Chapter 4

Getting the PDF

Kirsten M. Ø. Jensen and Simon J. L. Billinge

4.1 Introduction and Overview

In the previous chapter, you learned how to refine a structural model to a PDF. We had already obtained the PDF from synchrotron x-ray total scattering data, but in this chapter, you will do this yourself. Using either PDFgetX3 or xPDFsuite you will learn how to go from a 1D scattering pattern to the PDF. We will show this using the data from Ni as an example.

We will show all the main steps for both PDFgetX3 and xPDFsuite, both introduced in Chapter 1. While they basically do the same thing (as xPDFsuite works as a GUI for PDFgetX3), using xPDFsuite makes PDF life a lot easier and can highly increase productivity in getting data analysis done. However, xPDFsuite is commercial, while PDFgetX3 is freely available from diffpy.org. We will not reproduce the full manuals for the two programs here so it may be a good idea to go through quick-start guides or tutorials to get used to the software.

4.2 The Question

We will not really answer any scientific questions in this chapter, as we are still learning some of the fundamental steps in PDF analysis. However, we could pose the question as "How can I get a quantitatively accurate PDF from a 1D diffraction pattern?"

4.3 The Result

In this chapter the result is an optimal PDF of bulk nickel. We provide examples of Ni PDFs that have been optimised for resolution and optimised for low-noise and your job is to try and vary parameters until you get those results.

4.4 The Experiment

Total scattering data were measured from a Ni powder enclosed in a kapton capillary. We also collected data from an empty kapton capillary so that we can correct for background scattering. Everything was done at room temperature.

Facility	NSLS-II
Beamline	XPD
Detector type	Perkin Elmer amorphous silicon 2D detector
Sample geometry	powder in 1 mm ID kapton capillary
Sample environment	room temperature, ambient conditions
X-ray wavelength	0.1834 Å
Sample-detector distance	208.252 mm
exposure time	60 s

Table 4.1: Experimental conditions for data collection.

Filename	Note
`ni.chi`	chi format file of sample signal
`kapton.chi`	chi format file of sample container signal
`template.cfg`	a template config file for PDFgetX3
`ni-hires.gr`	target PDF optimised for resolution
`ni-lownoise.gr`	target PDF optimised for low noise

Table 4.2: Files for download

Note that `.chi` format files are text files that contain a header section and then data in two or three columns. In the two-column case, the first column is the momentum transfer Q or diffraction angle 2θ (depending on how the data were integrated) and the second column is the intensity. In the three column format, the third column contains the estimated standard uncertainty on the intensity values. Occasionally you will encounter four column format, in which case, most commonly, the third column is the standard error on the data in the first column and the fourth column contains the standard error on the data in the second column.

4.5 What next?

1. Download the files.
2. Plot the two target PDF files to see what they look like.
3. Use PDFgetX3 or xPDFsuite to generate PDFs:
   (a) Set the configuration parameters to start the data reduction.
   (b) Load the data files within PDFgetX3/xPDFsuite to look at it, and generate plots of $I(Q)$, $F(Q)$ and $G(r)$.
   (c) Dynamically play with the reduction parameters to see the effect on the PDF of

Q-range and other parameters.

(d) When you are satisfied, try and set the parameters to make a PDF that is optimised for high resolution (sharp peaks).

(e) Save the data as a `.gr` file.

(f) Repeat the last three steps but optimise the PDF for low-noise.

(g) Exit PDFGETX3/XPDFSUITE.

4. Make a plot of the target PDFs with the PDFs you obtained plotted on top for comparison.

4.6 Wait, what? How do I do that?

4.6.1 Download the files

Follow the instructions in Chapter 3 to download the files need for this chapter.

4.6.2 Plot the two target PDF files to see what they look like

If you have a preferred plotting program, then go for it. Plot those nickel files however you want. They should be easily readable as they are text files with two columns of data, the independent variable r in the first column and the dependent variable, $G(r)$, in the second column. Before the data columns start, there is a header of 23 lines.

