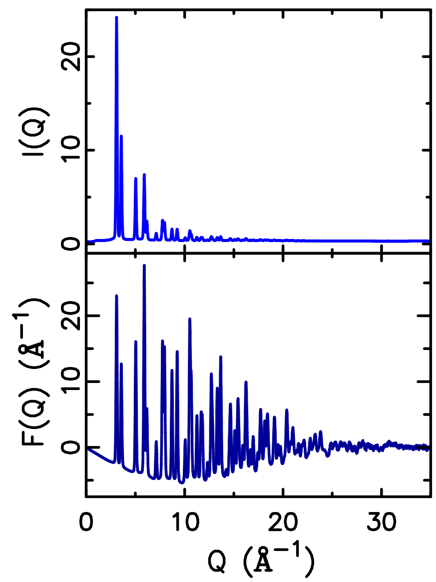
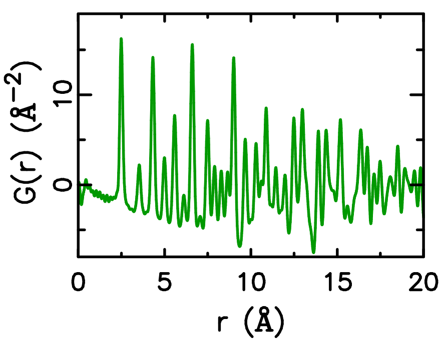
PDFgetX2 manual





written by

Xiangyun Qiu
email: xiangyun@pa.msu.edu

Simon Billinge
email: billinge@pa.msu.edu

Department of Physics and Astronomy, Michigan State University,
East Lansing, Michigan, 48824-2320, USA

Document created: July 22, 2004

Preface

PDFgetX2 Software License Agreement

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 What is PDFgetX2?

PDFgetX2 is a user friendly program to obtain the atomic pair distribution function (PDF) from X-ray powder scattering data. The interaction between users and data is facilitated by a rather extensive graphical user interface (GUI). The supported platforms include Linux, Windows, and Mac OSs. Please refer to Appendix A for detailed installation instructions. The supported file formats include SPEC, CHI, and free data format. Please refer to Appendix B for details.

The PDF, $G(r)$, is a real space function telling the probability of finding pairs of atoms separated by distance r in the material. The experimental PDF can be thought of as the sum of snapshots of the instantaneous atomic arrangements over the data collection time. Accordingly. the PDF can tell us about both the *local* structure (low r region) and average structure (high r region). The PDF method has been used as early as 1931 by Debye [1] to study liquid mercury, and has mostly focused on glass and amorphous materials [2; 3]. Not until recently, the PDF analysis has been successfully applied to crystalline and nanocrystalline materials [4]. We would like to direct you to a recent book by Egami and Billinge (2003) [5] for more technique details and recent PDF applications.

The PDF, $G(r)$, is obtained by a direct sine Fourier transformation of the structure function $S(Q)$, where Q is the magnitude of the scattering vector. However, the collected raw data during X-ray scattering experiments subject to various instrument and sample effects, which have to be properly corrected to extract the $S(Q)$. This program, PDFgetX2, lives to serve this purpose. Standard corrections [5; 6; 7] implemented include subtraction of background scattering, sample absorption, X-ray polarization, unwanted Compton intensity, and normalization by the average atomic scattering power. Particularly for the RA-PDF experiments recently developed by Chupas *et al.* [8] utilizing the image plate area detector, additional corrections due to oblique incident angle dependence [9] and the detector energy dependence [10] are also implemented. Statistical uncertainties due to limited intensity counts are estimated and are propagated up to the PDF, $G(r)$, which becomes very important when as the $S(Q)$ and $G(r)$ data are increasingly being modeled using regression algorithms. In addition, simple data smoothing and damping are also possible.

A standing alone functionality in this program is to preprocess the raw SPEC data to a single 4-column data set. You have the options to apply the detector dead time corrections on monitor and detector counts; measured intensities from different scans can be either normalized by their collection time or monitor counts. Scaling between different scans when merged into a single set is automatically computed to achieve the best overlap. Error propagation is also carried out.

All features of the program is accessible from the built-in GUI, however, manually editing data processing parameters are possible through a text window. One noteworthy feature is the automatic addition of data processing parameters into the $S(Q)$ and $G(r)$ files as header history information. The history information can also be saved alone. Restoration of the history information will reproduce the original $S(Q)$ and $G(r)$ which may become much desirable when careful examination of data processing is found necessary at a later stage.

1.2 What is new?

To find out about recent updates of PDFgetX2 or to get further information, please visit the PDFgetX2 homepage at the following site:

http://www.totalscattering.org/programs/PDFgetX2/

1.2.1 Version 1.0

Some bugs in the program have been fixed since the Beta release. Some GUIs are enhanced. The most notable change is $S(Q)$ optimization routine (still under test).

1.2.2 Beta release

The PDFgetX2 coding project started out as an upgrade to the Yorick program PDFgetX [6] developed in our group. The interactive data language (IDL[1]) was chosen for its cross-platform compatibility, and built-in GUI capability. The data correction methods used in PDFgetX were inherited. However, this release of PDFgetX2 is not a *translation* of program PDFgetX, i.e. with only an addition of GUI. The two programs differ a lot in programing structure and concepts, but do give quantitatively consistent results. As the coding efforts moved on, numerous new features have been constantly added, and some existing capabilities were expanded. A tour of the features available in PDFgetX2 is best made by going through the tutorial data provided in this manual.

1.3 Using PDFgetX2

Publication of results totally or partially obtained using the program PDFgetX2 should state that PDFgetX2 was used and contain the following reference:

[1]IDL is a registered trademark of Research Systems, Inc. for their Interactive Data Language software

X. QIU, J. W. THOMPSON, AND S. J. L. BILLINGE "PDFgetX2: A GUI driven program to obtain the pair distribution function from X-ray powder diffraction data" , *J. Appl. Cryst.* (2004), **37**, 678–678

1.4 Using this manual

Please note that this manual describes only the use of the program. For a detailed description of the implemented data corrections, please see the book by Egami and Billinge [5].

1.5 Acknowledgments

The authors are in great gratitude to Dr. Jeong for the generous sharing of his literature collection of X-ray data analysis methods, also for his valuable inputs. Great appreciations go to Dr. Proffen for his lateral helps with early coding consulting, latex template for this manual, and the windows self-installer program. We would to thanks all the pre-beta version users for their patience dealing with the malicious bugs in the program, and their kindness to report them in details. An incomplete list of names includes Mr. Masadeh, Ms. Kim, Mr. Malliakas, Mr. Mtshatawani. Dr. Petkov and Dr. Bozin should also be thanked for the useful discussions.

Some open source utilities written by David Fanning, and Craig Markwardt have been used in this program without modifications. Several X-ray scattering tabulated data files are taken from the DABAX data base from ESRF website (http://ftp.esrf.fr/pub/scisoft/xop/DabaxFiles/), as guided by the GNU general public library license.

Financial support from the National Science Foundation through the grants DMR-0304391 and CHE-0211029 and from U.S. Department of Energy through grant DE-FG02-97ER45651, as well as the Center for Fundamental Materials Research (CFMR), are gratefully acknowledged.

Chapter 2

Quick Start Tutorial

Please note that the data corrections carried out on the data are not described in this manual, just the use of the program. For a detailed description of the corrections, please see Egami and Billinge [5].

2.1 Installation and start-up

Detailed information about where and how to obtain the program PDFgetX2 can be found in Appendix A. Before you start, it might be a good idea to check for updates and/or bug fixes. Ways to invoke PDFgetX2 on for different platforms are also explained in Appendix A.

Now let's start the program, and the welcome window should appear as shown in Fig. 2.1. The look and feel of the GUIs varies slightly on different platforms.

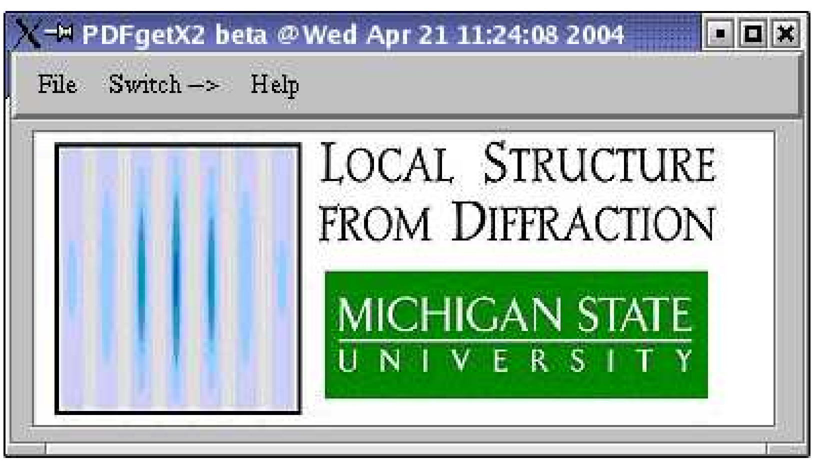


Figure 2.1: The welcome interface when starting PDFgetX2

How to use the tutorial

Step by step instructions will be given with the tutorial data. Some detailed explanations are added as *slanted text* with an offset. They are for further reading only, thus can be safely skipped. The directory and file names are typed as *italicized*. The **bold texts** refer to the widget names on the GUI.

2.2 Copy tutorial data

The data files used in this short tutorial can be found in the *tutorial* subdirectory under the PDFgetX2 installation directory (e.g. C:\Program Files\PDFgetX2 in Windows). You need to copy three files, *ingaas_tutorial.spec*, *ingaas_tutorial.int*, and *ingaas_tutorial.bkg* to a local directory of your choice. The *tutorial* directory also contains the data files we plan to obtain in this chapter for your reference only. For all supported data file formats, please refer to Appendix B.

2.2.1 Description of the tutorial experiment

The tutorial sample, semiconductor alloy $\mathrm{In}_{0.33}\mathrm{Ga}_{0.67}\mathrm{As}$, was measured at 10 K with flat-plate transmission geometry at the Cornell High Energy Synchrotron Source (CHESS). The x-ray energy used was 60.0 keV ($\lambda = 0.2067$ Å) selected with a Si(111) double-bounce monochromator. The scattering from the empty flat plate container was also measured, and found to be rather low. All data were collected using program SPEC^{TM}, a UNIX-based software package for instrument control and data acquisition.

2.3 Preprocess SPEC format data

2.3.1 Objective

The objective is to obtain usable (i.e. properly corrected) data for the general purpose PDF data processing in section 2.4 **Get X-ray PDF**.

¶ *Do I need to preprocess my data?*

This step, preprocess the data, is necessary if any one of the following conditions is met.

- *more than one scan exists, and merging is required (see Appendix B)*
- *the detector counts need to be corrected for dead time effects*
- *the data need to be normalized point by point, with the normalization factor stored in one column*
- *the data have been properly preprocessed, but don't have the estimated standard uncertainties.*

If none of the above describes your situation, you probably don't need to go through this step, and please jump to section 2.4. If preprocessing is required, the input file needs to be in SPEC format. ASCII files can be easily converted (see Appendix B).

The justification for this data preprocessing is the following. A SPEC data file usually contains more than one scans with each scan covering portions of the measured reciprocal space. Every scan is a n row by m column data set. One row represents a measured data point, and the columns record the $Q/2\theta$, detector counts, monitor counts, collection time, and some other values. However, what's of our interest is the scattering intensity versus $Q/2\theta$ over the entire measured range. This simply means to merge the scans together and extract the useful columns. However, different scans often have different scales due to either varying collection time or incident beam intensity changes. Even within one scan, each point may also require proper normalization due to the small fluctuations of the incident beam. Often more significantly, the non-linear response (counting rate dependence) of detectors needs to be corrected on the raw counts first. Thus each scan needs to be treated individually before being merged together.

2.3.2 Load SPEC data file

First, we need to load the SPEC file. The tutorial raw data from the sample were saved into the SPEC format file, *ingaas_tutorial.spec*. A simple but sufficient description of the SPEC format can be found in Appendix B.1. In principle, all supported file formats (Appendix B) can be used, but are all converted into SPEC format internally. If your file is not of one of those formats, a format conversion is necessary.

Following the menu **Switch–>** to **Preprocess SPEC format data** (shown in Fig. 2.2), you will activate the main preprocess SPEC format data GUI as in Fig. 2.3.

