

## Exercise 1

# The RMCProfile working environment

The aim of this exercise is to setup the basic working environment for RMCProfile and to check that the main three programs are working correctly. RMCProfile does not have a GUI at this time so everything is driven from the command line. Windows explorer can be used to set up and move or copy files, and a text editor such as wordpad will be used to edit files where necessary. The commands will then be run from a command prompt window (also known as a DOS window).

1. First, find and launch the X server – this may be called Exceed, XWIN32 or Xming. If you do not have an X server installed please install the free Xming software, a link can be found on [www.rmcpfile.org](http://www.rmcpfile.org). This program runs in the background and will be used to display graphics.
2. Next, find the **rmcpfile** folder—this will be where every you unzipped the files (e.g. **c:\rmcpfile**). Then double click the file **rmcpfile\_setup.bat**. This will launch a command prompt window with the correct environment variables set. From here on in this will be referred to as the **rmcpfile** window.
3. Create a working directory on your machine to keep all the RMCProfile files together. This can be done with windows explorer, and something along the lines of **c:\temp\myfolder** is fine.
4. Copy the folder **ex\_1** from the **rmcpfile\tutorial** folder to your new working directory.
5. In the **rmcpfile** window change the directory to this new folder; *i.e.* type

```
cd c:\temp\myfolder\ex_1
```

If your working directory is not on the **c:** drive then you will have to type the drive letter and hit return before running the 'cd' command; *i.e.*

```
H:  
cd h:\myfolder\ex_1
```

6. Then type 'dir' in the **rmcpfile** window. This will list all the files in the directory and it should look similar to the list below:

```

Directory of C:\temp\pdf_test\ex_1

31/03/2008  13:44    <DIR>          .
31/03/2008  13:44    <DIR>          ..
22/10/2007  22:56             40,495 sf6190k_all_rmc.gr
22/10/2007  22:56             138 rmcsf6_190k.back
22/10/2007  22:56            19,361 rmcsf6_190k.bragg
31/03/2008  13:49            14,089 rmcsf6_190k.cfg
31/03/2008  13:54             1,581 rmcsf6_190k.dat
22/10/2007  22:56             6,809 rmcsf6_190k.fs
22/10/2007  22:56              38 rmcsf6_190k.hkl
22/10/2007  22:56             471 rmcsf6_190k.inst
22/10/2007  22:56              47 rmcsf6_190k.poly
22/10/2007  22:56            2,758 rmcsf6_190k.sf
31/03/2008  13:59            19,704 rmcsf6_190keye.cfg
22/10/2007  22:56            32,104 sf6190k_all_conv8p82.dat
31/03/2008  13:49            30,907 rmcsf6_190k.his
31/03/2008  13:49            33,283 rmcsf6_190k.amp
31/03/2008  13:49           145,767 rmcsf6_190k.out
31/03/2008  13:49           87,320 rmcsf6_190k.braggout
          16 File(s)          434,872 bytes
          2 Dir(s)  13,156,237,312 bytes free

```

7. For now do not worry about what these file are—they are just a set of test files to ensure the programs work correctly on your machine. We will look at them in detail in the next exercise. For now, select the **rmcprofile** window, type

```
rmcprofile rmcsf6_190k
```

and hit return.

RMCPProfile should now start running. It may take a few minutes to initialise, but then it will display various messages explaining what it has been asked to do. It should finish with a set of parameters that indicate how well the data is being fitted. The final part of the output should appear similar to the text below:

```

Peaks calculated
scale 0.210434976038272
Bragg calculations initialised

----->
Time used 0.0 ,time limit 0.0 ,Last saved at 0.3

Bragg Chi**2/npts = 0.4830

10074 moves accepted 89690 generated and 85650 tested
Chi**2/dof 0.8880
Expt 1: Renorm 1.0000 Constant = 0.0000 Chi**2/nq = 2.320
Expt 2: Renorm 1.0000 Constant = 0.0000 Chi**2/nq =0.7584
Restraint energy = 3.465

Total time in loop so far : 4.6876073E-02s
Time per generated move is: 15.6253576278687 ms

Saving configuration and results to disk
*** DO NOT INTERRUPT UNTIL SAVING IS COMPLETED ***
Saving has been completed

Total time taken by program: 10.672

```

8. The next program to check is called **rmcplot**. This program displays the fits to data obtained with RMCProfile and also the partial structure factors calculated from the RMC configuration. Working again within the **rmcprofile** window, type

```
rmcplot.bat rmcsf6_190k
```

A new window should appear, called **PGPLOT Window1**—it may be hidden behind another window. This window is where the graphs will be displayed. In the **rmcprofile** window, enter '5' and hit return. The fit to the bragg profile should now be displayed. The fit will not look very good since the model being used to fit it is too small, but for now all that is important is that the fit is displayed at all.

You can use the left mouse click to zoom in on a region. Right-click will then restore the display and pressing 'q' with the graphics window active will return control to the **rmcprofile** window, from where you can make a new selection or indeed quit. This is a very basic plotting program designed to provide a quick look at the fits. For more comprehensive functionality feel free to use the csv file RMCProfile produces with your favourite plotting program, such as excel, sigmaplot, origin, *etc.*

9. The final program to check at this stage is **atomeye**. This can be used to display a ball and stick model of the RMCProfile configuration. It is a quick way to check that the atoms are not in unphysical places and also whether the configuration is physically sensible or not. This program was written by Ju Li at Ohio state university, USA and is not developed by the RMCProfile team; it is just a very useful program, so we use it.

To check that the program is working, type

```
atomeye.bat rmcsf6_190keye.cfg
```

in the **rmcprofile** window. You will have to quit **rmcplot** (by entering '0') if you have not already done so. If **atomeye** is working two new windows should appear: both are called **rmcsf6\_190k.cfg**. Don't worry if error messages are displayed in the **rmcprofile** window, so long as the new windows do in fact appear and they remain on the screen. Note that one or other or perhaps both of the new windows may initially be minimised. If this is the case, click on the new item in the task bar at the bottom of this screen.

The grey text window displays useful information, and is used to enter parameters. The new graphics window displays the image and is the main control window. With this new graphics window active the mouse can be use to move the view. Press 'F1' to display a list of option in the grey window. To quit the programs press 'q' with the graphics window active.

10. If everything worked correctly you are now ready to move on to exercise 2.

## Exercise 2

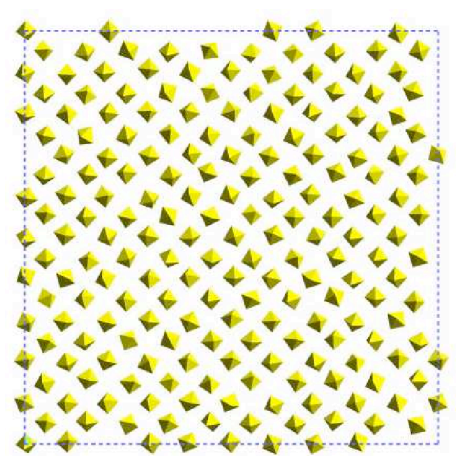
# Setting up an RMCProfile refinement

This exercise will guide you through the steps required to set up an RMCProfile refinement from scratch. It is assumed that you already have a GSAS refinement of your data and have produced corrected total scattering data. The **ex\_2** folder contains all the data you need, so please copy this to your working area. The **gsas** folder has the GSAS refinement and all data needed to provide bragg profile information and a starting configuration for RMCProfile.

The **data** folder contains time of flight neutron structure factor  $F(Q)$  data and the radial distribution  $g(r)$ . Together with the **gsas** folder, these two folders provide all the information required to run RMCProfile and produce a configuration of the system.

The **rmc** folder contains a completed run together with all the relevant output files; it is included to illustrate what the results will (hopefully) look like.

The example files supplied are for the molecular crystal  $\text{SF}_6$  and the following paragraph gives you some background information to help understand the aim of the RMCProfile refinement.



$\text{SF}_6$  is a disordered molecular crystal made up of  $\text{SF}_6$  octahedra arranged over the crystal lattice. Molecular crystals are often the most disordered crystalline systems, giving rise to large amounts of diffuse scattering (this is easily seen in a plot of the example data). The example data and files supplied are taken from neutron scattering experiments carried out at 190 K using the GEM diffractometer at ISIS. At this temperature the crystal structure has body centred cubic symmetry, with one molecule on the corner of the unit cell and one in the centre. However this average structure places the fluorine atoms of neighbouring molecules too close to one another, and so on a local scale the system tries to maximise the distance between these contacts by rotating the octahedra in different directions. The system is said to be frustrated, with fluorine–fluorine repulsion driving the motion of the molecules.

With the data supplied, you should be able to start with the “ideal” ordered structure and use RMCProfile to model these local deviations from the average bcc arrangement.

The preparation of the files for an RMCProfile refinement can be split into three stages: first, the information from a GSAS refinement is obtained, next the total scattering data is prepared,

and finally the control files for RMCProfile are produced. The next three sections take you through each of these stages in turn.

1. As with most RMC modelling, the process actually starts with a GSAS refinement. A GSAS refinement of SF<sub>6</sub> at 190 K is contained in the **gsas** folder and can be used to supply RMCProfile with all the information required to model the Bragg profile. It also contains the information need to produce a starting configuration of the bcc cell written out in *P1* symmetry. The program **data2config** should produce all the files you need. If however it doesn't work for your structure, GSAS supplies all the information required to build a starting configuration and it is a good exercise to try with your own system; at least it ensures that you start with a good understanding for the long-range average structure.

We will now extract the required information from the GSAS files using the **data2config** program.

In the **rmcprofile** window change the directory to the **gsas** folder; *i.e.*

```
cd c:\temp\myfolder\ex_2\gsas
```

or just

```
cd gsas
```

if you are already in the **ex\_2** directory.

Now type

```
data2config -noannotate -rmc6f -order[S F] -supercell[3 3 3] SF6_190K.TBL
```

and hit enter. The program should write out the ending message 'Job done by data2config. Enjoy your RMC journey!' and end. The command tells data2config to extract the information from the file SF6\_190K.TBL<sup>1</sup> with the following options: '-noannotate' tell the program to not request any metadata; '-rmc6f' will write out the starting configuration in the new rmc6f formation; '-order [S F]' means the file will have the atoms ordered with the sulphur atoms first and then fluorine atoms; '-supercell [3 3 3]' means generate a supercell 3×3×3 (in future if you wish to produce a larger supercell just change the numbers to the required size).