If you are using PDFGETX3 you can also use a plotting utility that comes with it. Open a command prompt or terminal window, navigate to the working directory and type

```
> plotdata ni-hires.gr
```

to plot one, or

```
> plotdata ni-hires.gr ni-lownoise.gr
```

to plot them both on top of each other. PLOTDATA will plot all the input and output files for/from PDFGETX3.

You can also use XPDFSUITE for simple plotting. The workflow in XPDFSUITE is to first use the file-system explorer to navigate to the working directory, then to select all files you want to work with (in this case just plot!) into the active work area, then select the dataset (or sets) you want to plot and click on the 2D plot icon. You want to load in the two `.gr` files in your directory and plot them on top of each other.

Now you know what your final PDFs are going to look like. The next step is to use our Q-space data, i.e. the `.chi` files provided, to obtain plots like these.

4.6.3 Set configuration parameters and load the data.

Both XPDFSUITE and PDFGETX3 use only a few parameters to carry out the data reductions and analysis. Configuring the data reduction is slightly different in the two programs, and we here give instructions to both.

4.6.3.1 PDFgetX3

In PDFGETX3 the values of the configuration parameters are stored in a config file with an extension `.cfg`. To make it easier to get started we have included a template, `template.cfg`, with the download files so it should already be there in your directory. The PDFGETX3 config file is a plain text file and you will use a text editor to edit it, so open `template.cfg` in your favourite text editor, for example Notepad on a Windows computer. It is a good idea to save it with a different name so it does not get overwritten by another template file, or muddled up in some other way. Maybe "save-as" `my-hires.cfg` since you will be working on parameters for a high-resolution Ni PDF shortly. Then explore around inside the file to see how it is laid out and see if you can figure out how to edit the relevant parameters in the file appropriately to get started. The first input line in the config file is the data format, where you can choose between diffraction angle (2θ), or momentum transfer Q (in Å^{-1} or

nm^{-1}). If you open the data file (ni.chi) in another text editor, you can see that the format in this case is Q in $Å^{-1}$ so give the input 'QA' in 'dataformat'. Go through the rest of the config file, and fill out the other relevant input lines: Give the background file name (given above in Table 4.4) and the sample composition (Ni). Keep the starting values for *bgscale*, $r_{pol}y$, $Q_{maxinst}$, Q_{min} and Q_{max} at default values - we will get to them in Section 13.6.5.

There are many more options for inputs in the config file that can easily be applied when you are more familiar with the PDF and the software. We will not need it now, but for more information, read through the PDFGETX3 manual that can be found on the diffpy.org website.

You now want to run PDFGETX3 with the ni.chi data file and your configuration file. You will have to again work in a terminal/command-line window. Reuse the one you used earlier, or open a new one and navigate to the working directory, then type:

```
> pdfgetx3 ni.chi --config='myfile.cfg'
```

Remember to replace 'myfile.cfg' with your actual filename. If everything is set up correctly, a window with plots of $I(Q)$, $F(Q)$, and $G(r)$ should now open. If you get an error message, check the config file again and make sure that all necessary inputs are given. You have now generated the PDF! Next step is to optimize it with appropriate parameters.

4.6.3.2 xPDFsuite

If you are using xPDFSUITE all parameters involved in the PDF calculation appear in the GUI interface and you can update them in there. Open the program from the desktop. In the right-hand pane under the `Basic` tab, you can give inputs on the data and sample you will be working with. The first input to give is the data format, where you can choose between diffraction angle (2θ), or momentum transfer Q (in $Å^{-1}$ or nm^{-1}). If you open the data file (ni.chi) in a text editor, you can see that the format in this case is Q in $Å^{-1}$

so choose Q in Å^{-1} in 'dataformat'. Next you can load the background file by navigating to the file obtained from measurements of only the sample container, with filename given in Table 4.4. The composition of the sample is simply Ni. You can also change the r-grid that the PDF will be calculated on, but for now, you can stick to the default values.