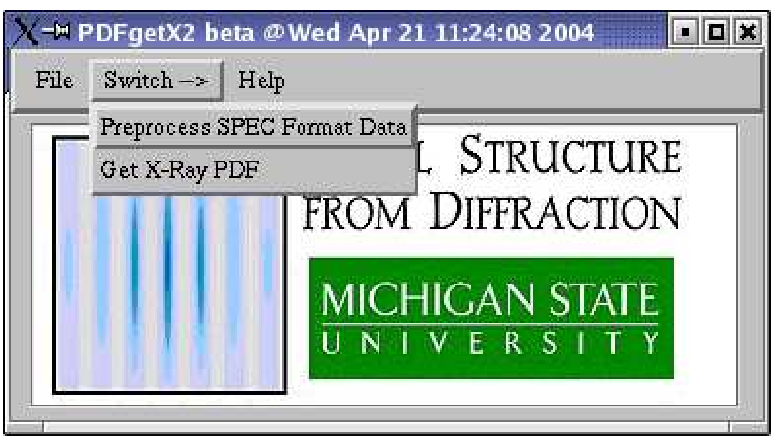


Figure 2.2: Click the **Preprocess SPEC Format Data** button to switch to the GUI in Fig. 2.3

Clicking on the button **SPEC Data File** (Fig. 2.3) will bring up a graphic dialog to select the desired file. You may need to go through directories to locate the tutorial file *ingaas_tutorial.spec*. Either double

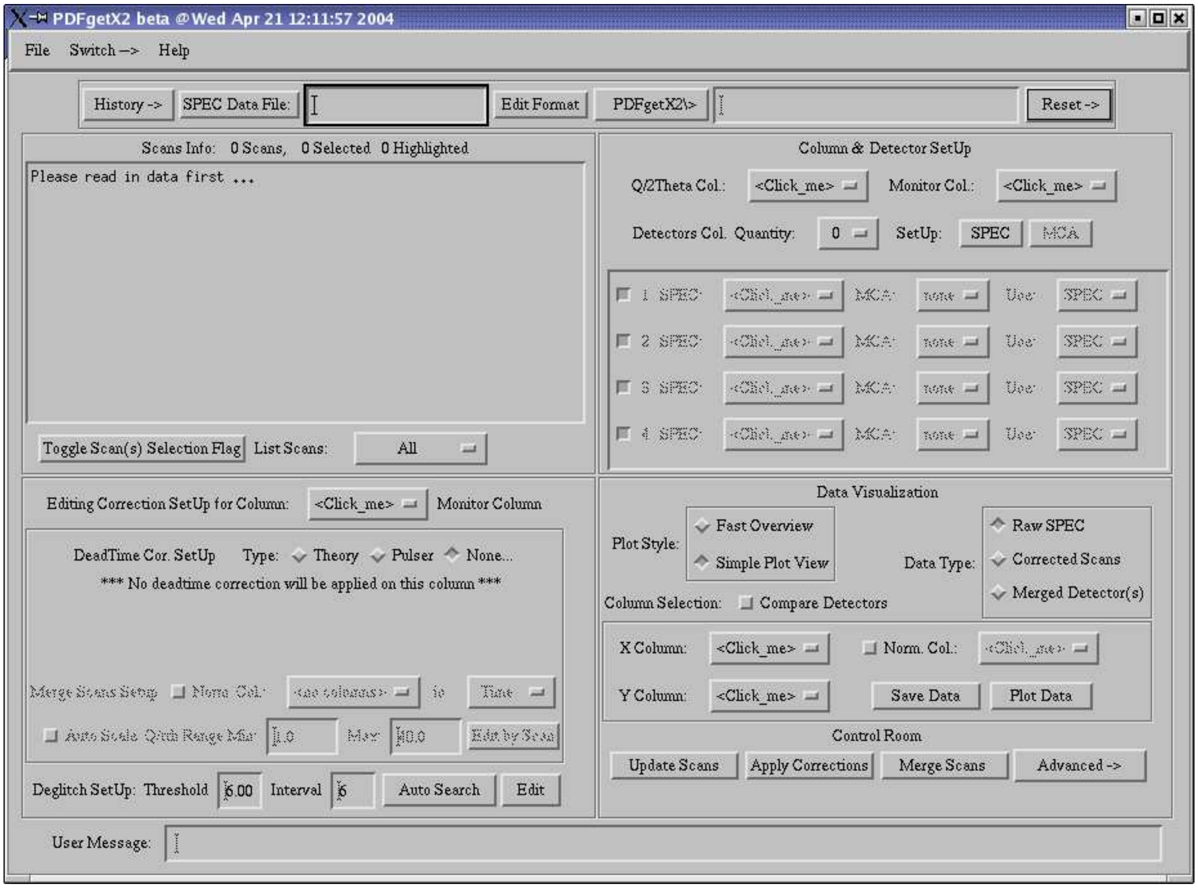


Figure 2.3: The main preprocess SPEC format data GUI

click on *ingaas_tutorial.spec* or press the **OK** button with *ingaas_tutorial.spec* highlighted will make the selection. After a successful data reading (waiting time depends on computer speed), the top left portion of the GUI should resemble Fig. 2.4 (more details on each GUI section come later).

2.3.3 Overview of scans information

In general a SPEC file contains multiple scans. The program reads the scan header line beginning with #S in the SPEC file (Appendix B and reproduces them in the GUI window (Fig. 2.4) allowing you to select the scans you want. Fig. 2.4 shows the GUI portion displaying the scans information. The label on the top tells us that this SPEC file contains five scans. The following **5 Selected** shows the total number of scans selected; **5 Highlighted** shows the total number of scans being highlighted. Each line/item in this list box represents one scan in the data file. Scan information such as the scan number, number of data points, value range of the column of interest (read on for details), and the selection flag are shown as text contents. We can see that all scans are selected (you can highlight all of them using the mouse), which explains the **5**

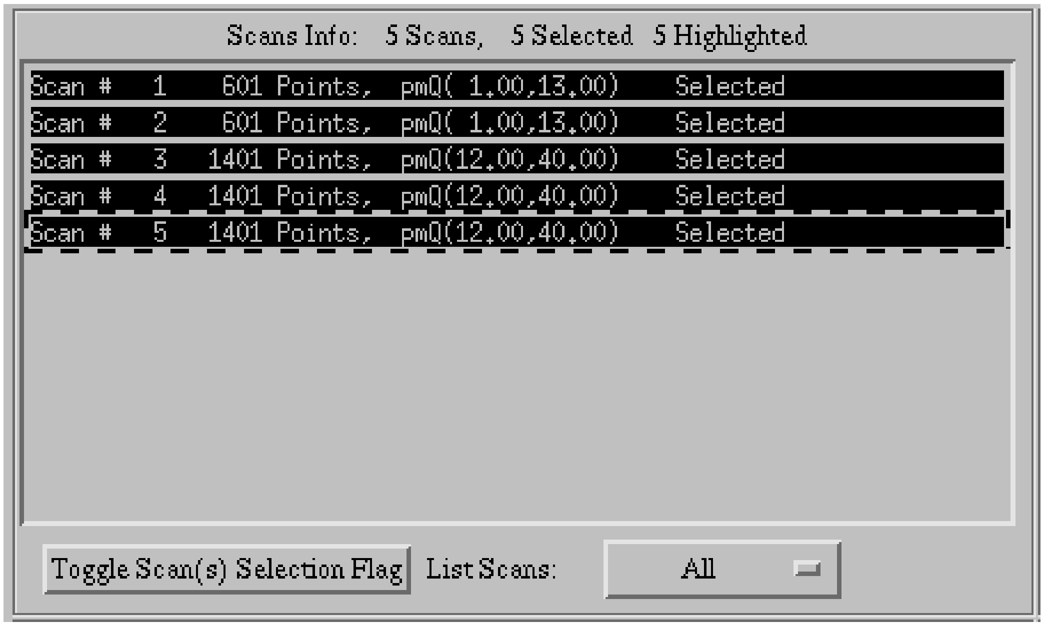


Figure 2.4: The GUI portion displaying scans information in the SPEC data.

Selected, 5 Highlighted in the label above the list.

On the last row, the button **Toggle Scan(s) Selection Flag** will toggle the selection flag of currently highlighted scans. When a scan is flagged as **Selected**, it will be used when merging scans (which happens at a much later stage). If **unselected**, it will not be included during merging. You also have the choice to show only the **selected** or **all** scans in the list through the droplist list next to the label **List Scans**. This only becomes necessary when you have many "bad" scans in the data.

2.3.4 Set up SPEC columns

The next step is to tell the program several column names of interest. This is done in the upper right side of the GUI as shown in Fig. 2.5. You need to specify the column names for $Q/2\theta$, monitor column, and detector column(s). Only the detector can have more than one columns, in the case that multiple detectors are used during measurements. The allowed maximum number of detectors is 10. For each detector, you need to specify its column name in the SPEC file

In our data, each scan has 13 columns, whose names are listed in the droplist for each column selection. $Q/2\theta$ column should be set to **pmQ**, while the monitor column set to **IC2**. Since only one detector was used during the experiment, the number of detectors is 1, whose SPEC column should be **ELASTIC**. In the end, the GUI should look like Fig. 2.5.

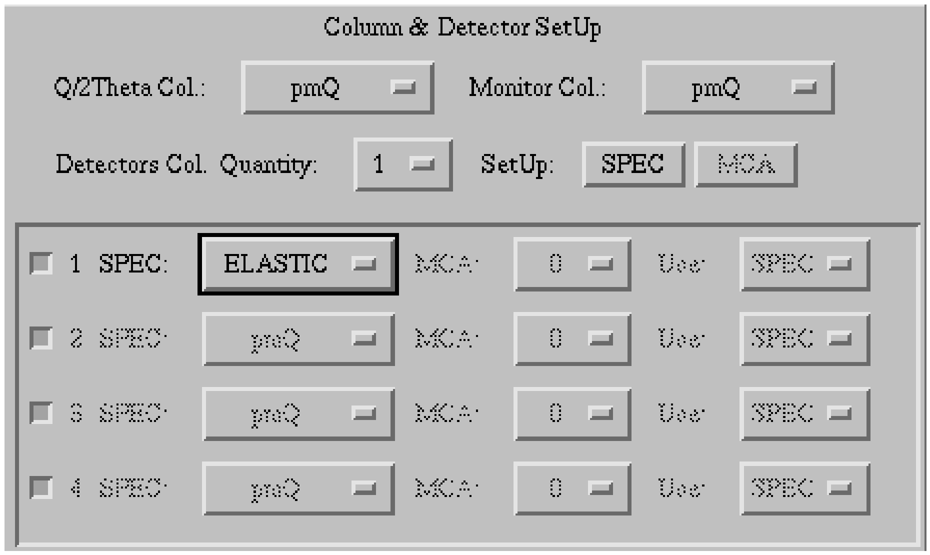


Figure 2.5: The GUI section to set up SPEC columns

2.3.5 Set up dead time correction

Once SPEC column names are set, we are ready to set up the dead time correction on the *monitor* and *detector* column(s). They are **IC2** and **ELASTIC** columns, respectively. The GUI for this job is located at the lower left panel, as shown in Fig. 2.6.

> ¶ *Is dead time correction necessary?*
>
> *This is determined by whether the count rate is within the linear response regime of the detector. Dead time correction is rarely needed for in-house X-ray measurements, while becomes necessary when data are collected at high flux synchrotron sources. If you are not sure, please check the specifications of the detector against your maximum count rate (i.e. the strongest Bragg peak). For more details on why and how detector dead time correction is applied, see the book by Egami and Billinge [5].*

From the GUI, the dead time correction setup for ONLY ONE column is directly displayed and editable. Switching between columns (the monitor and detector columns) is done via the **Editing Correction SetUp for Column** droplist on the top portion of this GUI section. Once a selection is made, the GUI below will be updated by selected column settings. Our monitor counts are from ion chambers, therefore doesn't need any dead time correction, so please select **None**. For the detector column **ELASTIC** coming from the solid state detector (SSD), you can either choose to use theory or pulser method (see Chapter 5 of [5]). If you use the theory method, you will need the the total counts in the detector and the measuring time and the detector dead time. Select the data column **TOTAL** and **Seconds** for the first two. For this Ge detector, a dead time of 11.6 μs works well. The pulser method refers to the use of an electronic pulser ([5]) that is propagated with the data. You will need to select the column of data containing the pulser counts "**PULSER**". Either the theory or pulser method will work well for our data, however the theory method is recommended because it introduces less noise. If you are interested, try them both for comparison. The GUI is shown in Fig. 2.6.

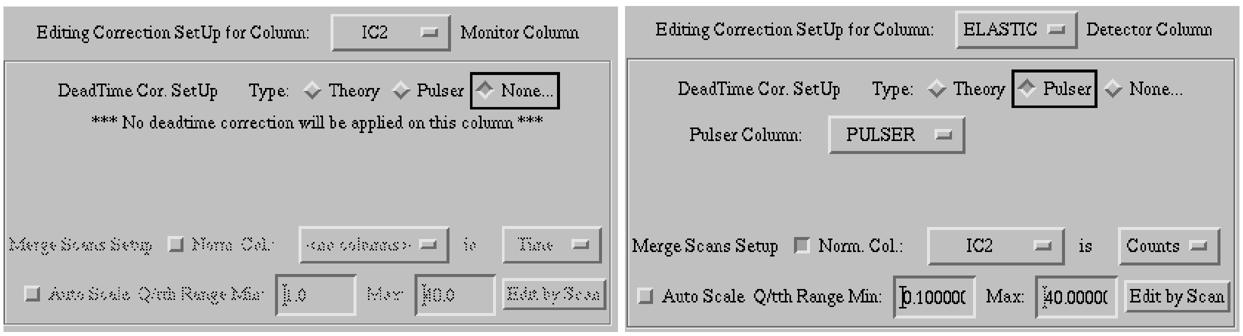


Figure 2.6: The GUI to set up dead time corrections for monitor (left) and detector columns (right)

2.3.6 Set up detector(s) merging

In general, when data are collected in more than one scan, the different scans need to be normalized by their respective monitor counts or measurement time before being merged. Therefore, to set up the program for merging scans, we must assign the data column for normalization, also the range of each scan, and scaling between scans We should set lower and upper cutoff values for each scan to exclude the experimental artifacts sometimes appearing at either end of the data. Be aware that you need to set up EVERY detector column in case of multiple detectors. Also for this reason, the GUI to set up each detector is accessed through the **Editing Correction SetUp for Column** droplist used in setting up the detector dead time corrections. The section for this job is in the lower section of Fig. 2.6.

In our case, we set the normalization column to **IC2**, which is the monitor counts, and opt not to use **auto scale**. **Qmin** should be 0.0 Å^{-1}, with **Qmax** of 40.0 Å^{-1}. The right panel of Fig. 2.6 shows the settings for detector column **ELASTIC**. Here the range of from 0.0 to 40.0 Å^{-1}is for the merged data. Setting individually used range for each scan is prompted by clicking on the **Edit by Scans** button.