At this point you should have four file types for RMCProfile located in the **gsas** folder: **.bragg**, **.back**, **.inst** and **.rmc6f**. These are the bragg profile data, gsas background parameters, instrument peak-shape profile parameters and the starting configuration respectively. We will copy these to the correct location a bit later on.

In the **rmcprofile** window, move to the top level of the **ex\_2** folder by entering 'cd ..'

You have now produced all the files required from the GSAS refinement.

---

<sup>1</sup>The **.TBL** file is created with GSAS and to find out more about producing the GSAS files required please see the appendix

2. The next stage is to prepare the total scattering data for RMCPProfile. How to collect and produce total scattering data is beyond the scope of this simple tutorial and so will not be covered here.

The data folder contains an  $F(Q)$  file (`sf6_190k_fq.dat`) produced from three of the detector banks of the GEM diffractometer at ISIS. It is quite common to combine several banks of data into one data set, although RMCPProfile can be used to fit many  $F(Q)$  data sets simultaneously.

Take a look at the  $F(Q)$  file, using excel or any other plotting program available on your system. A basic alternative, supplied with RMCPProfile, is called **datamplot** that plots multiple files in rmc data format. To use this program to plot the  $F(Q)$  data, select the **rmcprofile** window and move into the **data** folder by entering `'cd data'`. Enter

```
datamplot sf6_190k_fq.dat sf6_190k_fq.dat
```

(the filename is entered twice since the program expects two or more files). This should then bring up a plot of the data file. To exit the plot, press 'q' with the plot window selected, then select option '5' in the **rmcprofile** window.

The data contained in the  $F(Q)$  file are scaled from  $-0.27593$  to  $0$ ; the  $0.27593$  factor is the sum of the neutron partial scattering factors and it needs to have this scale in order to be suitable for RMCPProfile (although there are now a number of scale options as detailed in the RMCPProfile manual). The data also need to be convolved with the RMCPProfile configuration box function to enable a fair comparison, for a  $3 \times 3 \times 3$  supercell, the appropriate truncation distance is  $8.82 \text{ \AA}$ . However, this version of RMCPProfile will do this convolution for you so nothing needs to be done at this stage.

The next file we need from this folder is the  $G(r)$  file (`sf6_190k_gr.dat`). This file contains the fourier transform of the  $F(Q)$  data. Once again the preparation of the  $G(r)$  data themselves will not be described here; the data are normally obtained during processing of the total scattering data. Take a look at the file using either a windows plotting program or **datamplot**. Again the  $G(r)$  file is scaled from  $-0.27593$  to  $0$  and not  $0$  to  $1$  as is perhaps more standard. RMCPProfile comes with a basic program for rescaling data called **data\_rescale** if you have data that needs rescaling when trying to use your own data, although as mentioned above most formats can now be read directly by RMCPProfile.

Next we need to copy the  $F(Q)$  and  $G(r)$  files into the top level of the **ex\_2** folder. In the **rmcprofile** window type

```
copy * ..
```

this copies all the files in the current directory to the one above.

3. The next file to prepare is the main **.dat** control file for RMCPProfile. This can be copied from the **rmc** folder to the top level of the **ex\_2** folder and is called `rmcsf6_190k.dat`. The `'rmcsf6_190k'` is called the "stem name" and this needs to be the same for all the files RMCPProfile uses except the experimental data files.

Assuming you are still in the **data** directory, In the **rmcprofile** window type:

```
cd ..  
copy rmc\rmcsf6_190k.dat .
```

You should open the `rmcsf6_190k.dat` control file with a text editor (e.g., wordpad or notepad++) and check that the names of the data files in the `.dat` file match the names of the files from the **data** directory. A description of the `.dat` file format and contents is given in the RMCPProfile manual.

Next we need to copy the starting configuration and bragg files from the **gsas folder** as mentioned above these need to have the same “stem name” as the main `.dat` file.

In the **rmcprofile** window type `dir`

This should produce a list of the files in the `ex_2` folder and should look as follows:

```
RMCPProfile: ex_2 >dir /b
data gsas rmc rmcsf6_190k.dat sf6_190k_fq.dat sf6_190k_gr.dat
```

to copy the required file from the `gsas` folder enter the following commands:

```
copy gsas\SF6_190K.rmc6f rmcsf6_190k.rmc6f
copy gsas\SF6_190K.inst rmcsf6_190k.inst
copy gsas\SF6_190K_3.bragg rmcsf6_190k.bragg
copy gsas\SF6_190K_3.back rmcsf6_190k.back
```

Here we have copied the ‘\_3.back’ and ‘\_3.bragg’ files, as at present you can only fit one bragg dataset and so here we have chosen to fit the 3rd histogram in the `gsas` refinement (please note that `data2config` may have produced other `.back` files, so it is important to always use the same numbered file for the `.bragg` and `.back` in this case 3).

4. Finally, since the  $\text{SF}_6$  molecules form an octahedron, it makes sense to use a constraint to hold these units together. There are a number of restraints available for this such as polyhedral restraints, potential restraints and distance window constraints, here we will use the latter. The distance window constraint as the name implies defines distance windows in which pairs of atoms are allowed to move. Initially RMCPProfile searches for atom pairs within the ranges defined and produces a neighbours list. Then as the refinement process proceeds these neighbours are constrained to remain within the set distance windows. To turn on this constraint a `.dw` file needs to be present. This file contains two rows of numbers, the first row defines the shortest distance pairs of atoms can have and the second line defines the longest distances. A suitable file for  $\text{SF}_6$  would be:

```
0 1.37 2.00
0 1.74 2.42
```

Here the ‘0’s are the shortest and longest distances for ‘S-S’ pairs and are a special case where we do not want to constrain the sulphur atoms, so in this way no neighbours will be found and nothing is constrained. The ‘1.37’ and ‘1.74’ are the shortest and longest distances for ‘S-F’ atom pairs and the ‘2.0’ and 2.42 are for the ‘F-F’ atom pairs. The numbers are derived by looking at the peaks in the  $G(r)$  file and choosing values encompassing the relevant peak without cutting into it. So if you plot the `sf6_190k_gr.dat` using **datamplot** by typing `datamplot sf6_190k_gr.dat sf6_190k_gr.dat` you should see that the first two peaks which are for ‘S-F’ and ‘F-F’ fit within the windows defined above (you can zoom in on these peaks using the left mouse button and unzoom by clicking the right mouse button).

Please now create a `rmcsf6_190k.dwg` file containing the numbers above either by using a text editor or copying the file from the **rmc** directory.



To check you have copied and created all the files you need type `dir /b` and the output should be similar to:

```
RMCPProfile: ex_2 >dir /b
data rmcsf6_190k.back rmcsf6_190k.inst
gsas rmcsf6_190k.bragg rmcsf6_190k.rmc6f
rmc rmcsf6_190k.dat sf6_190k_fq.dat
rmcsf6_190k.dw sf6_190k_gr.dat
```

5. You should now have all the files you need to run RMCPProfile on this system. Both the data and the size of configuration (a  $3 \times 3 \times 3$  supercell) have been chosen to run quickly. A reasonable fit should be achievable within about 10 minutes. However, the configuration is too small to ever give a useful model or fit. Doing so would normally require a large configuration and 10–20 hours of run time. The next exercise will describe how to run RMCPProfile on these files.

## Exercise 3

# Basic RMC refinement

This exercise will guide you through the steps required to run an RMCProfile refinement and perform some basic analysis of the results. There tends to be two stages to the analysis. The first is to check that the minimisation is proceeding usefully, and that the model being produced is sensible—this is our main aim here. The second type of analysis tends to be sample specific, starting once a final configuration has been reached and aiming to decide what the model tells us about the system being studied. Typically this may involve writing custom programs or scripts, and as such the process is very hard to generalise.

1. To start, copy the **ex\_3** directory into your work area, then move to this directory

```
cd c:\temp\myfolder\ex_3
```

and make a directory called **start** in it (*i.e.* 'mkdir start'). Then copy the files you created from the top level of you **ex\_2** folder into the **start** folder (*i.e.*

```
copy ..\ex_2\* start
```

Next, create a copy of the **start** folder, calling it **initial\_fit**. Type

```
xcopy/I start initial_fit
```

In the **rmcprofile** window change the directory to **ex\_3\initial\_fit**; *i.e.*, enter

```
cd c:\temp\myfolder\ex_3\initial_fit
```

and then 'dir /b'. The listing that appears should be similar to the one below. Check that all the “stem names” have been kept updated and are the same:

```
RMCProfile: initial_fit >dir /b
rmcsf6_190k.back rmcsf6_190k.dw sf6_190k_fq.dat
rmcsf6_190k.bragg rmcsf6_190k.inst sf6_190k_gr.dat
rmcsf6_190k.dat rmcsf6_190k.rmc6f
```

Open the RMCPProfile **.dat** file (probably called **rmcsf6\_190k.dat**) with a text editor and change the line

```
TIME_LIMIT ::      10.00 MINUTES
SAVE_PERIOD ::      0.20 MINUTES
```

to

```
TIME_LIMIT ::      0.00 MINUTES
SAVE_PERIOD ::      0.00 MINUTES
```

and save the file. This change will tell RMCPProfile to initialise and generate an initial fit and then stop.

2. Next type 'rmcprofile [stem name]' to run RMCPProfile; e.g.,

```
rmcprofile rmcsf6_190k
```

RMCPProfile should now run for a little while, and if all seems OK then move on to the next step. However, if you see an error message as below:

```
Using files: rmcsf6_190k
```

```
-----
XML/CML output requested
Info> Read history file if present
  Reading configuration file rmcsf6_190k.his6f
Info> ntypes,ni: 2
           54          324
Info> Elements in configuration:
S
F
Info> Number of atom types = 2
Info> Number of partials = 3
Material: SF6
Phase: BCC
Sample temperature: 190K
Investigator: Matt
Info> Numbers of data sets = 1 1 0 0
Status> Calculating partial weighting coefficients.
Info> Assuming no renormalisation of neutron PDF data required
At line 75 of file ../main/rmcprofile_procedures.f90
Fortran runtime error: No such file or directory
```

then the name of the  $G(r)$  data file in the main **.dat** file is not correct or the file has not been copied. You would need to edit the **.dat** file and change the name to the correct one or copy the missing file into the current directory. Once you have saved the new file, rerun RMCPProfile as above.

