent method as follows. Minimization along the \bar{I}_i directions can be done analytically for fixed $\{\beta_j\}$ since the weighted average minimizes the χ^2 , hence given $\{\beta_j\}$ we write

$$\bar{I}_i = \frac{\sum_j \omega_{ij} \beta_j I_{ij}}{\sum_j \omega_{ij} \beta_j^2}. \quad (4.20)$$

Minimization along each of the β_j directions is done using the golden search method. The global minimization scheme is thus

1. Initialize $\beta_j = 1$ for $1 \leq j \leq N$, where N is the number of experiments
2. Calculate \bar{I}_i according to [Equation 4.20](#)
3. For each j , where $2 \leq j \leq N$, minimize [Equation 4.19](#) with respect to β_j (all other variables fixed) using a golden search
4. If the resulting change in (the norm of) $\{\beta_j\}$ is less than 0.01% then exit, otherwise go to 2.

Once the minimization has converged, \bar{I}_i gives the adopted intensity for the i -th gamma ray. This adopted intensity can be put onto the intensity scale of the j -th experiment by multiplying by β_j .