Now use the file browser to load the Ni.chi file from your folder. Double click on the .chi file. Now a window should appear with the PDF. Click 'xrd', and $F(Q)$ in the top right corner as well, to see both the raw data $I(Q)$ and the Reduced Total Scattering Structure Function $F(Q)$ along with $G(r)$.

4.6.4 Play with the reduction parameters

xPDFsuite and PDFgetX3 use the same parameters in the data reduction. We now want to play with them and see the effect on the resulting PDF, and we give instructions to both programs while describing the effect of the parameters.

In PDFgetX3 the parameters are adjusted by typing

```
> tuneconfig()
```

in the terminal from which PDFgetX3 is running. Now, a window with slide bars show up that can be used to adjust the configuration parameters that we will describe below.

In xPDFsuite the parameters are tuned directly through the GUI in the `PDF` tab in the right hand panel, where you also see slide bars for each of the parameters.

4.6.4.1 Background scale

We first want to adjust the background scale. In our case, the background is the scattering signal measured for an empty kapton capillary. The kapton pattern was measured for twice as long as the Ni pattern, and the x-ray flux may have been slightly different. We therefore need to adjust the background scale so that it matches the Ni pattern. If considering the data in Q-space (the top window in the plot panel), the kapton signal has a bump in the

low Q-region, which is seen both in the empty kapton data and the Ni sample data. This bump is the scattered intensity from the amorphous kapton tube, and we want to remove this signal from our Ni data before we obtain the PDF. Since we know that the Ni sample is completely crystalline, we do not expect its structure to give rise to any scattering intensity in that region, so we can adjust the background scale so that the Ni and kapton data match up in the bump. Note that if you are working with samples that do show diffuse scattering themselves, you have to be more careful with determining the value for the background scale as we will discuss more in Chapter 10.

Background scale tuning is easily done in xPDFsuite where you can directly plot the background data along with the sample data by checking the 'Plot background' box in the plot window. You can now adjust the background scale with the slider until the two bumps are at the same intensity.

When using the default `tuneconfig()` function in PDFgetX3 you will not see the raw scattering data or the background data directly in your plot, but we can change the settings for tuneconfig. Close your current plot window, and in your terminal, where you are now in the 'interactive' PDFgetX3 mode, you can type:

```
> t2 = pdfgetter.getTransformation(2))
> tuneconfig([t2, 'fq','gr'])
```

This should open new plot and `tuneconfig` windows, where you will see the measured scattered intensity from the sample, the measured background intensity, and the difference between the two plotted. This should make it easier to find a good value for the background scale, as you can now use the slide bar to find a value where the scattered intensity and the background intensity line up. We refer to the PDFgetX3 manual to learn about `pdfgetter` and `getTransformation`.

4.6.4.2 $Q_{max-inst}$ and r_{poly}

The next parameters to adjust are $Q_{max-inst}$ and r_{poly}. Both these parameters have to do with the correction algorithm used before performing the Fourier transform of the data. xPDFsuite and PDFgetX3 use an *ad hoc* approach to PDF reduction instead of doing explicit corrections for e.g. incoherent Compton scattering and flourescence. The *ad hoc* approach is described in detail in [Juhás *et al.*, 2013], and it is important that you understand this process when you get further into PDF analysis. For now, just note that the program will determine a polynomial function that approximates all contributions to the measured scattering intensity that is *not* coherent, elastic scattering. The polynomium will then be subtracted from the data, leaving only the contribution interesting to us. For the program to do this, two inputs are needed: 1) r-poly, which is the lower limit of r (in Å) where we need reliable PDF peaks after the *ad hoc* corrections and Fourier transform, and 2) $Q_{max-inst}$ which defines the maximum value of momentum transfer Q where we have meaningful intensities. $Q_{max-inst}$ is thus an instrument parameter, which is not related to your sample, but is given by the geometry of the measurement and how you acquired the data. The Ni data were measured using a square 2D detector in the RA-PDF setup, and have subsequently been integrated to yield the 1D pattern provided here. We describe the integration process in Appendix **??**. From the intensity file, we see that the highest Q value with data is 31.26 Å. However, the data obtained at the highest scattering angles, corresponding to the highest values of Q, come from only a few pixels in the corner of the square detector, and we will therefore not include these in the physical Q-range. Instead, we will use the highest value of Q where the detector covers the full azimuthal range of the diffraction pattern. For the current Ni data, we set $Q_{max-inst}$ to 29.5 $Å^{-1}$.The r_{poly} parameter can often be kept at the default value of 0.9 Å, so do not change that for now.