3. After a short while (being nice and vague as it depends on the machine being used), RMCPProfile should finish initialising and then display the level of the current fit, as below.

```
Chi**2/dof  1145.
Expt  1: Renorm 1.0000  Constant = 0.0000  Chi**2/nq = 565.2
Expt  2: Renorm 1.0000  Constant = 0.0000  Chi**2/nq = 114.7
denergy:  0.0000000000000000
Distance window summary
ipar   n(lt)   n(ok)   n(gt)
  1     0     0     0
  2     0    324    0
  3     0    648    0
Number of moves allowed by DW constraints:          0
Number of moves rejected by DW constraints:         0

Total time in loop so far :  4.5910478E-03 s
Time per generated move is:  1.53034925460815      ms

Saving configuration and results to disk
*** DO NOT INTERRUPT UNTIL SAVING IS COMPLETED ***
Updating neighbours file
Writing mamp file
Saving has been completed

Total time taken by program:          8.011
```

In the **rmcprofile** window, type `'dir /b'` and hit enter to produce a list of the files in the directory. You should see that several new files have been produced. The **.his6f** and **.amp** files are just history files for restarting a run. The results of the refinement are contained in the **.rmc6f** file (which now contains the new atom locations and the original information will have been replaced), the **.out** file and also the **.braggout** file. These last two files contain the fits to the data, together with other useful information such as the partial radial distribution functions. They can be plotted using the **rmcplot** program. The fits are also available in the comma separated files **.csv** for plotting in other programs

Now type `'rmcplot.bat [stem name]'`—e.g., `'rmcplot.bat rmcSF6_190k'`—and select '1' from the menu when it appears. A plot of the partial  $g(r)$  functions should appear in the graphics window (you may need to find it if it is hidden behind other windows). In this case, the plot should be a series of delta functions, since all the atoms are on the ideal positions given by the average structure. Press 'q' (with the graphics window selected) to return control to the menu.

Next, select '5' from the menu, and a plot of the Bragg profile fit should appear. The position of the peaks in the RMC fit should be correct, as should the peak shape and background, since these all come directly from the GSAS data. The only aspect that will be wrong is the intensity of the peaks. Again, this is because all the atoms are on their "ideal" positions, and hence there is no thermal motion or static disorder in the configuration. Consequently, the first thing RMCPProfile will do is to displace the atoms in order to reduce the intensity of these peaks. The distance window constraint will help maintain the SF<sub>6</sub> octahedral connectivity during this process, and the total  $G(r)$  will define the final local

structure of these polyhedra. Again, press 'q' with the graphics window selected in order to return control to the menu.

Finally, select '3' from the menu (answering '1' when asked which  $T(r)$  to plot), and a plot of the fit to the total  $G(r)$  will be displayed. As with the partial  $g(r)$  functions, the total  $G(r)$  function consists of a series of delta functions. The actual data (the red curve) is lurking somewhere at the bottom of these. Use the left-click of the mouse to zoom in on the low- $r$  region and produce a plot that spans roughly from 0 to 3.5 in  $x$  and  $-0.5$  to 6 in  $y$ . You should now be able to see the data more clearly and it will be easier to see that—unlike the Bragg profile fit—the positions of the peaks are not correct. This tells us straight away that the real local structure of the SF<sub>6</sub> octahedra and their local environment is not the same as that derived from the ideal average structure. We can say this even though we haven't yet refined anything. Press 'q' to return to the menu and select '0' to exit.

To finish our look at the initial model we will now produce a 3D model of the starting configuration in **atomeye**. Type 'rmc\_to\_atomeye' in the **rmcprofile** window and hit '[Enter]', followed by typing in 'rmcsf6\_190k.rmc6f', followed by entering the mass for each atom type as required by the program. This will create an **atomeye** file called `rmcsf6_190k_eye.cfg`. To plot this file, type 'atomeye rmcsf6\_190k\_eye.cfg' in the **rmcprofile** window. With the atomeye graphics window active, pressing 'tab' will toggle use of the perspective view. If you then press 'page down' the atom spheres will shrink. Next, press 'b', and the S–F bonds will be displayed; this will make it easier to visualise the SF<sub>6</sub> octahedra. This configuration should look like a bcc array of octahedra, as defined by the average structure. Examine the structure and the **atomeye** function for a while, then when you are happy to move on, press 'q' whilst the graphics window is active.

4. The next step is to use RMCProfile to refine this model. Make a copy of the initial fit directory called **first\_run**. Open this directory and edit the **[stem name].dat** file (e.g., `rmcsf6_190k.dat`), changing the line

```
TIME_LIMIT ::      0.00 MINUTES
SAVE_PERIOD ::     0.00 MINUTES
```

back to

```
TIME_LIMIT ::     10.00 MINUTES
SAVE_PERIOD ::     0.20 MINUTES
```

and save the file. This will tell RMCProfile to refine the structure for 10 minutes. The small configuration used here will not produce a very good fit but it should illustrate the minimisation process.

In the **rmcprofile** window, change to the **first\_run** directory and type 'rmcprofile [stem name]' (e.g., 'rmcprofile rmcsf6\_190k'). At this point text should start to appear indicating what RMCProfile is doing.. The run will take 10 minutes and during this time you can also follow the progress of the fits in the another rmcprofile window using **rmcplot**. Every time you choose a different plot, **rmcplot** reloads the file, so in this way you can follow the progress by re-plotting the various fits. RMCProfile will save the current fits every 12 second (0.2 minutes).

Initially the fit to the Bragg profile and the  $F(Q)$  improves at the expense of the  $G(r)$  fit and so, as predicted, the octahedra are distorted to reduce the Bragg intensities. Eventually the fit to the Bragg data get close enough that the  $G(r)$  fit starts to dominate and RMCProfile

starts to regularize the octahedra (whose connectivity has been maintained by the distance window constraint). As the refinement progresses the general trend is for all the  $\chi^2$  to reduce; however, if you look closely you will notice that some will increase at various times. This shows that the weighting of the data is correct and that local minima are being avoided.

Once the refinement has finished, examine the fitting parameters in the **RMCPProfile** window, the  $\chi^2$  values should be below 1.

If you haven't done so already, take a look at the final fits using **rmcplot**. The fits to the Bragg profile and  $F(Q)$  should look reasonable, although there are extra peaks in the Bragg profile due to the small configuration size. The fit to the  $G(r)$  should also be fairly reasonable although the fit may not quite have finished its minimisation process, and also the noise is caused by the limited number of atoms in the configuration.

The only analysis we will do here is to produce a new atomeye file (as above) and to look to see that the atoms are more disordered.

If you would like to see the fits produced by a quick-ish run with a  $10 \times 10 \times 10$  supercell then take a look at the files in the **ex\_3\sf6** folder. Here the Bragg profile fit is probably as good as it can be. The  $G(r)$  fit is better, but still not perfect.

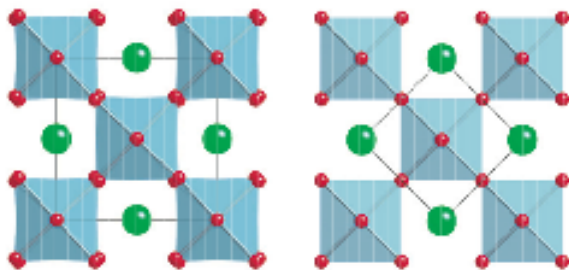
## Exercise 4

# Case Study: Using RMCProfile to study the 105K phase transition in strontium titanate

In this exercise you will use what you have learnt in the previous exercises to investigate a phase transition in the perovskite material strontium titanate ( $\text{SrTiO}_3$ ). Perovskites are made up from a network of corner sharing octahedra ( $\text{BO}_6$ ) with cations in the interstices (A) and so have the general formula  $\text{ABO}_3$ . The ideal structure is cubic with space group  $\text{Pm}\bar{3}\text{m}$ , where the A atoms are at the corners, the B atoms at the body centre and the O atoms on the cube faces. Perovskite materials have many useful physical properties, such as ferroelectricity and piezoelectricity, and are widely used in industry.

Perovskites have two main mechanisms to change their structure, the B atom can displace from the centre of the octahedra or the octahedra can tilt as a unit. These changes are usually driven by a size mismatch of the A and B atoms and often result in a distortion of the polyhedra, displacements of the A atoms and a change in space group and unit cell metric. These changes then produce the ferroelectric, piezoelectric and other useful properties these materials have.

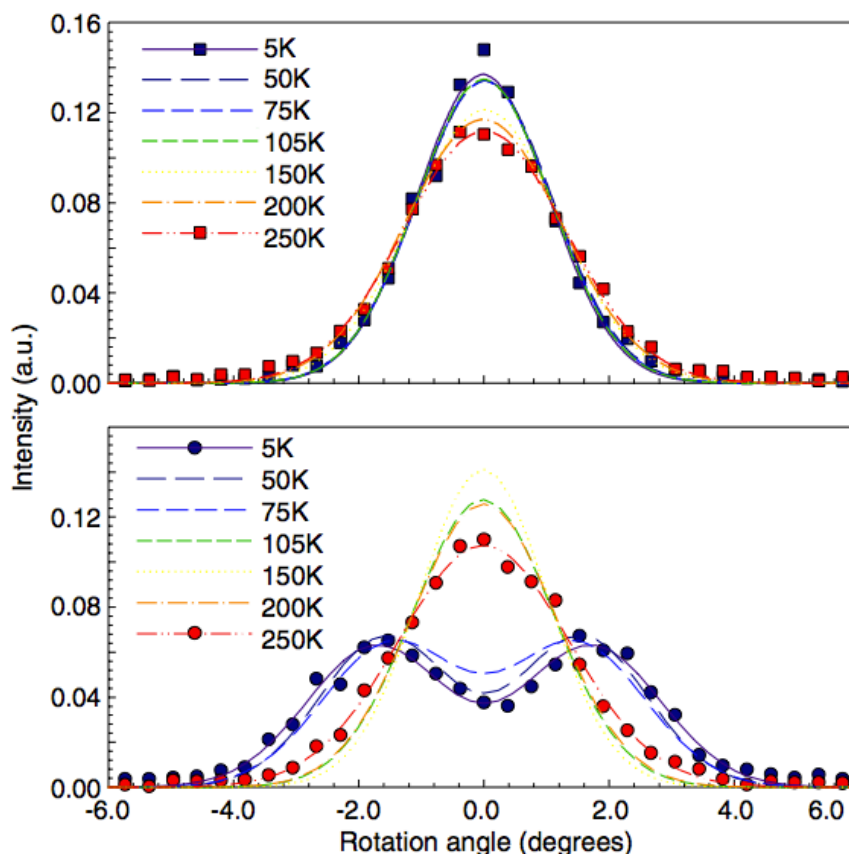
One of the simplest transition in a perovskite is the displacive transition in  $\text{SrTiO}_3$  at 105 K.  $\text{SrTiO}_3$  has the cubic  $\text{Pm}\bar{3}\text{m}$  structure at room temperature and on cooling below 105 K the octahedra begin to tilt to form a tetragonal structure, with space group  $\text{I4/mcm}$ .