2.3.7 Actions and visualization

Now, we have given all the needed information to prepare the raw SPEC data. We can turn to our attention to the GUI section of action and visualization as shown in Fig. 2.7.

The buttons to issue calculation commands are in the bottom row. Button **Update Scans** is only useful when MCA data are available (under development), though it doesn't hurt to click it. Clicking button **Apply Corrections** will apply the corrections (the detector dead time correction, normalization, etc) to monitor and detector columns as you specified. Clicking button **Merge Scans** will take all selected scans (flagged as "selected") and merge them into the single merged data, which is what we set out to obtain in the first

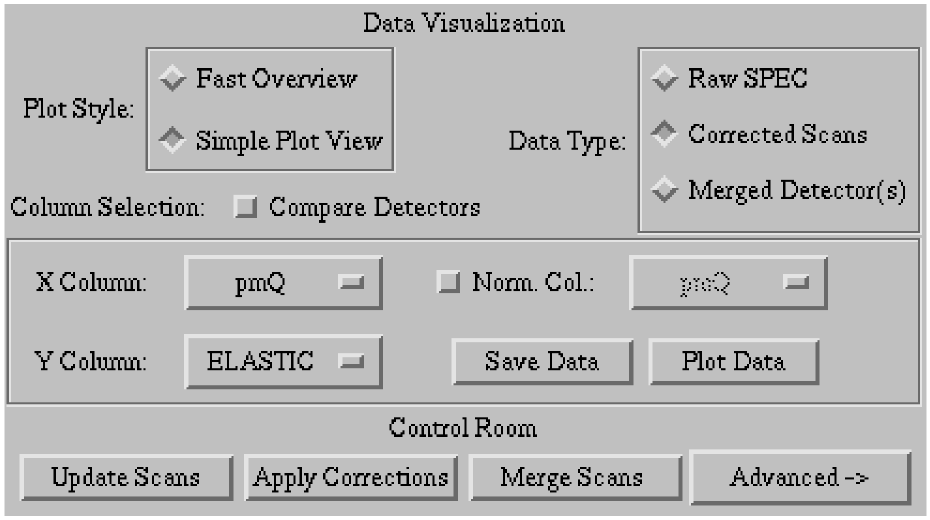


Figure 2.7: The GUI section to plot/save data and process data.

place. When multiple detectors are used, the merged data for each detector is stored as one scan in the total merged data in SPEC format. The merged data now has more than one scans, however conceptually, this is different from the usually meaning of SPEC data.

In the end, we have data at three stages, raw SPEC data, scans data after correction, and the merged detectors. We can either save or plot them after making our selections. If the **Save Data** button is clicked, you will be asked for a file name to save the selected data. If the **Plot Data** button is pressed, the same selected data will be plotted in a separate window, with additional settings described below. The **Plot Style** should be practically set the **Simple Plot View** at all times.

Data selection is made as the following. First, choose from the **Raw SPEC**, **Corrected Scans**, and **Merged Detectors** under button group **Data Type**. When either **Raw SPEC** or **Corrected Scans** is selected, the targeted data are those scans currently being *highlighted*. If the button **Compare Detectors** is selected, all the detector columns for the *first highlighted* scan are selected. In all cases, you have the freedom to select the x values to plot via the droplist **X Column**. Unless **Compare Detectors** is selected, you can also choose the y data to plot via droplist **Y Column**. In addition, if you prefer to have a normalization column, you can do so by enabling normalization column (by setting the **Norm. Col.** button). In this case, the plotted **Y Column** will be divided by the normalization column. If this sounds rather confusing, please play with the interface for better understanding.

After carrying out the actions (**Update Scans**, **Apply Corrections**, **Merge Scans**, we need to check the correction results. We should not select **Compare Detectors** since only one detector is present in the tutorial data. Set the **X Column** as **pmQ**, the **Y Column** as **ELASTIC**, the **Norm. Col.** as unselected, and highlight all the scans. To visualize the raw SPEC data, select **Raw SPEC** as the **Data Type**, then click on the **Plot Data** button. The plot window appear as Fig. 2.8 left panel. If the **Corrected Scans** is selected, we should see Fig. 2.8 right panel. All scans match quite nicely after the corrections and normalization.

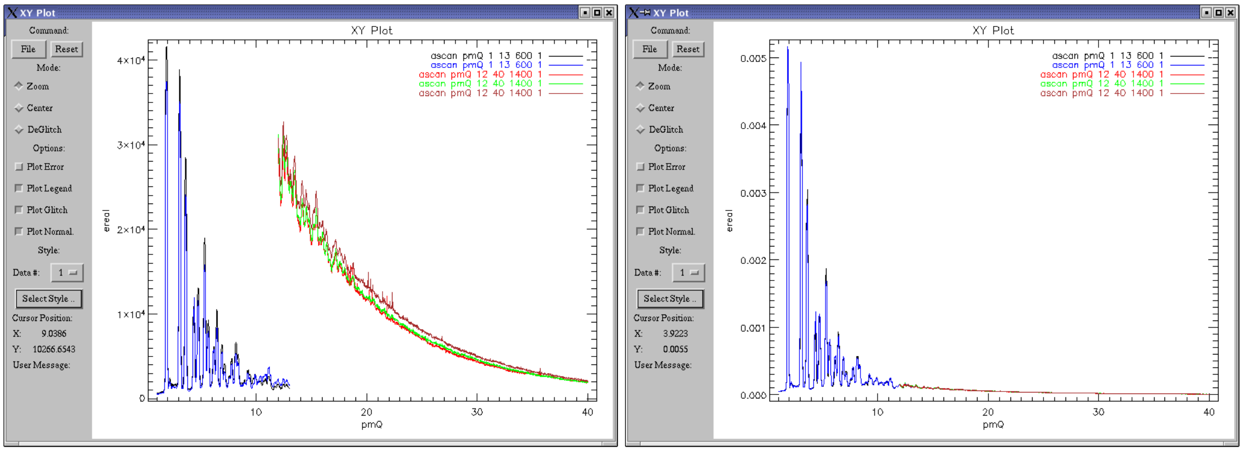


Figure 2.8: Visualization of the raw SPEC data (left) and the corrected SPEC data (right)

2.3.8 Save merged data and the history

Following the above section, if **Merged Data** is selected as the **Data Type**, we can visualize the merged data as in Fig. 2.9. Clicking on **Save Data** will invoke a dialog to select a file to save the merged data. You can use your favorite plot program to compare what you have just got with the provided reference file *ingaas_tutorial.int* (under the tutorial directory where we copied the tutorial data).

The merged data have four columns. The first is $Q/2\theta$ column; the second is the detector count column; the fourth column contains the error (standard deviation) of the detector count, i.e. the second column. The third column is present for compatibility reasons with all values set to zero. The detector count error is first estimated as the square root of the detector count, then propagated through the corrections and normalization steps. This format is also compatible (and the default format) with the next step **Get X-ray PDF** described in next section.

All the settings can be saved into a history file. Pushing the **History ->** and then the **Save ...** button will prompt to select the file to save the history information. The saved history file can be loaded back into the program by clicking on the **Load ...** button instead. This tutorial also contains an example history file *ingaas_specred.xhst*. Loading this file will set up what we have gone through in this section.

2.4 Get X-ray PDF

2.4.1 Objective

Our goal for this step is to obtain the structure function $S(Q)$, then the PDF, $G(r)$. Though only the elastic scattering intensities are what we are interested in, the collected raw scattering intensities inevitably are subject to various instrument and sample effects. The background should be subtracted. Standard corrections due to partial X-ray polarization, sample absorption, and double scattering need to be performed.

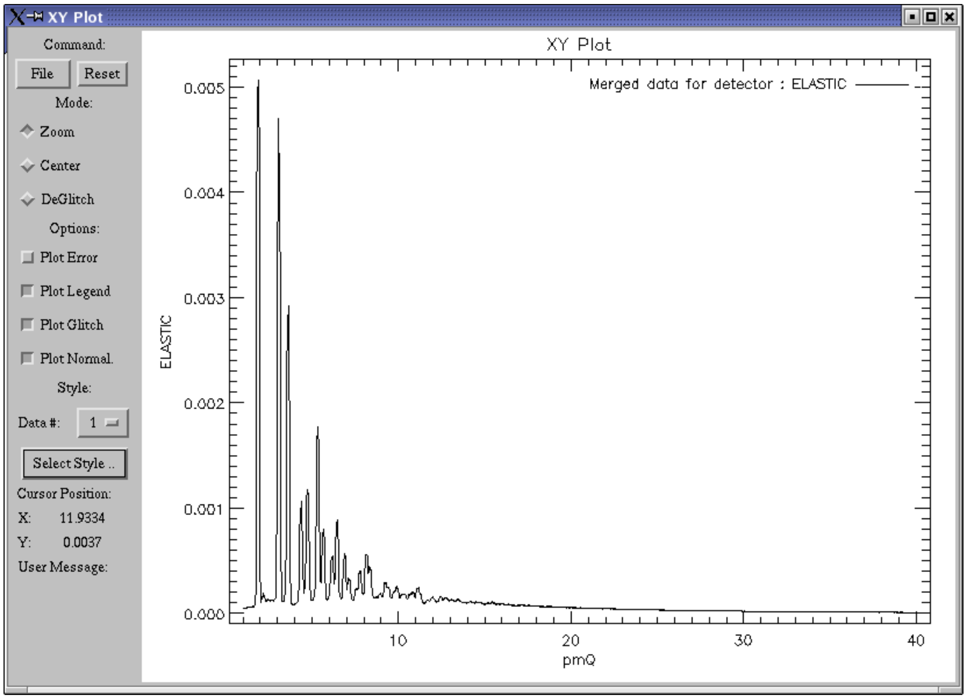


Figure 2.9: Visualization of the merged data obtained from preprocess SPEC format data

The unwanted Compton intensities should also be removed. This section will show you how these are done in this program.

2.4.2 Setup working directory

Following the menu **Switch–>** to **Get X-ray PDF** as Fig. 2.10, you will activate the get X-ray PDF GUI as in Fig. 2.11.

Changing working directory is done by clicking on the button **Working Directory** in the top portion of Fig. 2.11. This would bring up a directory selection window. We should select the local directory containing all the copied tutorial data files (see section 2.2).

The **Working Directory** displayed on the GUI (Fig. 2.11) can be very useful. First, when selecting data files, the directory to start the search will always be the **Working Directory**. More importantly, the program by default saves the $S(Q)$ and $G(r)$ files automatically under the **Working Directory**. A log file (*.pdfgetx2.log*) and temporary history file (*.pdfgetx2.xhst*) are also saved under the **Working Directory** when exiting the program. A recommended file organization is to have all the data files under one directory, then set this directory to be the **Working Directory**. This will also make data file selection easier, and all the autosaved files will also be in the same directory. The default working directory is where you start up the program PDFgetX2. In Linux/UNIX, it's where you run the *pdfgetx2* command; in Windows, double click on a **.xhst* will set its directory as the **Working Directory**.

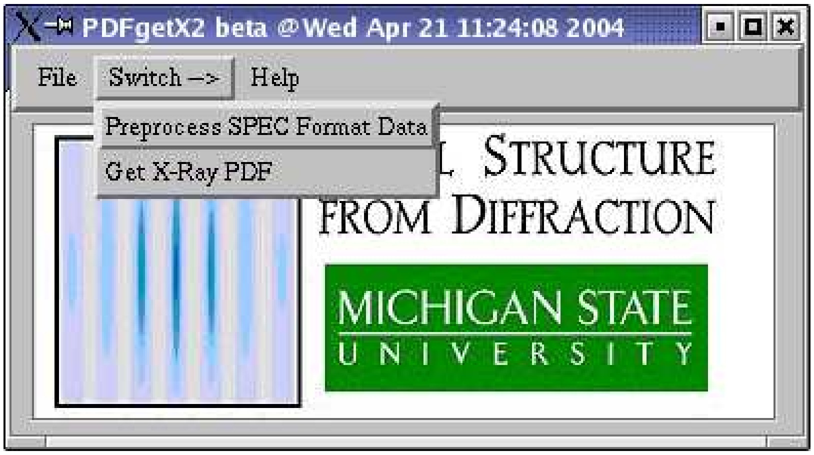


Figure 2.10: Click the **Get X-ray PDF** button to switch to the GUI in Fig. 2.11

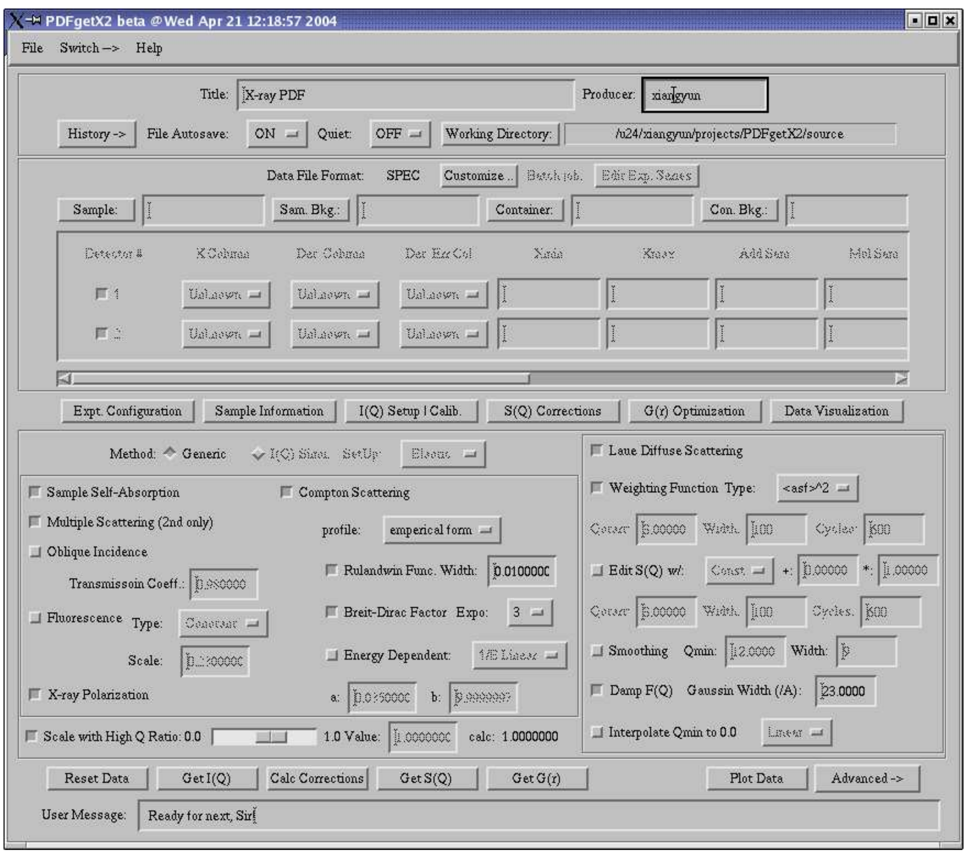


Figure 2.11: The main Get X-ray PDF GUI

2.4.3 Input data files

In this tutorial, we will use the provided file *ingaas_tutorial.int* as the sample data, *ingaas_tutorial.bkg* as the sample background data. Certainly you can also use the merged data obtained in section 2.3 as the sample data. Container and container background data are not used in the tutorial. Input of each data file is done by clicking on the corresponding button (as in Fig. 2.12), which will invoke a GUI to select files. Data files are not immediately read in after being selected. For your reference, the corresponding GUI section should look similar to Fig. 2.12.