4.6.4.3 Q_{min} and Q_{max}

Q_{min} and Q_{max} define the data interval that is used in the Fourier transform. Q_{min} should generally be set as low as possible, however, the beamstop (used to protect the detector from the direct x-ray beam) defines the minimum usable Q-value. By zooming into the low Q-region of the data, you can see that below 0.8 $Å^{-1}$, the data are very noisy, so we cut it off there.

Q_{max} is the upper data limit used in the Fourier transform. By extending the Q_{max} to high values, we minimise the effect of termination ripples in the PDF and get higher resolution in r, meaning that we can better distinguish neighbouring PDF peaks. However, at high Q values, the data will also often be very noisy due to the decreasing x-ray scattering power f at high momentum transfers. We do not want to include the noisiest part data in our PDF as this will affect our data analysis, so most often, the Q_{max} chosen is a compromise between resolution and noise. Use the slide bars to see what happens when you change the Q_{max}. As you will see, the Ni data used here are of high quality, so even at the highest values of Q, there is not much noise. This may be different when you start working on your own data from other samples, and in Chapter **??**, we will treat data with a lot more noise, where this compromise become more important.

4.6.5 Save the data as a .gr file

We first want to save a PDF with high resolution, so choose a suitable value for Q_{max}. In xPDFsuite you simply click on the gear icon in the top right corner, select "Save data files", chose an appropriate file name, and click the files you want to save: iq, sq, fq or gr.

In PDFgetX3 files were actually already saved when you first ran the `pdfgetx` command. However, this were done with the initial configuration values. Now that you have changed them, you have to make the program overwrite the new files. You can do that by running:

```
> processFiles()
```

in the terminal running PDFGETX3. This takes the current configuration and overwrites the old files, if 'force' is set to 'yes' in the config file. You can change what output you want to be saved in the config file under 'outputtype'.

4.6.6 Repeat the last three steps but optimise the PDF for low-noise.

Instead of optimising for high resolution, now try to limit the Q-range by choosing a lower value for Q_{max}, e.g. 15 Å^{-1} . Save your new PDF, as described above.

4.6.7 Exit PDFGETX3

Now close PDFGETX3 by typing

```
> exit()
```

If you are using XPDFSUITE just keep it open - we will use it for plotting.

4.6.8 Make a plot of the target PDFs with the PDFs you obtained

Now compare your two PDFs with eachother and with the ones we provided - plot the data however you like. If you are using XPDFSUITE you can plot the data there, as described above. If you are a PDFGETX3user, you can again use the PLOTDATA program from the command line by typing the following with appropriate file names:

```
> plotdata file1.gr file2.gr file3.gr file4.gr
```

4.7 Results

You have already seen the results - if you got good PDFs similar to the ones provided, you have succeeded!

4.8 Problems

Here, we will give you some problems to consider based on your experience in the above sections:

1. Consider your two PDFs. What gives the difference in peak width between the two?
2. How does noise in the measured data manifest itself in the PDF?
3. What factors affect the optimal value to choose for the Q_{max} parameter during data reduction?
4. What happens if you set the r_{poly} value to zero?
5. What happens if you do not subtract the kapton background from your data before obtaining your PDF?