The figure shows the low-temperature structure on the left and the high-temperature cubic structure on the right. If you look carefully at the low-temperature structure you should be able to see the slight displacement of the oxygen atoms (red) caused by the rotation of the octahedra. The tilting is very subtle and at 5K it is only of the order of 2 degrees from the ideal arrangement.

It is quite possible that the phase transition in  $\text{SrTiO}_3$ , together with that in quartz, is one of the best studied of all displacive phase transitions. Because of this, it is interesting to apply new techniques to this system as they become available, partly to provide a useful benchmark, and partly to see if the new technique can provide new insights. So it was natural to try running RMCProfile on this systems to see if

the method and the data were sensitive to such a small effect on a local level. The results were published in our paper (Hui et al JPCM **17** (2005) S111-S124) and the aim here is to recreate part of this analysis. The figure below shows the tilt angle distribution extracted from RMCProfile models produced by fitting neutron total scattering data. It illustrates one of the main advantage of large box modelling in that you obtain a distribution of values rather than just a single average value. The aim of this exercise is to recreate this plot for data at 5 K and 293 K.



The **ex\_4** folder contains all the files you will need for this exercise. In the top level folder there are two further folders **5K** and **293K** and these contain **data**, **gsas**, **rmc** and **gasp** folder that contain the files needed for each temperature.

1. Start by moving to the **5K** folder and using the files in the **gsas** folder to extract the bragg data and a starting configuration of a  $6 \times 6 \times 4$  supercell, with the atoms ordered as [Sr Ti O].
2. Next move back to the top level of the **5K** folder and prepare the main RMCProfile .dat control file, you can copy the file from the **RMC** folder if you don't want to start from scratch. I suggest you call it `srtio3_5k.dat` and set it to fit the neutron  $G(r)$  file `srtio3_wk_-5k_rmc.gr` and the neutron  $F(Q)$  file

```
srtio3_wk_5k_rmc.fq
```

and the bragg data with a DMIN of 0.5. Set the `TIME_LIMIT` to 10.0 minutes and the `SAVE_PERIOD` to 0.2 minutes. The .dat file in the rmc folder should be setup this way.



- Now you should copy the files you need from the **data** and **gsas** folders to the top level of the **5K** folder and you should be almost ready to run RMCPProfile. Remember the files from the **gsas** folder need to be copied to have the same “stem names” as the main `.dat` file *e.g.* `'srtio3_5k'`, in this case we will use bank 2 so you need the `'_2'` files. Also check that the names of the data files copied from the **data** folder are the same as those listed in the control file.
- Again we will use the distance window constraint to maintain the connectivity of the structure during the rmc refinement. Plot the `.gr` file with `datampplot` or another plotting program. Since this is neutron data and Ti has a negative neutron scattering length (-3.438 fm), peaks involving Ti and another atom will have a negative amplitude, such as the first peak in the data which is due to the Ti-O bonds. So the `.dw` file needs to have a first line defining the minimum distances and a second line defining the maximum distances for each pair, they are in the order (Sr-Sr, Sr-Ti, Sr-O, Ti-Ti, Ti-O and O-O) and in the data these are the 4th, 3rd, 2nd, 4th, 1st and 2nd peaks in the `.gr` file with increasing  $r$  respectively. Note that due to the symmetry of the structure some peaks are due to more than one correlation. So by examining where these peaks start and finish you should be able to decide in suitable values for the distance windows, such as:

```
3.7 3.1 2.50 3.7 1.68 2.50
4.1 3.6 3.05 4.1 2.21 3.05
```

A suitable `.dw` file containing these values is located in the `rmc` folder, so either copy this to top level of the **5K** folder once you have checked you are happy with the values (NB if you press 'c' with the graphics window of `datampplot` active a crosshairs will appear that you can use to read positions from the graph and the middle mouse button will exit this mode, or just `ctrl-C` out of the program) or produce your own files with similar values.

- You should now have all the files you need to run RMCPProfile and a listing of the directory should look similar to this:

```
data srtio3_5k.back srtio3_5k.inst
gasp srtio3_5k.bragg srtio3_5k.rmc6f
gsas srtio3_5k.dat srtio3_wk_5k_rmc.gr
rmc srtio3_5k.dw srtio3_wk_5k_rmc.fq
```

Make a directory called `first_run` and copy the RMCPProfile files to this directory *e.g.* `'copy sr* first_run'`. Now run RMCPProfile in this new folder and follow its progress in another RMCPProfile window with `rmcplot`. After 10 minutes the `Chi**2` values should be between 20 to 30. If you plot the PDF fit you will see, that while it is fairly good, it is also clear that in the rmc fit (the black line) the peaks are too intense as their areas are larger than those of the data (the red line). The area in the rmc fit is defined purely by the structure supplied (mainly the density of atoms in the configuration) so unless this is wrong the problem is that the data are not scaled correctly. This is a fairly common problem since when preparing the data it is often quite hard to be sure it is on the correct absolute scale and unfortunately, since the number of atoms in the rmc configuration is fixed, RMCPProfile is fairly intolerant to this scale issue and it will introduce extra disorder to attempt to reduce the peak areas. However, RMCPProfile can be used to indicated the correct scaling required and then the data can be corrected if possible.

Edit the main control file so the `TIME_LIMIT` and the `SAVE_PERIOD` are set to 0.0 minutes. Then in the two data sections change the keyword

NO\_FITTED\_SCALE

to

FITTED\_SCALE

then save the file. This will ask RMCPProfile to run through one cycle but scaling the data to give the best fit, so run RMCPProfile again with these settings.

After a short while RMCPProfile should end and the  $\chi^2$  values for the  $G(r)$  and  $F(Q)$  data (EXPT 1 and 2) should lower. The `Renorm` value, which was 1 before, should now be between  $\sim 1.2$  and  $\sim 1.1$ , this gives an indication of how much the data needs to be rescaled.

Since data processing is outside the scope of this tutorial we have prepared some rescaled data and put it in the data folder called `srtio3_wk_5k_sc_rmc.gr` and `srtio3_wk_5k_sc_rmc.fq`. (If of course you have already been fitting these files by mistake then you are wondering what I am wittering on about; in this case please try again fitting the correct data without the `_sc` in it.) Assuming everything so far makes sense and has worked please make a new directory called `second_run`, then copy the fresh RMCPProfile files from the top of the **5K** folder into it and the rescaled data from the **data** folder into it. Then edit the main control file so RMCPProfile will fit this new data and run RMCPProfile. If everything is correct the initial  $\chi^2$  values should be large (e.g. 'EXPT1 > 500'), otherwise the files being used are not "fresh" and the final result may not be as good.

After the 10 minute run the Bragg  $\chi^2$  should be less than 20 and the others for the  $G(r)$  and  $F(Q)$  data should be less than 7 and the fits to the data should look good. If this is not the case but they are close to these values then RMCPProfile may just need to run for a bit longer, so set the time limit to 5 minutes and rerun RMCPProfile. Once you are happy with the fits you are ready to analyse the configuration.

6. The first basic analysis, as with  $\text{SF}_6$ , is to look at the configuration visually as a ball and stick model with a suitable program such as `atomeye`. To produce the correct input file, enter `get_atomeye_srtio3 [stem name]` this should produce a file called `[stem name]eye.cfg`. Open this file in a text editor and change the line `A = 1.0 Angstrom` to `A = 10.0 Angstrom` then save the file. This is just to overcome a problem `atomeye` has with atoms that have more than 24 neighbours, you should be able to view the file now by entering `atomeye [stem name]eye.cfg`. You should be able to see the atoms have jiggled about a bit but they are basically very ordered.

Next we will analyse how the octahedra have rotated using a program called GASP, this stands for Geometric Analysis of Structural Polyhedra and as this suggests it uses geometric algebra to analyse whether the polyhedra have rotated or distorted and by how much. To do this GASP needs two configurations of atoms to compare and it will then see if you can match the atom positions by just rotating the polyhedral or if it needs to distort the polyhedra. The manual for GASP is in the **tutorial** folder of the main RMCPProfile distribution, so please look at this if you would like to know more about the program.

So to run GASP on the  $\text{SrTiO}_3$  results we need two configurations, the first is the `.rmc6f` we have just produced, so please copy this to the `gasp` folder in the **5K** folder, then move to the **gasp folder**. Next we need to create an ideal configuration with everything the same as the 5K structure apart from the tilt angle should be zero. In the **gasp** folder you should

find two crystallographic information files (.cif, if you compare these in a text editor you will see that the only difference is that the oxygen x and y positions in the

```
srtio3_5k_cubic.cif
```

are all exactly 0.25 or 0.75, whereas they should be slightly offset from this in the tetragonal structure due to the tilting, hence the '\_cubic' as it is now more like the untilted cubic structure. So we can use this file with data2config to produce the ideal configuration for the GASP analysis by entering `data2config -noannotate -order [Sr Ti O] -supercell [6 6 4] -rmc6f srtio3_5k_cubic.cif`.

Next we need to convert the two .rmc6f files into the correct format for gasp, so enter `rmctogasp` followed by entering the corresponding '.rmc6f' file name (in this case, it should be 'srtio3\_5k.rmc6f' and 'srtio3\_5k\_cubic.rmc6f', respectively for the two configurations that we are going to inspect). This will produce two new .dat files together with the corresponding .xtl files. Currently we only need the .dat files as the input for GASP.