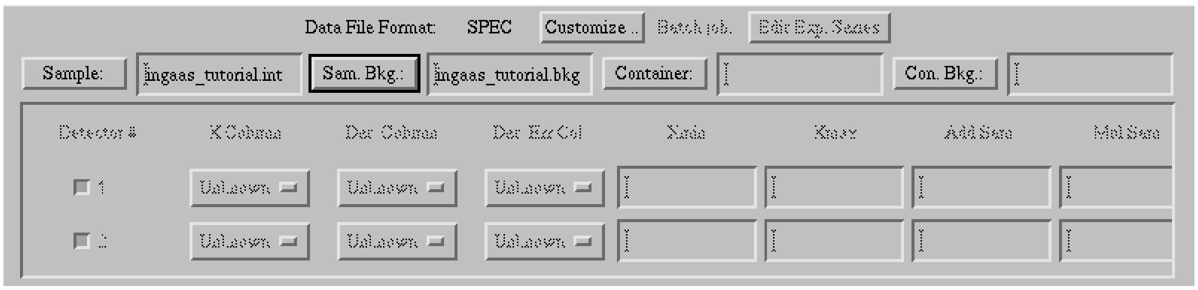


Figure 2.12: Read in the sample and sample background files.

¶ *What are those data types?*

Four kinds of data are used. Sample data refers to the scattering intensity when sample is measured, which means to contain contributions from sample, container (if used), and background scattering. Sample background data refers to the background scattering when sample (and container if used) is removed from the beam. Container data refers to the scattering when only container is positioned in the beam. The container background refers to the background when container is removed from the beam (this is usually the same as the sample background). Only sample data is mandatory to proceed the analysis, though background data is always preferred to have. The classification of data types shouldn't be restricted to their names. The essential differences between them are the types of corrections applied (see Egami and Billinge [5]).

2.4.4 Experiment setup

GUI section of experimental configurations is accessed by clicking the **Expt. Configuration** button (among the button group in the middle of the GUI). Similarly, you can switch to other GUI sections by clicking on the corresponding buttons. In our case, we only need to input the X-ray wavelength and polarization factor. No filter was used during the experiment. The absorption coefficient of the container is only necessary in case of strongly absorbing containers.

For our example data, the X-ray energy of 60.0 keV was used, giving the wavelength of 0.2067 Å (type is **Customize ...**). The data were collected at a synchrotron source (CHESS) where beam is highly linearly polarized. The polarization factor is around 0.94 (type is **Customize ...**) in this case. You would have something like in Fig. 2.13.

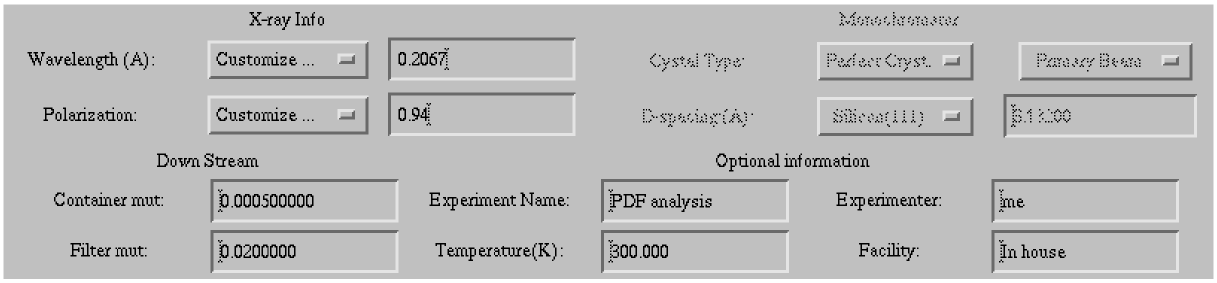


Figure 2.13: The experiment configuration

2.4.5 Sample information

Access of the sample information GUI is via clicking the **Sample Information** button. Then, clicking the **Add/Remove Atoms** button, will present you an interface mimicking the periodic table to specify the chemical species in your material. After the selection of atoms, it may take some time for the GUI to create text boxes to enter the fraction, valence, etc. of each atom. Only the atom fractions are required. The tabulated X-ray scattering data base used includes some common ions, so entering the valence information may help. Other fields such as f1 and f2 are the dispersion corrections to the scattering factor that become important in resonant scattering experiments. Sample measurement geometry is required, as well as the measured sample attenuation coefficient μt.

Fig. 2.14 shows the series of GUIs you should get and the values you need to enter. Atoms In, Ga, and As are first chosen, then the fractions are set to be 1, 2, and 3 respectively. Sample geometry is selected to be **Flat Plate Transmission**, and the value of sample attenuation coefficient μt, 1.11.

At the bottom are some optional parameters about the sample, from which the theoretical μt value can be automatically calculated. This can serve as a consistency check against the measured μt value. Also the number density here is used to compute the baseline for the $G(r)$, which also serves as an preliminary check of the PDF scale. Thus it's advised to fill out these information properly, though these numbers are not used in the corrections explicitly.

2.4.6 Set up I(Q) corrections

The only thing we need to specify here is the **X Column Format**, which is **Q(Å^{-1})** in our case. The filter correction doesn't apply to our data. The effective scattering volume correction is not significant due to the rather low background level. All other default values are fine. The setup GUI on the right are still under development.

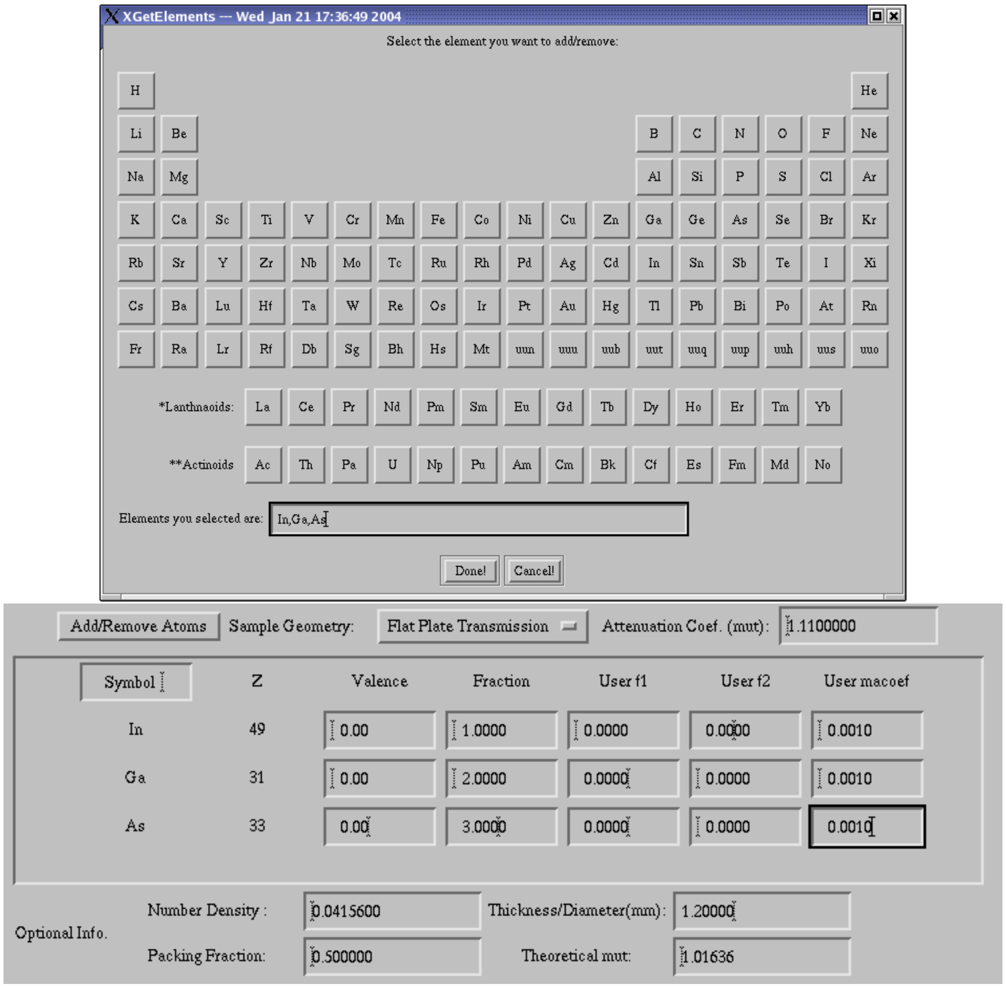


Figure 2.14: The GUI to input sample chemical information

2.4.7 Set up $\mathrm{S}(Q)$ corrections

This GUI section is shown by clicking on the **S(Q) Correction** button. Too many available corrections here are somewhat misleading because only a few are necessary in most cases. We should always set the method to be **Generic**, as the **I(Q) Simu** is not functioning yet. For this tutorial data, corrections to be applied are **Sample Self-Absorption**, **Multiple Scattering**, and **Compton Scattering**.

In addition, we need to apply the **Laue Diffuse Scattering** and **Weighting** correction. The weighting function type should be **<asf>**^{2}. Other corrections are mostly cosmetic, only needed in special situations. At this point in time, the setup is done! For the technical details, please refer to the book by Egami and Billinge [5]. At last, check whether what you have is the same as in Fig. 2.4.7.

¶ *What corrections should be applied to Compton scattering?*

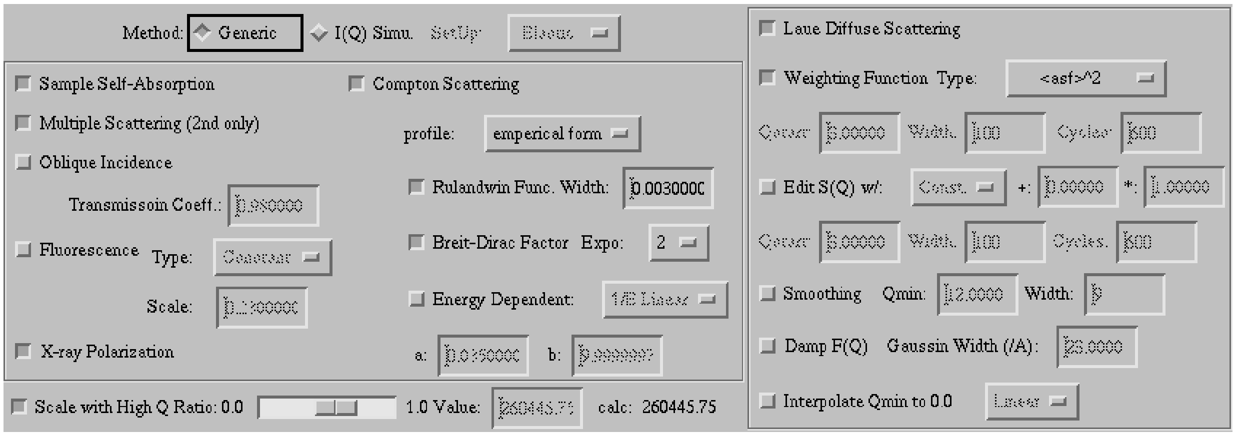


Figure 2.15: The GUI to set up $S(Q)$ corrections

Theoretical Compton intensities sometimes require additional corrections depending on the experimental setup and the detector characteristics. The Ruland window function [11] should be multiplied when the measured total intensities come from the integration within a finite energy window. When the energy window is properly set, as Q increases, the Compton scattering gradually falls out of the energy window, giving cleaner signals in high Q region. The Ruland window function is an empirical analytical formula to take account of this energy window effect, with only one parameter. When the data extend to high Q values (say, $\geq$ 25.0 Å^{-1}), the Breit-Dirac recoil factor becomes significant. The exponent should two if the number of photons is measured; three when X-ray intensity is measured [12]. The energy dependence correction applies when the detector counts depend on the X-ray energy, for example, the image plate area detectors. Two empirical forms are available: linear or quadratic with the X-ray energy.