4.9 Solution

In Fig. 4.1 we show the PDFs, i.e. the $G(r)$ functions, of Ni that were distributed with the data. The high resolution PDF is in black and the low resolution one is in red. The two PDFs shown in Fig. 4.1 were obtained from the same data-set and therefore from Ni at the same temperature. Nonetheless, careful inspection shows that the PDF peaks have different width! In the previous chapter, we told you that the width of the PDF peak depends on the atomic movement, which we described by ADPs. However, now we see that there is also a data reduction effect on the peak width. Because of the Fourier relationship between the measured data and the PDF, the real-space resolution, i.e. the r resolution is related not to the *Q-resolution* in reciprocal space, but on the *range* of data in reciprocal space. The red PDF is obtained from the same initial data-set but was processed with a lower Q_{max}, resulting in slightly broader PDF peaks. For this reason, you should *only* directly compare PDFs with each other when they have been processed with the same Q_{max}. Since the Q_{max}

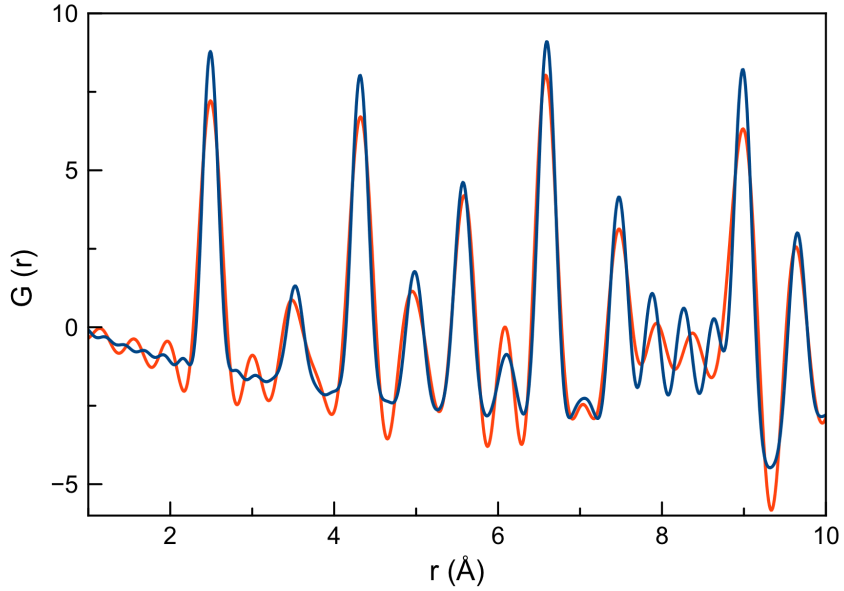

Figure 4.1: Our two PDFs: Red shows the PDF obtained with lowest Q_{max}, blue the PDF obtained with higest Q_{max}

is a known quantity, its value is given to PDF modelling programs which correct for the resolution effects so that accurate thermal/disorder distribution widths may be obtained from the modelling.

Having seen the advantage of a high Q_{max} (sharper peaks and a higher resolution measurement), why would we not always take data over the widest measured range possible? Well, we do tend to do that; however, another factor is that by extending the Q_{max} higher and higher, we introduce more and more measurement noise into the PDF because the signal-noise ratio falls off rapidly with increasing Q in the measurement. This is evident in the $F(Q)$ function plotted in Fig. 4.2 where the corresponding $I(Q)$ is also plotted. In the high-Q region the signal becomes a little more noisy. This effect is greatly exaggerated in weakly scattering samples, or for *in situ* data, where fast data collection is often important. At the cost of some real-space resolution, one can reduce the noise in the data by lowering