The main GASP control file is called `srtio3_gasp.cfg` and should contain:

```
SrTiO3
Sr Ti O
OCT Ti O 2.21
```

```
PRE srtio3_5k_cubic_gasp
fast srtio3_5k_gasp srtio3_5k_cubic_gasp

com srtio3_5k_gasp srtio3_5k_cubic_gasp
```

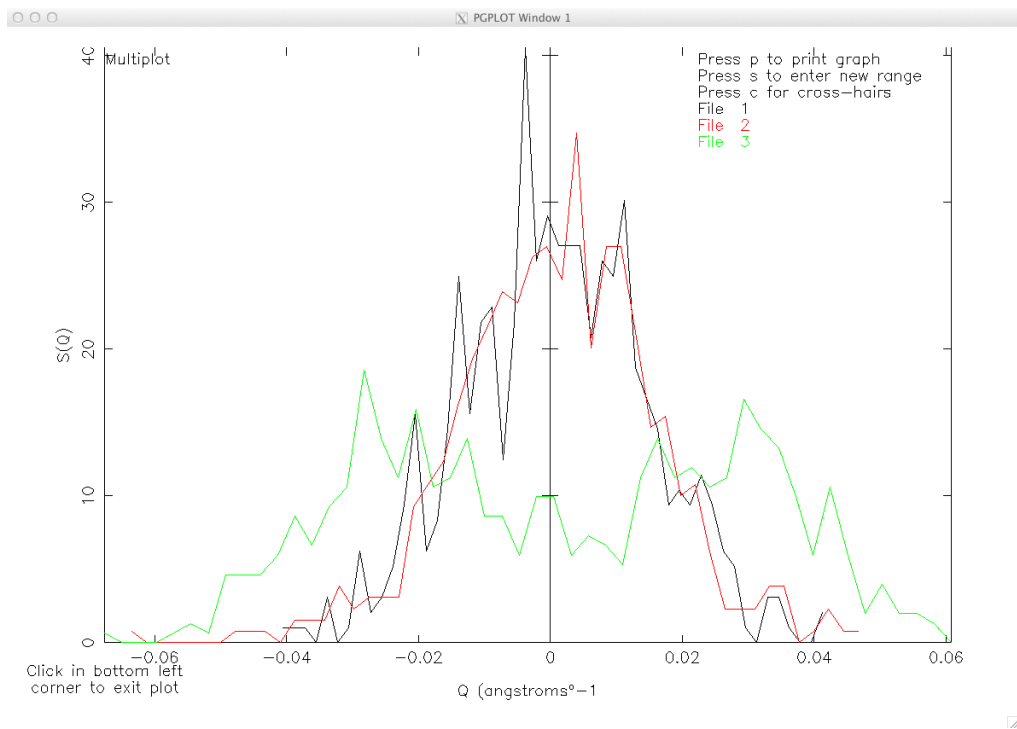
This file format is described in the GASP manual and here it is setup to compare the files `srtio3_5k` and `srtio3_5k_cubic`. so if your 'stem name' is not `srtio3_5k` you will need to change this in the file to the correct name.

To run this file with GASP just enter `gasp` and enter `srtio3_gasp` when asked Please give the name of the configuration file. Don't worry if it ends with some malloc error and/or a strange phrase possibly about Elvis, the first is just a system message and the second is from the author Steve's sense of humour! If everything has worked there would be some new .con, .pol and .com files in the folder, the .com file has the rotor comparisons we are after.

To extract the information about the octahedral tilt enter

```
get_gasp_rotors srtio3_5k_gasp_srtio3_5k_cubic_gasp.com srtio3_5k_rot
```

again this assumes you have used the 'stem name' suggested and so the .com file may have a different name, in this case change the command as needed. If everything runs well this should produce three .rot files for the x, y and z rotation direction. These are just text files containing a list of values for each of the 576 octahedra. To be able to plot these we need to histogram the data, so enter `histogram_data` followed by the name of a .rot file for each of the files, e.g. 'srtio3\_5k\_rot\_x.rot'. All being well you now



have three `.bin.dat` files and these can be plotted with **datamplot** or some other plotting program. Entering the following command (assuming the files are called as follows):

```
datamplot srtio3_5k_rot_xbin.dat srtio3_5k_rot_ybin.dat srtio3_5k_rot_zbin.dat
```

should produce a plot very similar to the one in the figure above. Clearly the distributions in files 1 and 2 are centred about zero, whereas the file 3 distribution has a maximum at about 0.03 and -0.03, these are the rotations for the z rotor in radians, so about 1.7 degrees. This is very close to the expected  $\sim 2$  degrees.

7. So the obvious next question is: does RMCPProfile always produce this double maximum distribution in one direction or is it coming from the data? To answer this question we need to analyse some data from a sample in the cubic phase. The **293K** folder contains all the data we need to do this, so please repeat the process above but with this data. This time start with the `_sc_rmc.gr` and `_sc_rmc.sq` (unless you want to repeat the rescaling part as well) and use a supercell of  $8 \times 8 \times 8$ . Again the **rmc** folder contains a control file you can start with and since the start configuration has zero tilt, the unrefined configuration can be used as the second configuration in the GASP analysis.

## Exercise 5

### Other things to try

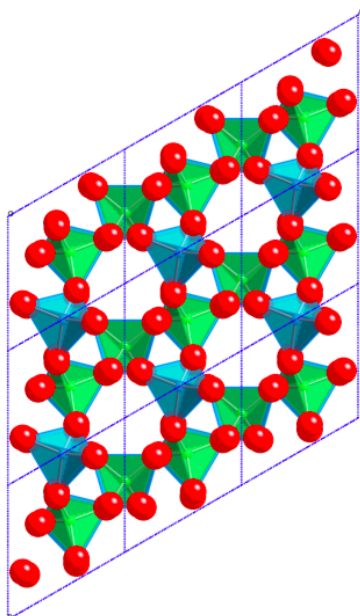
So well done if you have completed everything and reached this point! By now you should be able to happily run RMCPProfile and analyse the data to some extent at least. However, you have not used many of the features of RMCPProfile and we have skipped over quite a few things by supplying starting files. So please consult the manual for anything you don't understand (please be aware it is still a work in progress) and/or perhaps try out some of the following:

1. What effect does using the wrongly scaled data have on the final octahedral tilt results?
2. What is the effect of just fitting one or two of the different data types i.e just Bragg or just Bragg and the  $G(r)$ ?
3. If you rerun RMCPProfile with the same starting files you will get a slightly different final configuration. This can and should be used to improve the statistics of your results. So rerun RMCPProfile several time and then combine the .rot files before histogram-ing them. This should produce less noisy distributions. You can also use the runs from other people or run a single large configuration.
4. What does the **CONVOLVE ::** option do when fitting reciprocal space data (the  $F(Q)$  file) and what does the fit look like if you turn it off?
5. You can fit different functions by specifying the **FIT\_TYPE ::**. Try fitting a  $D(r)$  or  $T(r)$  from the  $G(r)$  data or a  $QF(Q)$  from the  $F(Q)$  data. What effect does this have on the final fits and results?

## Exercise 6

# Using RMCProfile with X-ray data

In the exercises so far we have been using RMCProfile with neutron data. The aim of this exercise is to guide you through the extra steps required to set up an RMCProfile refinement using x-ray data. It is assumed that you have already completed the previous exercises.



GaPO<sub>4</sub> is a piezoelectric crystal made up of corner sharing and alternating GaO<sub>4</sub> and PO<sub>4</sub> tetrahedra arranged to form an  $\alpha$ -quartz trigonal structure with a doubled  $c$ -axis. As temperature is increased, these tetrahedra undergo increasing amounts of rigid-unit motion and at high temperature undergo a phase transition to a hexagonal  $\beta$ -phase which has the  $\beta$ -quartz structure. There is an apparent discrepancy between the Ga-O and P-O bond lengths obtained from Rietveld refinements and from the PDF functions as temperature is increased, reflecting the increasing thermal disorder and difficulty of effectively modelling this using standard anisotropic atomic displacement parameters within Rietveld refinement. The example data and files supplied are taken from x-ray scattering experiments carried out at room temperature on beamline I15 at the Diamond Light Source. (The data folder also includes neutron data collected using the POLARIS diffractometer at ISIS.) At this temperature the disorder is not very pronounced, with sharp Bragg peaks observable to high- $Q$  and limited diffuse scattering in the total scattering function. The PDFs show two distinct peaks at low- $r$

corresponding to the P-O and Ga-O bonds, which at this temperature should be at very similar distances to the intra-tetrahedral bond lengths obtained from Rietveld refinement. With the data supplied, you should be able to start with the ideal ordered doubled  $\alpha$ -quartz structure and use RMCProfile to model any slight changes in bond lengths from the average structure.

The **ex\_6** folder contains all the data you need, so please copy this to your working area. In this folder you will find four subfolders: **data**, **gsas**, **rnc** and **rnc\_neutron**. It is a good idea to create a new folder in which to work, perhaps called **start**

1. The first step is to process the Bragg data, so begin by moving to the **gsas** folder. Unfortunately, it is not possible to use **data2config** with X-ray Bragg data at present, and the alternative requires that GSAS is installed on your machine. If this is not the case you will

need to copy the necessary files from the **rmc** folder and continue from step 2. If you do have GSAS installed please type `get_gsas_bragg c:\gsas\exe\` changing the directory to the location of your GSAS installation if different. Enter the name of the .EXP file, select histogram 1 when asked, and enter 0.4 as the minimum *d*-spacing. After this the program will run and produce the **.bragg**, **.back** and **.inst** files that you have seen before, as well as a **.xray** file which contains the X-ray form factor information.

2. We will now use `data2config` to generate the RMC configuration. Please type:

```
data2config -rmc6f -noannotate -order [Ga P O] -supercell [4 4 2]
gapo4_xray.cif
```

which should produce a **.rmc6f** file. Move this and the Bragg files you have created into to your working directory.

You may find that the stem name is not consistent amongst the files you have created (remember that only the total scattering data files can have different names) so you will need to rename the others to be consistent. `gapo4_xray` is a good choice.