Coming to our experiment, a single-channel pulse-height analyzer (SCA) was used to set the energy discrimination window. Therefore, the Ruland window function needs to be applied. It was found the window width value of 0.003 works fine. The number of photons was counted by a solid state detector, giving the exponent of the Breit-Dirac factor of two. There shall be no energy dependence correction to the solid state detector.

2.4.8 Set up G(r) transformation

Only a direct sine Fourier transform is performed when going from $S(Q)$ to $G(r)$. Thus, the only setup for $G(r)$ is the Fourier transform range of $S(Q)$, and the calculated r range of $G(r)$. Simply go to the GUI section by clicking on the **G(r) Optimization**, fill up the numbers as in Fig. 2.4.8. You may be distracted by other widgets in the same section. They are mostly for batch processing testings when you have a large number of data set. The batch processing is not user friendly yet, and still under testing. They may also be taken out of the GUI.

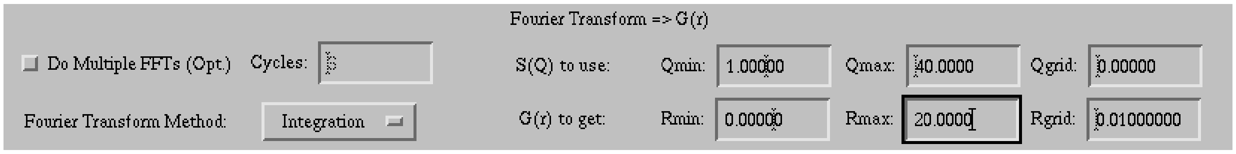


Figure 2.16: the GUI for $G(r)$ Fourier transform settings.

2.4.9 Process X-ray data

Once you've filled in the required information, it's time for data processing, which is controlled by the button group above the **User Message** display (refer to Fig. 2.18). There are totally seven buttons. **Reset Data** will invalidate the read-in data, thus forces the program to (re-)read the data files when **Get I(Q)** button is clicked next time. When **Get I(Q)** button is pressed, jobs to do are, 1) read the data files if they haven't been read in yet or **Reset Data** has been pressed since the last **Get I(Q)** action; 2) apply the additive and/or multiplicative corrections to all data; 3) subtract the background, container, scattering from the sample if applicable; 4) 2θ to Q conversion if applicable, and finally obtain $I(Q)$ with the specified **Xmin** and **Xmax**.

Set up experimental data

After the data files are read in (by pressing the **Get I(Q)** button), we may need to set up the experimental data if the data files are not of the default format (four columns, $Q/2\theta$, $I(Q)$, δQ(not used), $\delta I(Q)$). For data files of all supported formats, if the number of columns is less than four, extra dummy columns with all values of 1.0 will be added to have totally four columns. All columns are still identified by their names. If the data files don't have column names, such as the $n \times m$ ASCII format, default column names will be given based on their column numbers. Provided that all data file readings are successful, all available column names will be shown in the droplists in Fig. 2.4.9.

Figure 2.17: the GUI for setting up experimental data

Three columns are required to be specified, the X ($Q/2\theta$), detector count, and the error (standard deviation) of detector count. If the detector count error column is not present in the data, just specify **None...**. In this case, the square root of the detector count will be used as the error. Error propagation is carried out at all steps all the way to $S(Q)$ and $G(r)$. Here we also have the freedom to do some simple additive and multiplicative operations on all four data types. The function of each widget should be self-explanatory.

The program will do a consistency check on some most apparent parameters, and report the exceptions. You have to correct those parameters to be able to proceed. However, the parameter check is no where near complete. For any other situation, an error message will be shown. Please report this error message to the authors with detailed description how to reproduce this error message, for example, attaching the history file and related data files. Also please check the **User Message** and the program output on the terminal for any hint about the possible error sources. Invoking **Calc Corrections** would do several early stage calculations, including 1) absorption correction calculation; 2) multiple scattering calculation, 3) polarization correction calculation, 4) calculation of theoretical elastic scattering and Compton scattering intensities. **Get S(Q)** button will take the $I(Q)$ data, apply the specified corrections to get $S(Q)$ and $F(Q) = Q(S(Q) - 1)$. The sine Fourier transform is performed to compute $G(r)$ when the **Get G(r)** button is clicked. After all those buttons have been pressed, we are ready to look at the results.

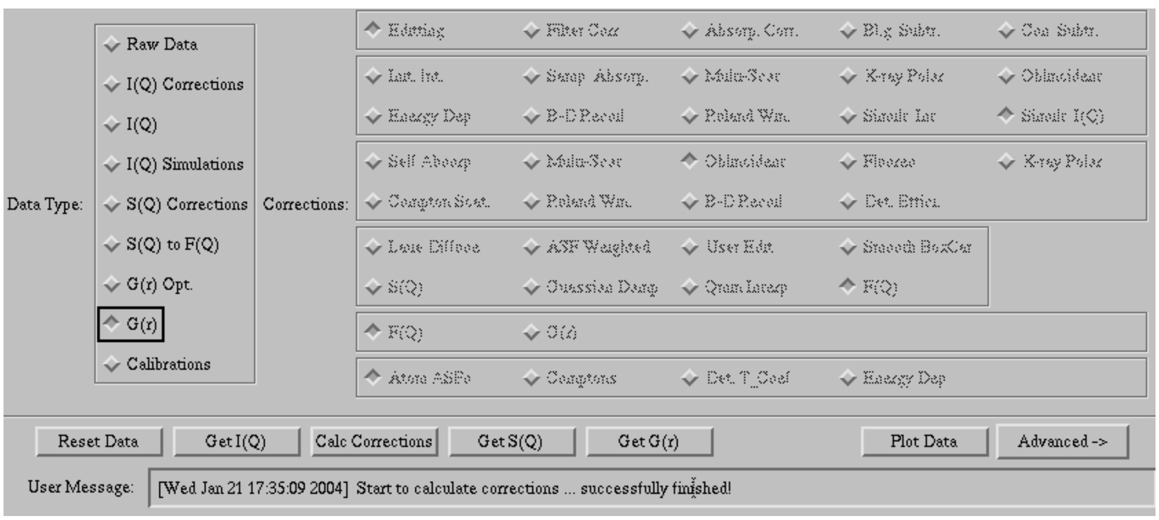


Figure 2.18: From bottom, the **User Message** display, the action buttons, the data visualization selections

2.4.10 Data visualization

The data flow diagram in the program goes from the four data files (sample, sample background, container, container background), to the sample only scattering intensity $I(Q)$, then to the structure function $S(Q)$, finally to the PDF, $G(r)$. Within each step, there are maybe also several corrections applied. In consideration of this, **Data Type** is first selected for the **Data Visualization**. Then, depending on the **Data Type**, some choices may be available to select the more specific corrections. Almost all internal data used in the program can be directly plotted, which can be very useful for problem diagnosis. To give you some examples, the **Raw Data**, **I(Q)**, and **G(r)** represent directly the data as the name suggests. Others refer to correction or calculation steps involving more intermediate data sets. We hope the meaning of each selection is self explanatory enough. One thing to note, all correction terms are calculated regardless of whether they will be applied or not. Thus, you can visualize all correction terms, keeping in mind that it doesn't necessary mean that they have been applied.

Clicking the **Plot Data** button will plot the selected data. The *same* plot window will be reused for subsequent visualizations, thus there is no need to close the window. If your $F(Q)$ and $G(r)$ data look like Fig. 2.19, chances are you have successfully finished this tutorial; congratulations. The data have been published in [13] and [14].

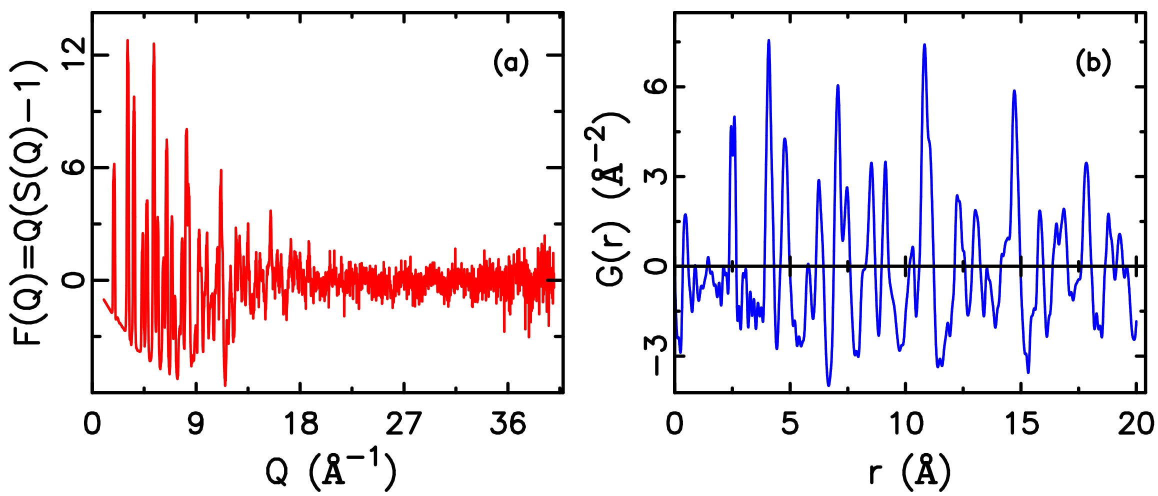


Figure 2.19: The $F(Q)$ (a) and the $G(r)$ (b) for your reference

2.4.11 History files and data saving

The full GUI session can be saved into a history file, and reloading the history file will restore the saved GUI session. History saving and loading is via the menu button **History**, located at the upper right of the GUI. The **Load History...** and **Save History...** from the main menu **File** will do this too. This built redundancy was to make the **Get X-ray PDF** somehow independent. When exiting the program, a temporary history file (*.pdfgetx2.xhst*) is also saved under the **Working Directory**. An example history file *ingaas_getxpdf.xhst* is included under the *tutorial* directory. You can load it up to check whether you have the same setup.

During data processing, the $S(Q)$, $F(Q)$, and $G(r)$ data are automatically saved by default. You can reset this behavior by toggling the **File AutoSave** droplist. The default file names have the same base as the sample data file, with suffixes of *.sq* and *.gr*, respectively. The history information is also saved into the three autosaved data files explicitly in the header. However, it's still advised to save history files because those autosaved files are simply *overwritten* during data processing. If you want to save other data types, the authors have to ask you to use the **Export ASCII** function of the plotting window. This means you first select the data type to save, then plot it, and export it to a file of your choice.

Chapter 3

GUI Widgets Reference

Quite a few tables will be shown to describe the GUI widgets. They all contain three columns. The first column is the name of the widget, which is usually the text displayed on the widget or the label in its front. The second column indicates the widget type. The third column explains the function of the widget. Table 3 explains what each type means in this context.

Name	Description
menu	a container of one group of next level widgets, mouse click usually displays submenus. The descriptions of its submenus are shown right after below, and preceded by $\hookrightarrow$
action	usually a button leading to value changes of data or parameters, or display of another window
label	a text object usually for the sole purpose to display useful information
list	a multiple item user selectable field by mouse
droplist	a multiple item user selectable field, however the selectable items need to be activated by clicking on it first to see available choices
checkbox	a flag to indicate the choice of a double-valued variable
radiobox	a group of exclusive choices, selection of one will deselect all others
text/edit	a text field where user can change the value directly with the keyboard

Table 3.1: Explanation of the widget types

3.1 Preprocess the SPEC format data

3.1.1 The top row

The top row of the GUI in Fig. 2.3 right panel deals directly with files and some less frequently performed actions. The descriptions of the widget functions are listed in Table 3.2.

Widget	Type	Description
History->	menu	show the submenu of load and save history buttons (two below)
↪Load ...	action	display a GUI to select a history file to load
↪Save ...	action	display a GUI to select a history file to save
SPEC Data File	action	display a GUI to select the data file to read
Edit Format	action	customize the format of SPEC data file to be read
PDFgetX2	menu	intended to run saved user commands (under development)
Reset->	menu	Show the submenu to choose what to reset (four below)
↪All Defaults	action	reset the defaults of the three choices below
↪Scans Select	action	reset the scan selection flags to defaults
↪SPEC/MCA Setup	action	reset the SPEC/MCA columns settings to defaults
↪Corrections Setup	action	reset the dead time correction and merging settings for all columns

Table 3.2: Descriptions of top GUI portion of the preprocess SPEC format data

3.1.2 List of scans information in SPEC data

When reading a SPEC data file, the program only looks for the scan starting identifier "#S", the column description identifier "#L", and the following data of each scan. "#S" and "#L" *must* exist for each scan, and are the *only* header information checked by the program. The scan number immediately following the "#S" should be unique throughout the file. The blank space is the default field separator. After you read in the data, all the parameter values will be set to defaults (which are usually wrong, e.g. the first column will be set to be $Q/2\theta$ as the default. The scans list in the GUI will tell you how many points in each scan, and the Q/2θ data range. The last field in the scan list is an indicator of whether that scan will be used during merging. The default is to use all the scans which has more than one point. When you highlight some or all the scans in the list, the column names that are common to all the highlighted scans will be shown as available columns for selection. As you may notice the change of values in Q/2θ, monitor column droplist and others. The label on top of the scan list gives a brief overview of current scan data status. See Table 3.3 for details.

3.1.3 Set up SPEC column assignments

Columns in the scan data are identified by their names after the "#L". The column names shown in the droplists are the *common* column names of the currently *highlighted* scans. The GUI operation should be mostly self-explanatory, with descriptions in Table 3.4. For each detector you can specify whether it'll used during merging, its column in the SPEC data file, and its column in MCA data file (Table 3.5).