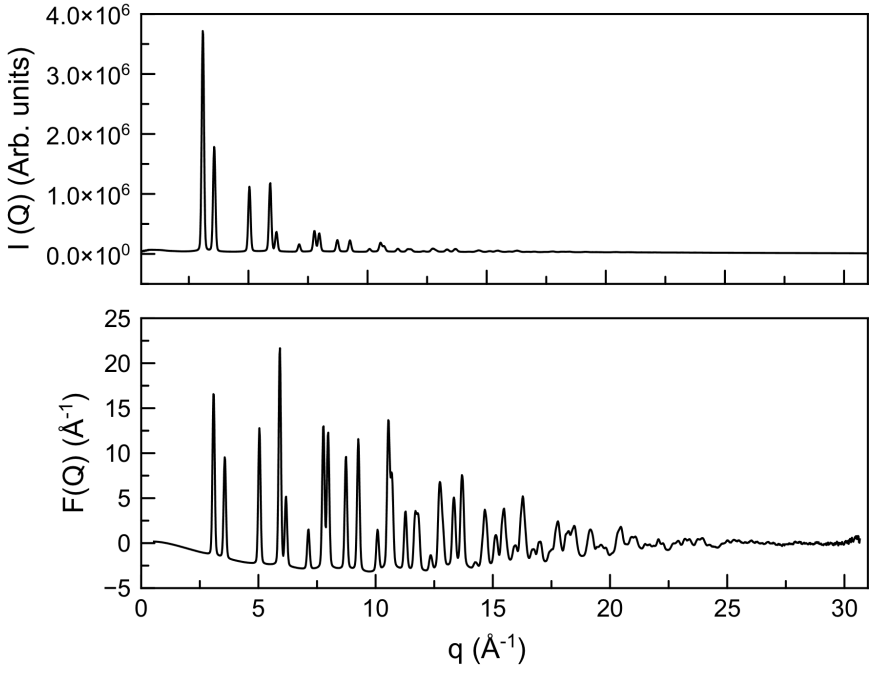

Figure 4.2: Ni $I(Q)$(top) and $F(Q)$ (bottom)

Q_{max} somewhat. Where to place Q_{max} then becomes a compromise between these two ideals, and the value chosen will depend to some extent on the scientific question under study in a particular situation. We will revisit this point in Chapter 10. If answering the question depends sensitively on seeing small features close to the baseline that are not overlapped with other peaks in the PDF, then it is better to err on the side of a lower Q_{max}. On the other hand, if it is important to separate signals from two bond-lengths that are very close to each other in distance, then a higher Q_{max} is preferred. If the data become too noisy at the required Q_{max}, then it is necessary to measure for longer, or with a more powerful beam, to get adequate counting statistics at the desired Q_{max}.

We already discussed a little bit the role of $Q_{max-inst}$ and r_{poly} when obtaining the PDF, as these are used in the *ad hoc* data correction applied in xPDFsuite and PDFgetX3. If you set the value of r_{poly} to 0 Å, you are basically telling the program not to do any corrections to the data, and you will see that unphysical features will show up in the low

r range. For more weakly scattering sample with less clear Bragg peaks, these effects are much worse, and will completely mess up your PDF. As mentioned above, the r-poly is the lower limit of r (in Å) where we need reliable PDF peaks after the *ad hoc* corrections and Fourier transform, so you should also not use a high value. The default value of 0.9 Å is usually a good starting point, as this is shorter than most chemical bonds.

If considering the data from the Ni sample, you can see that the Ni Bragg peaks are much, much more intense than the scattering signal from the kapton capillary. For this reason, you may not see much difference in the PDF with and without background subtraction, as the contribution from the kapton tube is very small. This is not the case for weakly scattering samples, where the background signal can contribute just as much or even more than the sample itself. For many experiments, the background signal may even contribute to the large majority of the signal - this could be the case e.g. in *in situ* experiments of particles and clusters in solution, or in PDF measurements of thin films on a substrate, as we will see in Chapter 10. Especially for such data, the background subtraction is a crucial step in the data reduction.

With this chapter, you now know how to generate PDFs from integrated total scattering data. We will now move on to more advanced PDF modelling.