3. The structure of  $\text{GaPO}_4$  consists of a network of corner-sharing tetrahedra, so we are going to use a polyhedral restraint to maintain the connectivity during RMC refinement. The polyhedral restraints are a somewhat historic part of the code and are system-type specific. Several auxiliary files are required which you can produce by going to your working directory and typing `get_poly_gapo4 gapo4_xray` (supply your own stem name if different). This should produce files with extensions **.alo**, **.oal**, **.po**, **.op** and **.poly**. The first four files contain the information about how the pairs of atoms are connected (the file extensions reflect the fact that the constraint is named for the analogous  $\text{AlPO}_4$  system), and the **.poly** file contains the ideal bond lengths and weights for the bond length and angle parts of the restraint.
4. Next, from the **data** folder, copy two files to your working directory:

```
gapo4_xray_rmc.gr
gapo4_xray_rmc.fq
```

5. You should now have the following files in your working directory:

```
gapo4_xray.rmc6f
gapo4_xray.xray
gapo4_xray.bragg
gapo4_xray.back
gapo4_xray.inst
gapo4_xray.alo
gapo4_xray.oal
gapo4_xray.po
gapo4_xray.op
gapo4_xray.poly
gapo4_xray_rmc.gr
gapo4_xray_rmc.fq
```

You are almost ready to run `RMCPProfile`, but first you will need a **.dat** control file. You can copy `gapo4_xray.dat` from the **rmc\start** folder if you don't want to make one from scratch. You should find several blocks of information in this file. First is the `BRAGG ::` keyword, which should be self-explanatory at this point, although note that we are using GSAS profile function 'xray2' in this case. Second is:

```

XRAY_RECIPROCAL_SPACE_DATA ::
  > FILENAME :: gapo4_xray_rmc.fq
  > DATA_TYPE :: F(Q)
  > FIT_TYPE :: F(Q)
  > START_POINT :: 1
  > END_POINT :: 1000
  > CONVOLVE ::
  > RECIPROCAL_SPACE_FIT :: 1 1000 1
  > RECIPROCAL_SPACE_PARAMETERS :: 1 1000 0.01
  > REAL_SPACE_FIT :: 1 1500 1
  > REAL_SPACE_PARAMETERS :: 1 1500 0.01

```

In which you must supply the name of your processed data file, if different (e.g. here it is probably `gapo4_xray_rmc_processed.fq`).

Starting from version 6.7.3, the X-ray PDF data is processed internally by `RMCPProfile`, which means users now do not need to provide the real space data themselves. The only data needed is the reciprocal space data, and fit the real space data users need to specify some corresponding sub-keywords under the main 'XRAY\_RECIPROCAL\_SPACE\_DATA' keyword, as shown above.

Here the crucial parameters for the real space X-ray fitting is 'REAL\_SPACE\_FIT' and 'REAL\_SPACE\_PARAMETERS'. The three numbers following 'REAL\_SPACE\_FIT' sub-keyword specifies the starting point, ending point and the number of sections (say  $N$ ) we want separate the whole range into, respectively. Then we should accordingly have  $N$  following 'REAL\_SPACE\_PARAMETERS' sub-keywords, specifying the starting and ending points together with the corresponding weight. The output for the X-ray real space fitting then can be found in the last section of the '.out' file. Also it should be mentioned that from version 6.7.4, the way to fit the reciprocal space X-ray data is changed to be the same with the real space fitting, as can be seen from the sub-keywords setup given above. Following this way, we only allow one 'XRAY\_RECIPROCAL\_SPACE\_DATA' section in the main input file. If users do have their own real space data suitable for `RMCPProfile` running, they can introduce the 'XRAY\_REAL\_SPACE\_DATA' keyword. This will not be covered in current tutorial and users are recommended to read the corresponding section of the main manual if needed.

Here it should be mentioned that for better fitting the X-ray data (especially to high- $r$  range), the facility resolution effect should be taken into account. Currently the resolution correction (including both broadening and dampening effect in  $r$ -space) has already been implemented in `RMCPProfile` but is still under active development. The users are welcome to get in touch with us if they want to try it out. To use such a resolution correction functionality, it is required that the users should measure the Si standard sample together with the measurement for their samples. Importantly, the container used for the Si standard and sample measurement should be consistent (e.g. diameter of the capillary should be identical).

The corresponding '.dat' file can be found in the `rmc\start_xray` folder, and a test running is given in the `rmc\run_xray` folder.

There is one additional line in the `.dat` file which you have not seen before:

```
POLYHEDRAL_RESTRAINT :: 5
```

This tells the program that we want to use a polyhedral restraint to maintain the integrity of the  $\text{GaPO}_4$  framework. There are several polyhedral restraints available in `RMCPProfile`,



each designed for a specific type of connectivity, and more information is available in the `RMCPprofile` manual.

6. You should now have all the files you need to run `RMCPprofile`. In your **start** folder, try running for 0 time to see how the fit looks before any refinement, and then make a copy of this folder and try running for 10 minutes or so.

We hope that this example serves to illustrate how to work with X-ray data, after such a short refinement, the fit will not be perfect by any means. The fits after several hours of running can be seen in the **rmc\run** folder. For an interesting comparison, have a look in the **rmc\_neutron** folder which contains fits to neutron data collected on the upgraded POLARIS instrument at ISIS.

## Exercise 7

# Using RMCProfile with EXAFS data

This exercise will, step by step, go through the whole process of using RMCProfile with the EXAFS data. Although neutron scattering data will also be used for the fitting, this exercise will only focus on the EXAFS part. The users are suggested to go through exercise 1–6 if you want to get familiar with using RMCProfile for scattering data.

To get the properly background-subtracted EXAFS data, you are required to have some basic knowledge about the data reduction and spectra fitting for EXAFS (e.g. using the Athena and Artemis program). Specifically for this exercise, all the experimental data including both the EXAFS and neutron scattering data have already been provided, so you don't need to worry about the data reduction for now. You can refer to the appendix-C for the introduction about extracting the suitable background-subtracted EXAFS data.

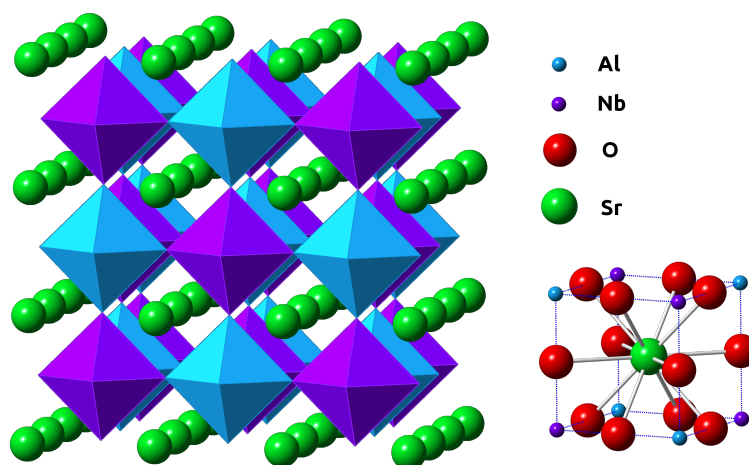


Figure 7.1: The crystal structure of SrAl<sub>0.5</sub>Nb<sub>0.5</sub>O<sub>3</sub>.

Before starting, it is necessary to point out that for using RMCProfile with EXAFS data, FEFF program is required. Therefore if you have the license for FEFF (e.g. FEFF96), you can use your own version (detailed how-to will be given later on). If you don't have the licence for FEFF, we have included the light version of FEFF (FEFF85L) in the toolkit of RMCProfile so you don't have to worry about it.

The material that will be used here is the SrAl<sub>0.5</sub>Nb<sub>0.5</sub>O<sub>3</sub> double perovskite material (see Fig. 7.1). For this material, the Al-O and Nb-O distances overlap nearly completely thus cannot be

resolved in the PDF obtained from total scattering. The same problem happens in between the Sr-O and O-O bonding. Owing to the site-selective property of X-ray absorption spectroscopy, the EXAFS spectra measured separately for the Nb and Sr K-edge could then be used to solve the problem. In this tutorial, the EXAFS spectra will be fitted altogether with the neutron data for  $\text{SrAl}_{0.5}\text{Nb}_{0.5}\text{O}_3$  to extract the structure model.

## Step-1

To start, copy the **ex\_7** directory into your work area, then move to this directory in the **rmcprofile** window (you need to launch 'RMCProfile\_setup.bat' first)

```
cd c:\temp\myfolder\ex_7
```

Type 'dir' in the **rmcprofile** window and it will show the list of items in the **ex\_7** directory:

|            |       |           |                           |
|------------|-------|-----------|---------------------------|
| 02/16/2027 | 16:18 | < DIR >   | .                         |
| 02/16/2027 | 16:18 | < DIR >   | ..                        |
| 02/16/2027 | 16:18 | < DIR >   | Configs                   |
| 02/16/2027 | 16:18 | < DIR >   | Exp                       |
| 02/16/2027 | 16:18 | < DIR >   | First_Run                 |
| 02/16/2027 | 16:18 | < DIR >   | Prep                      |
| 02/16/2027 | 16:18 | < DIR >   | RMC                       |
|            |       | 0 file(s) | 0 bytes                   |
|            |       | 7 Dir(s)  | 32,121,516,032 bytes free |

The **Configs** folder contains some input control or configuration files that will be used during either the preparation steps or the final RMC simulation. The **Exp** folder contains all the experimental data files which will be used for the RMC simulation. The **Prep** folder stores some already prepared input files for the convenience of the users to directly copy and paste. For the moment, don't worry if you cannot recognize what are included in this folder, and it will become clear as we go through this exercise. The **RMC** folder contains a complete run together with all the relevant output files. The **First\_Run** folder is probably a good place to begin from, where we have included some initial files to start the next step.

## Step-2

On terminal, go into the **First\_Run** directory,

```
cd First_Run
```

## Step-3

In this step, the program '**exafs\_inter.exe**' will be used for interpolating the scattering amplitude and phase shift of specified scattering paths. Detailed description about the scattering path in EXAFS theory (implemented by FEFF) can be found in Appendix-D. The basic idea here is that the real FEFF calculation for large system may be time consuming therefore we use the '**exafs\_inter.exe**' program to help us interpolating the scattering amplitude and phase shift of certain

scattering paths to certain grid. Then during the simulation after each atom move we don't need recalculate the spectra from scratch, which saves us the computation time.