Widget	Type	Description
Scans Info	label	displays the numbers of total, selected, and highlighted scans
Scan List	list	displays the summary of each scan as one line. One or multiple scans can be highlighted with help of Control or Shift keys. Highlighted scans are used to find the common columns. All the related column name droplists will be updated. The GUI may respond somewhat slow because many widgets have to be updated
Toggle Selection Flag	action	this will reverse the selection flags for currently *highlighted* scans
List Scans	droplist	choose either to list **All** scans or scans **Selected Only**

Table 3.3: Descriptions of the GUI displaying scans information

Widget	Type	Description
Q/2Theta Col	droplist	specify which column is Q/2θ. Note that its change will update the scan information in the list to the left, e.g. column name and data range. Certainly the column can also be θ or d spacing, not restricted by its name only
Monitor Col	droplist	specify which column is monitor. This monitor should effectively measure the incident beam flux on the sample, i.e. usually the closest upstream monitor. Detector dead time correction can be applied to the monitor column
Detector Quantity	droplist	specify the number of detectors to collect scattering intensities. You will notice the number of editable detectors below will change accordingly. Every detector can also be set up for detector dead time correction. The detector columns will be the columns to merge
SPEC	action	activate the GUI to set up the detector columns in SPEC file
MCA	action	activate the GUI to set up MCA data (under development)

Table 3.4: Descriptions of the GUI to set up SPEC columns

3.1.4 Set up dead time correction

Detector dead time correction can be applied to monitor column and all detector columns. Table 3.1.3 shows the descriptions of the GUI elements,

3.1.5 Set up detector(s) merging

Before merging the scans, you may need to do two things. i) normalization by the monitor counts or measurement time point by point; ii) remove the spurious features sometimes appearing at the ends of scans. Please refer to Table 3.1.4 for usage. One feature still under development is the automatic removal of

Widget	Type	Description
use flag	checkbox	indicate whether this detector is used for merging data or not. One trick can be done here if you want to apply some corrections to non-detector columns: you can simply add one more detector, but set it to be not used so that you won't merge it anyway. However, you can still do corrections to it as a normal detector column
SPEC	droplist	specify the detector's column name in the SPEC data
MCA	droplist	specify which column corresponds to this detector in MCA data. (under development)
Use	droplist	choose whether to use the integrated data from MCA or just the SPEC for this detector (under development)

Table 3.5: Descriptions of the GUI to set columns for each detector

Widget	Type	Description
Active column selection	droplist	the available selections are monitor and detector column(s). The selected column name will be the target to set up the widgets below
Theory	radiobox	choose to use the theory method of dead time correction
Pulser	radiobox	choose to use the pulser method of dead time correction
None	radiobox	choose not to perform any dead time correction
Detector Total	droplist	specify the total counts column for the target
Collection Time	droplist	specify the measuring time column for the target
Dead Time	text	input the measured dead time of the target
Pulser Column	droplist	specify the pulser column for the target

Table 3.6: Descriptions of the GUI to set up dead time correction

the spurious single bin-width noise spikes from the electronics.

3.1.6 Actions and visualization

The commands to process the data corrections are issued from this GUI section. The intermediate and final data can be plotted with some levels of flexibility. Their simple usage is explained in Table 3.1.6. The last row of the GUI is **User Message**, which displays some information during the data processing, as well as warning or error messages.

3.2 Get X-ray PDF

The top frame of the GUI shows some general settings, as explained in Table 3.2. Please refer to Table 3 for the column descriptions in all tables.

Widget	Type	Description
Normalize?	checkbox	choose whether to have a normalization column or not
Norm. Col	droplist	Specify the normalization column for the current detector
is ?	droplist	specify the type of the normalization column, which affects the error propagation
Auto Scale?	checkbox	indicate whether to scale the scans automatically. For autoscaling to work, each used scan must have some $Q/2\theta$ overlap with at least one other used scan. In most case, this should not be necessary. Considerable mismatch between scans after corrections probably signals improper corrections
Q/tth Min	text	set the minimum value of $Q/2\theta$ for the current detector
Q/tth Max	text	set the maximum value of $Q/2\theta$ for the current detector
Edit by Scan	action	This will fire up a GUI to input the Q/2θ range of each scan for current detector, you can also input change the addictive and multiplicative constants for each scan (auto scaling may work better with presence of sufficient overlaps between scans).
Deglitch SetUp	GUI	This part is under development
Threshold	text	to be able to identify glitches, the variation of each point is evaluated and compared to its neighbors. if the ratio is over this threshold value, it's judged to be a glitch (under development)
Interval	text	The variation of each point is only compared to its close neighbors, this width values gives the number of neighbors for comparison (under development)
Auto Search	action	run the auto deglitch algorithm to search for glitches automatically either for correction scan data or the merged data (under development)
Edit	action	Manually edit the glitches (under development)

Table 3.7: Descriptions of the GUI to set up merging detectors

3.2.1 Input data files

You can specify the file format by clicking on the **Customize...** button. Please see Appendix B for all supported file formats. It should be straightforward to load data and background files. One restriction is that the four kinds of file (sample, sample background, container, container background) must have exactly the same format, but can have different number of data points. In the case of multiple scans in SPEC format, each scan is considered as the merged data for one detector. Data from more than one detectors will be blended together to get $I(Q)$. For each detector, some degrees of freedom is provided for both sample and background data. Table 3.2.1 shows the descriptions, noting that they only become editable after data files are read successfully (by clicking on the **Get I(Q)** button).

Widget	Type	Description
Fast Overview	radiobox	in this plot type, each scan is plotted in a single small window, so that you can easily browse through all the scans and make judgments on which scans are "bad"
Simple Plot View	radiobox	this should be used in general. The **XYPLOT** window offers useful features such as zooming, recentering, printing, and exporting to ASCII data or postscript figures
Raw SPEC	radiobox	raw SPEC data will be plotted
Corrected Scans	radiobox	scan data after detector dead time correction and normalization will be plotted
Merged Detector(s)	radiobox	the merged data for detector(s) will be plotted. Note that in this case the X, Y, error columns in the data are already predefined, thus not selectable
Compare Detectors	checkbox	it's only useful in case of more than one detectors. In this case, only the detectors of the first highlighted scan will be plotted
X Column	droplist	the column that will be plotted as X
Y Column	droplist	the column that will be plotted as Y
Norm. Col?	checkbox	choose whether to use a normalization column (for plotting only)
Norm. Col	droplist	the column that will be plotted as normalization column to Y column
Save Data	action	this will save the to be plotted data in a user selected file in SPEC format
Plot Data	action	this will plot the data in a graphic window
Update Scans	action	it only does one thing. Program will check each detector to see whether SPEC or MCA data is selected for use. The active data will be set to the same as SPEC data or reset the value of detector counts to be the result of MCA data integration accordingly
Apply Corrections	action	This will perform dead time correction as set up to monitor and detector counts. Then, perform normalization to each detector counts
Merge Scans	action	this will merge the scans for each detector with use flag on
Advanced ->	menu	(under development)

Table 3.8: Descriptions of the GUI for action and visualization

3.2.2 Experimental setup

In this section, X-ray wavelength and polarization factor can be selected from some predefined values, or entered manually when choosing type as **Customize ....** In case of non-polarized X-ray (in-house source), you will be asked to input the monochromator information, such as the crystal type, where it's positioned, and the d-spacing of the reflection plane used.

Widget	Type	Description
Title	text	choose a title
Producer	text	meant to be the person who analyzes the data
History->	menu	show the sub-menu of load and save history buttons (two below)
$\hookrightarrow$Load ...	action	display a GUI to select a history file to load
$\hookrightarrow$Save ...	action	display a GUI to select a history file to save
File Autosave	droplist	choose whether $S(Q)$ $F(Q)$, and $G(r)$ files should be automatically saved during data processing
Quiet	droplist	choose whether to suppress output information during data processing, this will speed the computing time
Working Directory	action	bring up a GUI to select working directory

Table 3.9: Descriptions of the GUI top frame of Get X-ray PDF

Widget	Type	Description
use flag	checkbox	choose whether to use this detector data or not
X Column	droplist	specify which column is the X column
Det. Column	droplist	specify which column is detector column
Det. Err Col	droplist	specify which column is the standard errors of the detector counts. If not existing, select **none**. In this case, the square root of the detector count will be used the error
Xmin	text	set the lower cutoff value of data X range to use
Xmax	text	set the upper cutoff value of data X range to use
Add Sam	text	an additive constant to sample detector counts
Mul. Sam	text	a multiplicative constant to sample detector counts
...	text	the meaning of other text inputs is similar to Add Sam or Mul. Sam, but applied to different data type

Table 3.10: Descriptions of the GUI to input and edit data files

The absorption coefficient of the container only becomes important when it's large. When a filter is used after the sample (either to suppress the fluorescence or Compton contributions), the absorption of the filter is needed if elastic scattering is also significantly attenuated. Others are optional, mostly for the purpose of keeping a good record.

3.2.3 Sample information

In this section, you need to input the chemical compositions of your sample. First thing you need to do is to specify the elements. See Table 3.2.3 for details.

Widget	Type	Description
Add/Remove Atoms	action	this will bring up a GUI to input the elements with mouse
Sample Geometry	droplist	This is very important for absorption, multiple scattering corrections, the supported geometries are flat plate transmission/reflection, flat plate transmission, and cylindrical capillary
Attenuation Coff.	text	this is a value you should measure on the sample
Number Density	text	input the number density of your material
Packing Fraction	text	usually the value should be around 0.5. This value is used to estimate the theoretical μt.
Thickness/Diameter	text	how thick is the sample, it should be the inner diameter of the glass tube in case of capillary
Theoretical μt	text	program will display the estimated μt here (not editable)

Table 3.11: Descriptions of the GUI to set up sample information

Widget	Type	Description
Sample Self-Absorption	checkbox	should be turned on all the time, this correction takes the sample absorption coefficient and sample geometry as parameters
Multiple Scattering	checkbox	multiple scattering is only calculated up to double scattering, which is often the dominant contribution to multiple scattering intensity. You only need to apply this correction when the sample is quite absorbing.
Oblique incidence	checkbox	this is only applicable to area detectors such as image plates. This correction only becomes significant at high angles and high X-ray energies, where the absorption of image plate is quite small. It takes the X-ray transmission coefficient of the detector as one parameter
Fluorescence	checkbox	in case of photon counting without any energy discrimination, and the X-ray energy is close to one or more atom absorption edge, you need to subtract the fluorescence intensity. Now, only the image plate geometry is supported
X-ray Polarization	checkbox	it should be turned on unless the polarization factor is 1.0.

Table 3.12: Descriptions of the GUI to set up $S(Q)$ corrections (I)

Widget	Type	Description
Compton Scattering	checkbox	Compton scattering correction is one of the most important corrections for proper data crunching. Theoretical Compton scattering profile will be calculated from tabulated values. This should be turned on all the time.
Compton Profile	droplist	choose which theoretical profile to use. This has little effect on the data analysis
Rulandwin Func	checkbox	It's only used when Compton contribution is only present at low Q, e.g. energy discrimination by channel analyzers was used. In this case, the theoretical Compton profile is multiplied by an empirical Ruland window function to emulate the functionality of channel analyzers. This window width parameter defines the width of the window. Therefore, you should give a larger value in case of wide energy window width in your analyzer setup, and vice versa
Breit-Dirac Factor	checkbox	we need to apply this correction especially if you have Compton at high Q. Otherwise the effect is minor. Different mechanisms of photon detection use different exponents, in case of counting the number of photons, the exponent should be 2; in case of measuring X-ray intensities, the exponent should be 3
Energy Dependent	checkbox	this correction basically takes consideration of the energy dependence of quantum efficiency of detectors. It becomes quite important in case of image plates, the best way is to measure it and give it as a user provided curve. At present, either a linear or quadratic empirical formula is used as an approximation
Scaling with High Q	checkbox	scaling is always performed. However, when this knob is turned on, the program will automatically scale the data using the high Q method. Otherwise, user specified value is used.
Laue Diffuse	checkbox	it should be on in general
Weighting Function	checkbox	it should on in general. The function type should be **<ASF>**^{2}
Edit S(Q)	checkbox	this should not be used in most times unless you have sound reasons to do so. It simply applies an additive constant and one multiplicative constant to resulted $S(Q)$. You can turn to this for help when you really can't proceed otherwise
Smoothing	checkbox	a convenient way to smooth very noisy data. The starting Q value and the width of the smoothing boxcar are needed
Damping F(Q)	checkbox	a convenient way to damp the $F(Q)$ with a Gaussian function when data with very limited Q range are processed
Interpolate to 0.0	checkbox	a simple linear interpolation to $Q = 0.0$

Table 3.13: Descriptions of the GUI to set up $S(Q)$ corrections (II)

3.2.4 Set up S(Q) corrections

This GUI section tells the program what corrections to apply, as well as some parameter values. It's not as complicated as it may look like. At this moment, only the **Generic** correction method is implemented. So keep it set all the time. For all other corrections, please see Table 3.2.3 and 3.2.3.

3.2.5 Set up G(r) transformation

Under the GUI section of **G(r) Optimization**, what's of most interest is the range of $S(Q)$ to use and the range of $G(r)$ to generate. It should be straightforward to input those values. In particular, **Qgrid** of 0.0 means all Q values in the data will be used (either regular or irregular).

3.2.6 Data processing and visualization

Data visualization has bee described in details in Section 2.4.10. Data processing is triggered by the row of buttons near the bottom. What they do is explained in Table 3.2.6.

Widget	Type	Description
Reset Data	action	initial data will be invalidated, data files will be read again during next **Get I(Q)** event
Get I(Q)	action	data files will be read in if initial data are invalid. First, the additive and multiplicative constants will be applied on corresponding data types. The, background will be subtracted, as well as the container. Multiple detectors will be blended if applicable. X column is converted into Q. The final data is $I(Q)$, which will be automatically saved if the **File AutoSave** is on
Calculate corrections	action	This is a time saver since most corrections are simply the direct results of the known experimental configurations. You do *not* need to redo this when you change anything in the **S(Q) Corrections** GUI section
Get S(Q)	action	This will take the $I(Q)$ data, then apply corrections based on your setup. Both the structure and reduced structure functions are calculated, and may also be automatically saved if the **File AutoSave** is on
Get G(r)	action	Simply do the Fourier transform of $F(Q)$ to obtain $G(r)$
Save/Plot Data	action	a window will show the data as you specified in the **Data Visualization** session. **Save Data** will save the same data into a file you select

Table 3.14: Descriptions of the GUI to visualize data and process data

The last row of the GUI is **User Message**, which displays some information during the data processing, as well as warning or error messages. When running on Linux/UNIX platforms, the program also writes messages to the terminal.

Appendix A

Installation and Start-up

A.1 Obtain PDFgetX2 and IDL

The recommended way to obtain program PDFgetX2 is to download it directly from its homepage at

http://www.pa.msu.edu/cmp/billinge-group/programs/PDFgetX2/

Recent updates, bug fixes, and other related information are also posted on-line. Please check back from time to time. If the web is not accessible, please send us email (*billinge@pa.msu.edu*) to arrange alternative shipment.

A commercial IDL licensed distribution (version 6.0 or higher), or the freely downloadable IDL Virtual Machine (IDL^{VM}), is the only prerequisite to run PDFgetX2. The IDL^{VM} is available from the download section of the IDL website http://www.rsinc.com/download/. Supported platforms by IDL include Linux/UNIX, WINDOWS, and MACINTOSH. No specific hardware is used by PDFgetX2. In principle, PDFgetX2 can run on any platform where IDL can be installed and widgets are supported.

A.2 How to install

PDFgetX2 takes about 17 MB hard disk space including this manual, while a typical installation of IDL may take up to 200 MB disk space for either the licensed distribution or the virtual machine.

A.2.1 Linux/UNIX system

The file you should download is *pdfgetx2-xxxx-yymmdd-linux.sh*. After a complete download, go to the saved directory, and run command *sh pdfgetx2-xx-yymmdd-linux.sh* to initialize the installation. You will be asked where to put the source and the executable, just be sure that you have the right permission. For example, with my computer running Red Hat Linux 9.0, I got the following when installing the PDFgetX2 into a local directory.

```
[xiangyun@kneesave ~/tmp]$ sh ./pdfgetx2-beta-040416-linux.sh
Verifying archive integrity... All good.
Uncompressing pdfgetx2-beta-040416.......................................
Current directory: /tmp/selfgz31649
Source directory : .
Welcome to installation of PDFgetX2 beta
The liscense statement will be show first, use SPACE to scroll!
Press ENTER to continue ...
PDFgetX2
Software License Agreement


By downloading and installing a copy of the PDFgetX2 Software and
Documentation, you agree to the following terms.

Notification of Copyright: PDFgetX2 is a proprietary product of
Michigan State University ("MSU") and is protected by copyright laws
and international treaty.  You (as "End User") must treat PDFgetX2
like any other copyrighted materials.  Copyright laws prohibit making
copies of the Software for any reason.  You may make copies of the
Documentation for use with a licensed version of the Software;
however, MSU notifications of copyright must be left intact.  If you
have any questions concerning this agreement, please contact the
Copyright Licensing Office, MSU, East Lansing, Michigan 48824
U.S.A. (517) 355-2186.

License: MSU grants End User the royalty-free, non-exclusive,
non-transferable right to use PDFgetX2 software for research or
educational purposes.  You may not redistribute, transfer, rent,
lease, sell, lend, sub-license, prepare derivative works, decompile,
or reverse-engineer the PDFgetX2 Software without prior express
written consent of MSU at the above address.

MSU retains title to PDFgetX2, including without limitation the
Software and Documentation.  End User agrees to use reasonable efforts
to protect the Software and Documentation from unauthorized use,
reproduction, distribution, or publication.  All rights not
specifically granted in this Agreement are reserved by MSU.
```

```
Warranty: PDFgetX2 Software and Documentation are provided "as is."
MSU MAKES NO WARRANTY, EXPRESS OR IMPLIED, TO END USER OR TO ANY OTHER
PERSON OR ENTITY.  SPECIFICALLY, MSU MAKES NO WARRANTY OF
MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OF PDFGETX2
SOFTWARE OR DOCUMENTATION.  MSU WILL NOT BE LIABLE FOR SPECIAL,
INCIDENTAL, CONSEQUENTIAL, INDIRECT OR OTHER SIMILAR DAMAGES, EVEN IF
MSU OR ITS EMPLOYEES HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH
DAMAGES.  IN NO EVENT WILL MSU LIABILITY FOR ANY DAMAGES TO END USER
OR ANY PERSON EVER EXCEED THE FEE PAID FOR THE LICENSE TO USE THE
SOFTWARE, REGARDLESS OF THE FORM OF THE CLAIM.

General: If any provision of this Agreement is unlawful, void, or for
any reason unenforceable, it shall be deemed severable from, and shall
in no way affect the validity or enforceability of the remaining
provisions of this Agreement. This Agreement shall be governed by
Michigan law.

Do you ACCEPT the above statement? (Y/N)y
Default installation directory: /usr/local
Default location for executeable: /usr/local/bin
Use different directories? (Y/N)y
NO space is allowed in the path!!!
New installation directory:/u24/xiangyun/softsrc
New executeable directory:/u24/xiangyun/bin
Start installation ...
      installation directory: /u24/xiangyun/softsrc
      location for executeable: /u24/xiangyun/bin
coping files ...
making link ...
installation successfully completed, thanks!
please type: pdfgetx2 to run the program
```

A.2.2 Windows

The file you should download is *pdfgetx2-xxxx-yymmdd-win.exe*. Double click on the file will start the self-installation. An example of this starting step is shown in Fig. A.2.2. Following the GUI instructions should complete the installation. A shortcut will be created on the desktop. The file extension *.xhst* will be associated with program PDFgetX2. The PDFgetX2 group is also inserted to your programs menu.

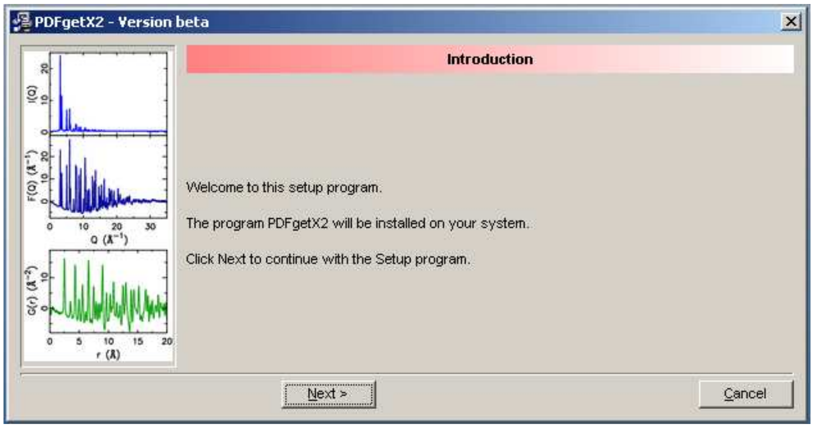


Figure A.1: The windows self-installer GUI

A.2.3 Others

If the above two methods do not apply for your system, the last resort is to manually install the program PDFgetX2, which is in fact also rather straightforward. In this case, you can either download the *pdfgetx2-xxxx-yymmdd.tar.gz* or *pdfgetx2-xxxx-yymmdd.zip*. Once the file downloading is done, unzip and/or untar all file into a directory where you want to install the program. The file *pdfgetx2.sav* should be able to run now, provided IDL is already installed.

A.3 How to start

Proper starting of PDFgetX2 will first display some message to screen and then show the logo window as in Fig. A.3

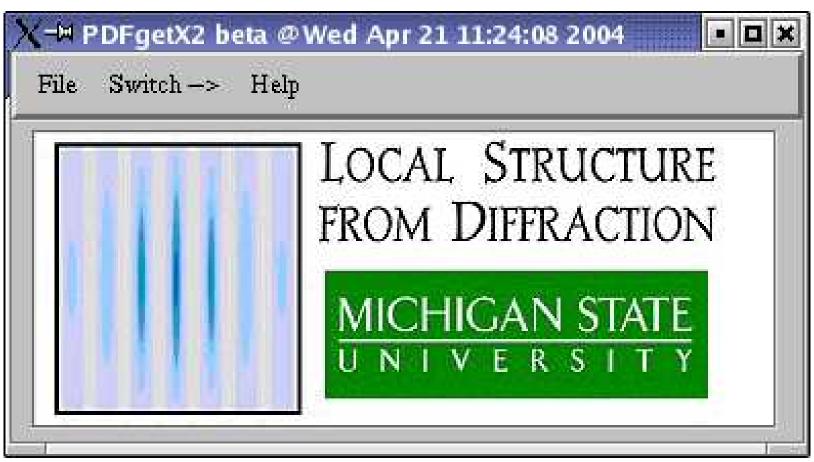


Figure A.2: The welcome interface when starting PDFgetX2

A.3.1 Linux/UNIX system

Run *pdfgetx2* to invoke the program from terminals. Also make sure the executable directory is included in your PATH variable. You can also give one PDFgetX2 history file as an argument from the command line, and the history file will be automatically loaded. The directory where you start the program will become the **Working Directory**.

A.3.2 Windows

If you installed the PDFgetX2 with the self-installer, a shortcut will be created on your desktop, and the extension *.xhst* is associated with program PDFgetX2. Double click on the desktop shortcut or on any **.xhst* file will invoke the program. The double clicked history file will be automatically loaded, and its directory will become the **Working Directory**.

A.3.3 Others

If none of the above works for you, you may have to go to the installation directory, and click on *pdfgetx2.sav* to fire it up. If you want to run it from command-line, you may try this: *idl -vm=[the full path]pdfgetx2.sav*.

A.4 Report problems and suggestions

If you have any problems in installing & running PDFgetX2 or have any suggestions, please send email to the following address:

billinge@pa.msu.edu

Good luck! Please remember to cite the paper describing this program (see Section 1.3 of this manual) in your publications where you made use of PDFgetX2.

Appendix B

Supported File Formats

Data files have to be in ASCII format. The program doesn't set the single line length limit, which would depend on your system. For all supported file formats, you can first specify the number of lines to skip from the beginning. You can also specify the comment identifier which starts a comment line instead of data. Changing the file format is done by clicking on either the **Edit Format** (Fig. 2.3) or **Customize ...** (Fig. 2.11) button , a GUI is provided to facilitate format selection.

B.1 SPEC file format

The SPEC file format here refers to one with compatible data format with the output file from the common diffractometer control and data acquisition program SPEC^{TM}. First, let's show an example SPEC file.

```
#S 1  ascan  pmQ 1 13  600 1
#L pmQ ereal elive Epoch Seconds IC1 IC3 I_CESR PULSER TOTAL COMPTON IC2 ELASTIC
1  2.07 1.967 75931 2.11758 556914 396634 394.395 416 2866 233 31718 606
1.02  2.07 1.968 75934 2.11849 558523 396548 394.159 432 3000 217 31791 610
1.04  2.06 1.962 75936 2.10892 555188 394768 392.324 414 3030 253 31569 591
1.06  2.07 1.969 75939 2.11886 558933 396616 394.023 417 3138 240 31776 647
1.08  2.07 1.977 75942 2.1189 559126 396636 393.919 419 2923 246 31839 639

#S 2  ascan  pmQ 1 13  600 1
#L pmQ ereal elive Epoch Seconds IC1 IC3 I_CESR PULSER TOTAL COMPTON IC2 ELASTIC
1  2.07 1.999 77606 2.11876 490517 396566 353.616 418 2397 186 27129 533
1.02  2.069 1.997 77609 2.11807 490872 396438 353.319 415 2486 194 27167 558
1.04  2.07 1.989 77612 2.11884 489377 396583 353.419 416 2672 177 27045 536
1.06  2.07 1.996 77614 2.11884 492200 396585 353.414 428 2551 195 27218 551
1.08  2.06 1.989 77617 2.10866 488500 394682 351.707 419 2458 199 26993 550
```

Each scan is a data block delimited by spaces. The first line of each block must start with #S, followed by one or more spaces then the scan number, which MUST be unique throughout the file. All other texts after the scan number are taken as the scan title. The second line in each block (except comment lines) MUST start with #L, followed by column names. The column names need to be space separated, and have the *same number* of columns as the data. The data are simply multiple-column numbers. You can have comment lines with # as the first letter anywhere in the file, however, the first two letters cannot be either #S or #L.

SPEC is the only format allowed to contain more than one scan, and multiple scans of one measurement have to be in *one* file (assumption of the program). If a file format conversion or files merging becomes necessary, strictly follow the SPEC format described in the above paragraph. In the simplest case, you only need to add two lines in front of each scan, the "#S <scan number> <scan title>" and "#L <column names>".

B.2 n×m ASCII format

Only two kinds of information can be in this file format: comment lines and data lines. All data lines must have the same number of columns. If the number of lines to skip is specified, those starting lines will be quietly ignored.

B.3 CHI file format

In the CHI file format, the number of lines to skip is set to be three. Otherwise, it's the same as the n×m ASCII format.