First of all, we want to create separate folders for each type of absorption atom under the **First\_Run** folder. Here we will take the case where Nb is the absorption atom as the example and we have already created the folder for you, with the name of 'Nb\_EXAFS'. Next, we should follow the steps shown below to prepare the corresponding scattering paths files.

**Step-3-1:** Located in **First\_Run** on terminal, we want to go into '**Nb\_EXAFS**' by typing

```
cd Nb_EXAFS.
```

Then executing `dir` will give you the result similar to that shown as following,

|            |       |           |                            |
|------------|-------|-----------|----------------------------|
| 02/17/2027 | 12:18 | < DIR >   | .                          |
| 02/17/2027 | 12:18 | < DIR >   | ..                         |
| 10/24/2026 | 11:15 |           | 8,110 feff.inp             |
| 10/24/2026 | 11:15 |           | 5,220 Nb_EXAFS(k).dat      |
| 10/24/2026 | 11:15 |           | 1,393,220 snao.rmc6f       |
| 02/17/2027 | 13:56 | < DIR >   | store                      |
|            |       | 5 file(s) | 2,463,954 bytes            |
|            |       | 3 Dir(s)  | 387,319,246,848 bytes free |

**Step-3-2 (optional):** If you have your own version of FEFF (e.g., feff820.exe), you can copy it to current directory. If you don't, we packaged up a free light version of FEFF with RMCProfile. It should work per needs by RMCProfile, so you can safely skip this step.

**Step-3-3:** If you use your own FEFF following step-3-2, you can run FEFF calculation by typing the following command:

```
feff820.exe feff.inp
```

Otherwise you can use the embedded light version of FEFF (feff85L) in RMCProfile by executing:

```
feff85L.exe feff.inp
```

The simulation will run for 1 to 2 minutes depending on your machine performance. After the FEFF calculation is completed, we need to rename the generated 'paths.dat' file to 'pathsbackup.dat' by executing:

```
rename paths.dat pathsbackup.dat
```

**Step-3-4:** Open the 'feff.inp' file with any text editor you like and change the CONTROL card to:

```
CONTROL 0 0 0 0 1 1
```

**Step-3-5:** Run '**exafs\_inter.exe**' from the **rmcprofile** window, and then you will be asked for a series of questions. So please patiently follow the long steps given as below:

**Step-3-5-1:** Enter the FEFF version, e.g. '820' for feff820.exe, or '85' for the embedded 'feff85L.exe', etc. as instructed.

**Step-3-5-2:** Enter the stem name (of the structure configuration file, for current tutorial, it is 'snao').

**Step-3-5-3:** Enter the type of absorbing atom. For current tutorial, the atoms are indexed as following:

1 – Sr, 2 – Nb, 3 – Al, 4 – O

Since for the moment we are taking Nb as the absorbing atom, so we need to enter '2' in this case.

**Step-3-5-4:** Enter the reduction factor, which is the theoretical estimation for the passive amplitude factor  $S_0^2$ . This can be obtained by inspecting the 'chi.dat' file generated from the FEFF running – open the 'chi.dat' file and you can find a line that looks like 'S02 = 0.933 ...'. Therefore here we need to input '0.933'.

**Step-3-5-5:** Enter the experimental file name – Nb\_EXAFS(k) .dat.

**Step-3-5-6:** Several options will then pop up for selecting the type of the scattering paths to be interpolated. We first type '1' to select the single scattering paths.

**Step-3-5-7:** Each single scattering path has its corresponding output file (from FEFF running). For example, the file for the first scattering path is 'feff0001.dat'. First of all, the program will ask for the file number for the single scattering path 'Nb -->Sr -->Nb', which represents the scattering shown in Fig. 7.2. Now we need to open

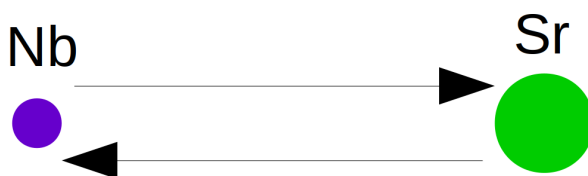


Figure 7.2: The Nb-Sr single scattering path.

the 'pathsbackup.dat' file to inspect which file corresponds to which scattering path. Here we pick up all paths that are going to be used as shown in Fig. 7.3. Then we know that for the first scattering path, we need to enter '3' (ignoring all the starting zeros) – corresponding to the 'feff0003.dat' file. The number that we should enter for the next few single scattering paths should be as following:

30  $\Rightarrow$  Nb -->Nb -->Nb, 4  $\Rightarrow$  Nb -->Al -->Nb  
1  $\Rightarrow$  Nb -->O -->Nb

After finishing the inputs for all the single scattering paths, the file named 'snao.2sc1' will be generated in the 'store' directory.

**Step-3-5-8:** Once we finish the single scattering paths, enter '2' for triple scattering paths. Then the program will ask for how many types of triple scattering paths we are going to consider, since we could have several possible triple scattering paths, namely between Nb and O, between Nb and Sr and that between Nb and Al. Here we enter '1' to consider only one of them. When asked for the type of the scatterer, enter '4' for O atom (i.e. we are only going to consider the triple scattering between Nb and O atoms), followed by entering '9' when asked for the corresponding file number (refer to Fig. 7.3). After that, the file named 'snao.2sc3' will be generated in the 'store' directory.

**Step-3-5-9:** Next, we should enter '3' to select the 'chain-like scattering paths'. The program will first take care of the following scattering path (refer to Fig. 7.3):

Nb -->O -->Al -->Nb

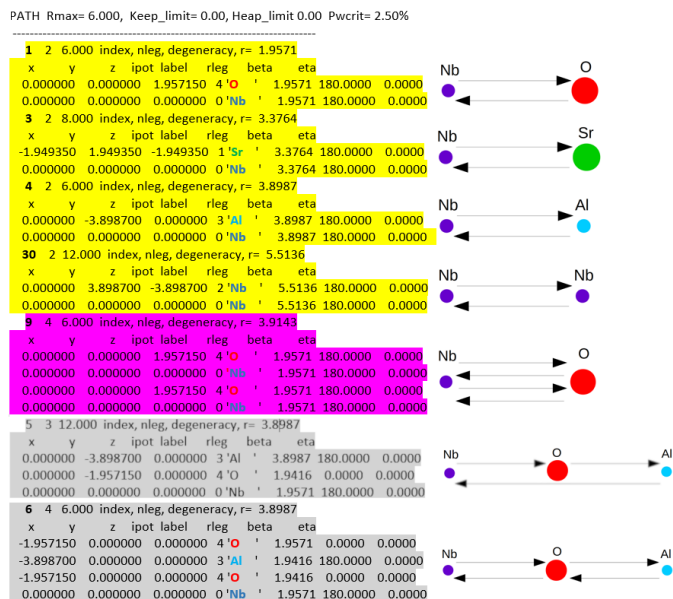


Figure 7.3: All scattering paths that are going to be used in the fitting – inspection of the 'pathsbackup.dat' file. The yellow shaded region corresponds to the single scattering path. The magenta shaded region corresponds to the triple scattering path, and the grey shaded region corresponds to the chain-like multiple scattering paths.

So when asked for the type of distant scatterer, enter '3' for Al atom, followed by entering '4' for O atom as the intervening scatterer. Next enter '5' when asked for the corresponding path file number. Then the program will ask for the minimum and maximum scattering angle. Enter, for example, '0 5' to specify the minimum and maximum scattering angle as 0° and 5°, respectively. The file named 'snao.234c2' will be generated in the 'store' directory accordingly and we are done with the first chain-like scattering path.

**Step-3-5-10:** Then enter '6' when asked for file number of the next chain-like scattering path:



Another file named 'snao.234c3' will be generated in the 'store' directory, accordingly. Here we have finished the interpolation for the scattering paths, so enter '0' to exit the **exafs\_inter.exe** program.

**Step-3-5-11:** Now if we go into the 'store' directory by typing `cd store` (hopefully you are still in the 'ex\_7\First\_Run\Nb\_EXAFS' directory in the **rmcprofile** window). Typing `dir` will show something similar as below:

```

02/17/2017  13:56  < DIR >      .
02/17/2017  12:18  < DIR >      ..
02/17/2017  14:34                59,337  snao.234c2
02/17/2017  14:34                59,337  snao.234c3
02/17/2017  14:31                29,878  snao.2sc1
02/17/2017  14:32                6,497   snao.2sc3
                4 file(s)                155,049 bytes
                2 Dir(s)                386,731,974,656 bytes free

```

Then we need to copy all files in the 'store' directory to the **First\_Run** directory by executing:

```
copy .\* ..\..\ /Y
```

**Step-3-5-12:** Next, we need to repeat all the steps from Step-3-5-1 to Step-3-5-11 for Sr as the absorbing atom, and the users are suggested to do this on their own for practice. If for the moment you don't want to follow the tedious steps above, you can skip it by copying the already generated files for Sr as the absorbing atom stored in the '**ex\_7\Prep\Sr\_EXAFS**' directory to the '**ex\_7\First\_Run**' directory.

**Step-3-5-13 (Optional):** Remove the **Nb\_EXAFS** and **Sr\_EXAFS** directories (if they exist) under the **First\_Run** directory

```
rmdir /Q /S Nb_EXAFS Sr_EXAFS
```

After **Step-3**, assuming you are located in '**ex\_7\First\_Run**' directory in the '**rmcprofile**' command window (suppose you are still in 'store' folder, executing `cd ..\..\` from command line will bring you to the **First\_Run** folder), typing `dir` will give something similar to this:

```

02/07/2017  11:32  < DIR >      .
02/07/2017  11:32  < DIR >      ..
10/24/2016  10:23                5,220   Nb_EXAFS(k).dat
10/24/2016  10:23                50,509  pdf_03305gr.txt
02/17/2017  14:36                29,542  snao.1sc1
02/17/2017  14:34                59,337  snao.234c2
02/17/2017  14:34                59,337  snao.234c3
02/17/2017  14:31                29,878  snao.2sc1
02/17/2017  14:32                6,497   snao.2sc3
10/24/2017  11:30                5,316   Sr_EXAFS(k).dat
                8 file(s)                245,636 bytes
                2 Dir(s)                387,229,655,040 bytes free

```

## Step-4

In this step, we are going to prepare the absorber and scatterer list files for the RMC simulation. The idea is that when one atom is moved during the RMC fitting, the program need to know about which atoms are influenced by the move (i.e. those atoms locally surrounding the moved atom). The file to be generated here in this step is just to store such information for each single atom in the structure configuration.