Appendix C

PDFgetX2 History File

Any PDFgetX2 session at any stage can be saved into a history file, and loading of the history file will restore the saved session. Each internal parameter is given a key identified by a logical keyword in the history file. You could edit the parameter values, but this should rarely happen since all essential parameters are accessible from the GUI.

For your reference, a history file is included below. Some lines are truncated for readability.

```
# Format:
#   Two sections are in this history file:
#             1) Preprocess SPEC Data, 2)Get X-ray PDF
# First delimiter is #######(7) which separates the two,
# second is #####(5) which starts each subsection following.
# Note:
#     You may modify this history file using a text editor, however,
#   extreme caution must be paid to preserve the format. There should
#   no blank before or after "=" sign, except no info available, please
#   don not change the number of strings in each line even that line is
#   too long or too short!
#
History written: Sun Jan 18 15:53:29 2004
Produced by xiangyun
####### SPECData_Reduction #######
##### ScansData_Setup   ViewSelect_Isa=0
SpecFile=ingaas_tutorial.spec    Num_Scans=5
#L   NO.    Scannum  Valid  Select  Highlight Has_MCA
      0        1       1       1        1        0
      1        2       1       1        1        0
      2        3       1       1        1        0
```

```
      3        4       1       1        1        0
      4        5       1       1        1        0
##### SPECMCA_Setup
QColumn=pmQ    MonitorColumn=IC2    Num_Detectors=1
#L   Det_#  Used  SPEC_Col  MCACol USE_SPEC  Use_MCA
       0     1   ELASTIC          0       1       0
##### SPECCorrection_Setup
### Monitor_SetUp   TargetColumn=IC2    DTC_Isa=1
Theory_Isa=0   TotoalCountsColumn=<click_me>
MeasuringTimeColumn=<click_me>   DeadTime=3.000000
Pulser_Isa=0   PulserColumn=<click_me>
None_Isa=1
### Detector_SetUp   TargetColumn=ELASTIC    DTC_Isa=1    IDetector=0
Theory_Isa=0   TotoalCountsColumn=TOTAL
MeasuringTimeColumn=Seconds   DeadTime=11.600000
Pulser_Isa=1   PulserColumn=PULSER
None_Isa=0
## Merge_Setup  Normalize_Isa=1   NormColumn=IC2   NormColumnType=1
   AutoScale_Isa=0   Qmin=0.100000   Qmax=100.000000
#L Scannum   Add   Multiply   Qmin  Qmax  Imin Imax
      -1  0.000000   1.000000   1.000000   12.900000   0.000000   0.000000
      -1  0.000000   1.000000   1.000000   12.900000   0.000000   0.000000
      -1  0.000000   1.000000   12.00000   40.000000   0.000000   0.000000
      -1  0.000000   1.000000   12.00000   40.000000   0.000000   0.000000
      -1  0.000000   1.000000   12.00000   40.000000   0.000000   0.000000
### ScanData_DeGlitching     Num_Glitches=42
#L  scannum   column   imin  imax
  -1    -1    -1    -1
### MergedData_DeGlitching     Num_Glitches=1
#L  scannum   column   imin  imax
  -1    -1    -1    -1
##### Plot_Setup
  PlotStyle=1   DataType=2  CompareDetectors_Isa=0
  XColumn=<click_me>    YColumn=ELASTIC  NColumn_Isa=0  NColumn=IC2

####### Get_XPDF #######

##### General_Setting
title=X-ray PDF
```

```
workingdirectory=/u24/xiangyun/projects/PDFGetX/tutorial/
quiet=0    debug=0    autosave_isa=1    savefilenamebase=ingaas_tutorial
iqfilesurfix=.iq   sqfilesurfix=.sq   fqfilesurfix=.fq   grfilesurfix=.gr

##### DataFileFormat
datatype=0    (0:SPEC, 1:CHI, 2:nxm column, 3:unknown)
num_skiplines=0    comment_id=#    delimiter=
### SPEC Format    scan_id=#S    scan_delimiter=
columnname_id=#L    columnname_delimiter=
data_id=    data_delimiter=
### CHI Format
### nxm column Format
### End of file format

##### Data&Background
samfile=ingaas_tutorial.int    num_sams=1
sambkgfile=ingaas_tutorial.bkg    num_sambkgs=1
confile=    num_cons=1
conbkgfile=    num_conbkgs=1
det#  used  xcol detcol deterrcol xmin  xmax  add_det  mul_det   add_bkg  mul_bkg
  0    1    0    1    3    1.00000    40.0000    0.00000    1.00000    0.00000    1

##### Experiment_Setup
title=InGaAs local structure
user=Dr. Jeong
facility=CHESS
temperature=300.000    containermut=0.000500000    filtermut=0.0200000
## X-Ray   radiationtype=3
           (0: Ag K_alpha, 1:Cu K_alpha, 2:Mo K_alpha, 3:Customize)
lambda=0.207800    energy=59.6655    polartype=2    polardegree=0.950000
## MonoChromator  crystaltype=0    (0:Perfect, 1:Mosaic, 2:None)
position=0    (0:Primary beam, 1:Diffracted beam)
dspacetype=2    (0:Si{111}, 1:Ge{111}, 2:Customize)    dspacing=3.13500

##### Sample_Setup information    num_atoms=3
#L symbol valence fractions z weight macoef asf_a[5] asf_b[5] asf_c compcoef_a[5] c
  Ga   0.000000   2.000000  31.000000  69.723000  39.956245  15.758946   6.841123
  As   0.000000   3.000000  33.000000  74.921590  46.441498  17.025642   4.503441
  In   0.000000   1.000000  49.000000 114.820000  20.441433   6.196477  18.816183
```

```
geometry=1    absorpCoef=1.1100000    numberdensity=0.00000
thickness=0.00000    packingFraction=0.00000    mut_theory=0.00000

##### GetIQ_Setup
xformat=2
filtercorr_isa=0    samfiltercorr_isa=0    sambkgfiltercorr_isa=0
confiltercorr_isa=0    conbkgfiltercorr_isa=0
scatveffcorr_isa=0    samconveffcorr_isa=1    sambkgveffcorr_isa=0
conbkgveffcorr_isa=0
nonegative_isa=1    negativevalue=-1.00000

##### Calibration_Data
## Detection efficiency energy dependence    detedepxaxis=0
detedepfunctype=0    detedep_elastic=1.00000    detedep_fluores=1.80000
detedep_quadra=0.00000    detedep_spline=0.00000    detedep_file=
## Detector transmission energy dependence    dettcoefxaxis=0
dettcoeffunctype=0    dettcoef_elastic=0.950000    dettcoef_fluores=0.60
dettcoef_quadra=0.00000    dettcoef_spline=0.00000    dettcoef_file=

##### IQ_Simulation
### Elastic    used_isa=1    mymethod=1
do_samabsorp=1    do_multscat=1    do_conabsorp=0    do_airabsorp=0
do_polarization=1    do_oblincident=0    do_energydep=0
do_breitdirac=0    breitdiracexpo=2.00000
do_rulandwin=0    rulandwinwidth=0.00100000
do_useredit=0    add_user=0.00000    mul_user=1.00000
### Compton    used_isa=1    mymethod=0
do_samabsorp=1    do_multscat=1    do_conabsorp=0    do_airabsorp=0
do_polarization=1    do_oblincident=0    do_energydep=0
do_breitdirac=1    breitdiracexpo=2.00000
do_rulandwin=1    rulandwinwidth=0.00300000
do_useredit=0    add_user=0.00000    mul_user=1.00000
### Fluores    used_isa=0    mymethod=1
do_samabsorp=1    do_multscat=1    do_conabsorp=0    do_airabsorp=0
do_polarization=1    do_oblincident=0    do_energydep=0
do_breitdirac=0    breitdiracexpo=2.00000
do_rulandwin=0    rulandwinwidth=0.00100000
do_useredit=0    add_user=0.00000    mul_user=1.00000
```

```
##### Correction_Setup    corrmethod=1
oblincident_isa=0    dettranscoef=0.999000    samfluore_isa=0
samfluoretype=0    samfluorescale=0.23000000
multiscat_isa=1    xraypolar_isa=1    samabsorp_isa=1
highqscale_isa=1    highqratio=0.490000    scaleconst=0.00159925
scaleconst_theory=0.00159925
comptonscat_isa=1    rulandwin_isa=1    rulandintewidth=0.00300000
comptonmethod=1    breitdirac_isa=1    breitdiracexponent=2
detefficiency_isa=0    detefficiencytype=0    (0-1: linear, 2-3: quadratic)
detefficiency_a=1.0500000    detefficiency_b=9.9999997e-05
lauediffuse_isa=1
weight_isa=1    weighttype=0    (0: <ASF>^2, 1: <ASF^2>, 2: Data Smoothed)
weightsmoothrmin=3.00000    weightsmoothwidth=100    weightsmoothcycles=600
editsq_isa=0    editsqtype=0    add_sq=0.00000    mul_sq=1.00000
editsqsmoothrmin=3.00000    editsqsmoothwidth=100    editsqsmoothcycles=600
smoothdata_isa=0   smoothfunctype=0   smoothqmin=12.0000   smoothboxwidth=9
interpolateqmin_isa=0    qmininterpolationtype=0
dampfq_isa=0    dampfqwidth=23.0000

##### SqGr_Optimization Setup
ftmethod=0    varcovarfile=grvar_covar.dat
## S(q)  qmin=1.000000    qmax=40.000000    qgrid=0.000000
## G(r)  rmin=0.000000    rmax=20.000000    rgrid=0.010000
## SqOptimization  sqoptfunction=1
optqmin=20.0000    optqmax=40.0000    optqgrid=0.00000
optrmin=0.00000    optrmax=2.20000    optrgrid=0.01000000
maxiter=20    relstep=0.00000    weighttype=2    weightfunc=0
fitbkgmult_isa=0    fitsampmut_isa=1    fitpolariz_isa=1
fitoblique_isa=0    fitfluores_isa=0
fitrulandw_isa=1    fitenergya_isa=0    fitenergyb_isa=0
fitsimurulandw_isa=1    fitDetEdepfluores_isa=0    fitDetEdepquadra_isa=0
fitDetEdepspline_isa=0   fitDetTCoefElastic_isa=0  fitDetTCoefFluores_isa=0
fitDetTcoefquadra_isa=0    fitDetTcoefspline_isa=0

##### Save&Plot Settings
datatype=SqtoFq    iqcorrtype=BkgSub    iqsimutype=SimuIq
sqcorrtype=DetEff    sqtofqtype=FqData
gropttype=OptFq    miscdatatype=AtomASF
```

Bibliography

[1] P. Debye and H. Menke, Ergeb. Tech. Röntgenk. **2**, 1 (1931).

[2] L. P. Tarasov and B. E. Warren, J. Chem. Phys. **4**, 236 (1936).

[3] B. E. Warren, H. Krutter, and O. Morningstar, **Fourier analysis of x-ray patterns of vitreous SiO_2 and B_2O_3.**, J. Am. Ceram. Soc. **19**, 202–6 (1936).

[4] S. J. L. Billinge and M. G. Kanatzidis, **Beyond Crystallography: the study of disorder, nanocrystallinity and crystallographically challenged materials**, Chem. Commun. , 749–760 (2004).

[5] T. Egami and S. J. L. Billinge, **Underneath the Bragg Peaks: Structural analysis of complex materials**, *Underneath the Bragg Peaks: Structural analysis of complex materials*, Pergamon Press, Elsevier, Oxford, England, 2003.

[6] I.-K. Jeong, J. Thompson, A. M. P. Turner, and S. J. L. Billinge, **PDFgetX: a program for determining the atomic pair distribution function from x-ray powder diffraction data**, J. Appl. Crystallogr. **34**, 536 (2001).

[7] B. J. Thijsse, **The accuracy of experimental radial distribution functions for metallic glasses**, J. Appl. Crystallogr. **17**, 61 (1984).

[8] P. J. Chupas, X. Qiu, J. C. Hanson, P. L. Lee, C. P. Grey, and S. J. L. Billinge, **Rapid Acquisition Pair Distribution Function Analysis (RA-PDF)**, J. Appl. Crystallogr. **36**, 1342–1347 (2003).

[9] J. Zaleski, G. Wu, and P. Coppens, **On the correction of reflection intensities recorded on imaging plates for incomplete absorption in the phosphor layer**, J. Appl. Crystallogr. **31**, 302 (1998).

[10] M. Ito and Y. Amemiya, **X-Ray-Energy Dependence And Uniformity Of An Imaging Plate Detector**, Nuclear Instruments & Methods In Physics Research Section A-Accelerators Spectrometers Detectors And Associated Equipment **310**, 369–372 (1991).

[11] W. Ruland, **The separation of coherent and incoherent Compton x-ray scattering**, Brit. J. Appl. Phys. **15**, 1301 (1964).

[12] S. Ergun, in *Chemistry and Physics of Carbon*, edited by J. P. L. Walker, chapter Vol. 3, pages 211–288, Marcel Dekker, New York, 1968.

[13] V. Petkov, I-K. Jeong, J. S. Chung, M. F. Thorpe, S. Kycia, and S. J. L. Billinge, **High real-space resolution measurement of the local structure of $Ga_{1-x}In_xAs$ using x-ray diffraction**, Phys. Rev. Lett. **83**, 4089 (1999).

[14] I.-K. Jeong, F. Mohiuddin-Jacobs, V. Petkov, S. J. L. Billinge, and S. Kycia, **Local structure study of $In_xGa_{1-x}As$ semiconductor alloys using high energy synchrotron x-ray diffraction**, Phys. Rev. B **63**, 205202 (2001).