First of all, create a new folder (again, assuming you are located in '**ex\_7\Fist\_Run**' folder in the **rmcprofile** window), e.g. named '**Abs\_Scat\_List**', by typing `mkdir Abs_Scat_List`. Next

copy the 'ex\_7\Configs\scat\_input.txt' and 'ex\_7\Configs\snao.rmc6f' files to the just created **Abs\_Scat\_List** folder

```
copy ..\Configs\scat_input.txt Abs_Scat_List
copy ..\Configs\snao.rmc6f Abs_Scat_List
```

Here the **scat\_input.txt** is the input file for the **scat\_abs.exe** program. You can refer to Appendix-E for the explanation about the format of this file. Next, go into the '**Abs\_Scat\_List**' folder and run the **scat\_abs.exe** program from there.

```
cd Abs_Scat_List && scat_abs.exe
```

After running the **scat\_abs.exe** program, two files will be generated – **absorlist.dat** and **scattlist.dat**, which stores the absorber and scatterer list, respectively. We need to copy these two files to the **ex\_7\First\_Run** folder

```
copy absorlist.dat .. && copy scattlist.dat ..
```

so by executing

```
cd .. && rmdir /Q /S Abs_Scat_List && dir
```

you will have something similar to

|            |       |            |                            |
|------------|-------|------------|----------------------------|
| 02/17/2017 | 15:31 | < DIR >    | .                          |
| 02/17/2017 | 15:31 | < DIR >    | ..                         |
| 02/17/2017 | 15:55 | 1,282,140  | absorlist.dat              |
| 10/24/2016 | 10:23 | 5,220      | Nb_EXAFS(k).dat            |
| 10/24/2016 | 10:23 | 50,509     | pdf_03305gr.txt            |
| 02/17/2017 | 15:55 | 2,390,109  | scattlist.dat              |
| 02/17/2017 | 14:36 | 29,542     | snao.1sc1                  |
| 02/17/2017 | 14:34 | 59,337     | snao.234c2                 |
| 02/17/2017 | 14:34 | 59,337     | snao.234c3                 |
| 02/17/2017 | 14:31 | 29,878     | snao.2sc1                  |
| 02/17/2017 | 14:32 | 6,497      | snao.2sc3                  |
| 10/24/2016 | 10:23 | 5,316      | Sr_EXAFS(k).dat            |
|            |       | 10 file(s) | 3,917,85 bytes             |
|            |       | 2 Dir(s)   | 387,162,169,344 bytes free |

## Step-5

The last preparation step is to copy the main RMCPProfile input control file 'snao.dat' and the initial structure configuration file 'snao.rmc6f' from **ex\_7\Configs** to **ex\_7\First\_Run**

```
copy ..\Configs\snao.dat . && copy ..\Configs\snao.rmc6f .
```

So right before the real RMCPProfile running, typing **dir** will give you:



|            |       |            |                            |
|------------|-------|------------|----------------------------|
| 02/17/2017 | 15:31 | < DIR >    | .                          |
| 02/17/2017 | 15:31 | < DIR >    | ..                         |
| 02/17/2017 | 15:55 |            | 1,282,140 absorlist.dat    |
| 10/24/2016 | 10:23 |            | 5,220 Nb_EXAFS(k).dat      |
| 10/24/2016 | 10:23 |            | 50,509 pdf_03305gr.txt     |
| 02/17/2017 | 15:55 |            | 2,390,109 scattlist.dat    |
| 02/17/2017 | 14:36 |            | 29,542 snao.1sc1           |
| 02/17/2017 | 14:34 |            | 59,337 snao.234c2          |
| 02/17/2017 | 14:34 |            | 59,337 snao.234c3          |
| 02/17/2017 | 14:31 |            | 29,878 snao.2sc1           |
| 02/17/2017 | 14:32 |            | 6,497 snao.2sc3            |
| 02/06/2017 | 14:17 |            | 1,939 snao.dat             |
| 10/24/2016 | 11:55 |            | 1,393,220 snao.rmc6f       |
| 10/24/2016 | 10:23 |            | 5,316 Sr_EXAFS(k).dat      |
|            |       | 10 file(s) | 3,917,85 bytes             |
|            |       | 2 Dir(s)   | 387,162,169,344 bytes free |

Here it's worth mentioning that a brief introduction to all the files prepared from **Step-3** and **Step-4** is presented in the table of Appendix-F.

## Step-6

Now we are done with the preparations, it's time to start the RMCProfile simulation by typing `rmcprofile snao` in the **rmcprofile** window. The time limit for the RMCProfile simulation was set to 5 minutes for this tutorial (refer to the 'snao.dat' file). Therefore after 5 minutes, the simulation will be terminated so that we can examine the fitting result by making some plots.

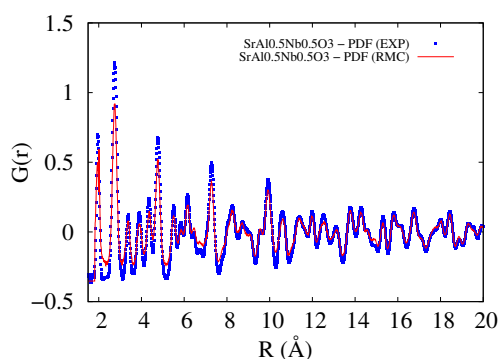


Figure 7.4: The experimental and fitted PDF pattern for  $\text{SrAl}_{0.5}\text{Nb}_{0.5}\text{O}_3$ .

The PDF pattern is stored in the **snao\_PDF1.csv** file, and the EXAFS output for both of the two atoms as the absorbing atom is stored separately in the **\*\_OUTPUT.dat** file (for current tutorial, the files are 'Nb-EXAFS-1\_OUTPUT.dat' and 'Sr-EXAFS-2\_OUTPUT.dat' for Nb and Sr as the absorbing atom, respectively). Also, for each absorber, the output in  $R$ - and  $k$ -space is given in separate files with the filenames ending with 'R\_OUTPUT.csv' and 'Q\_OUTPUT.csv', respectively. Currently the plotting functionality for EXAFS data is still under development so probably have to grab the EXAFS data from the **EXAFS\_OUTPUT.dat** file and plot it using any

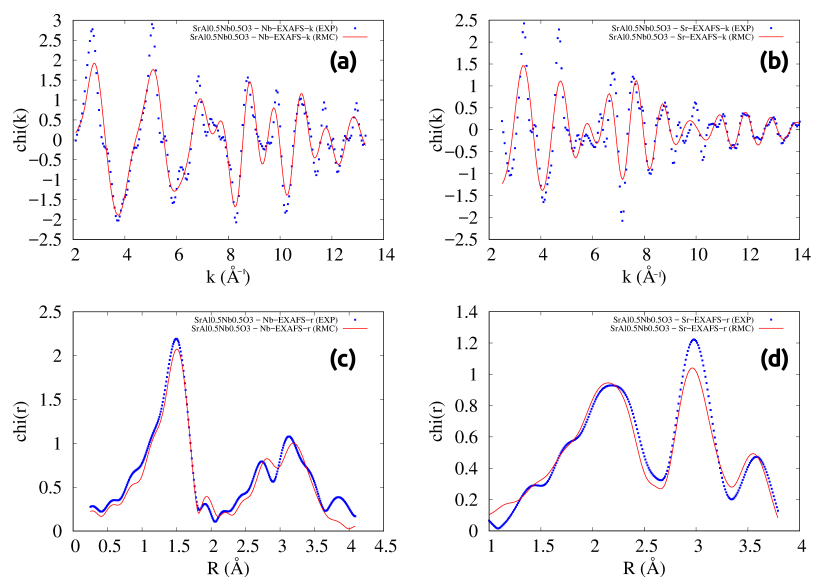


Figure 7.5: The experimental and fitted EXAFS spectra for  $\text{SrAl}_{0.5}\text{Nb}_{0.5}\text{O}_3$ . (a) and (b):  $k$ -space spectra with Nb and Sr as the absorbing atom, respectively. (c) and (d):  $R$ -space spectra with Nb and Sr as the absorbing atom, respectively.

programs you like. Here the PDF pattern is given in Fig. 7.4, and the EXAFS fitting results are presented in Fig. 7.5.

Finally, it should be mentioned that all the RMCProfile simulation input and output files for current tutorial have been included in the 'ex\_7RMC' directory so the users can compare with their own running results.

## Appendix A

# GSAS refinements

As mentioned above most RMCPprofile refinements start by producing a configuration from a Rietveld refined average structure. For historical reasons the program we use for our Rietveld refinements is GSAS and so data2config is designed to read the required information from the GSAS files. However, for data2config to be able to acquire the information it needs GSAS must be run in the following way:

1. It is best that the refinement only contains one phase, since additional phase information in the files may confuse things.
2. At present the only background function supported is the Chebyshev polynomial.
3. The d spacing range of the refinement should be set to the same range as you intend to fit in RMCPprofile.
4. Once you are happy with the refinement a `.TBL` file needs to be produced using the “pubtable” function.
5. The histogram you would like to fit in RMCPprofile needs to be written to the `.LST` file using the “hstdmp” function.

## Appendix B

# X-ray coefficients

To use `RMCPProfile` with real-space X-ray data requires an approximation which involves treating the data as neutron data with different weightings. This is possible because we have removed the  $Q$ -dependent part and can treat the atoms as having a constant 'scattering length' equivalent to the atomic number,  $Z$ . There are a few points to note when using this method.

1. You will need to normalise your X-ray  $G(r)$  so that it is scaled from 0 to 1
2. Use the `NEUTRON_REAL_SPACE_DATA` keyword in the RMC .dat file to introduce these data to your refinement.
3. You will need to provide a `NEUTRON_COEFFICIENTS ::` line under in the `NEUTRON_REAL_SPACE_DATA` block.
4. The coefficients for this line can be calculated in the same way as Faber-Ziman coefficients for neutron data, but replacing  $b$  with  $Z$  and normalising the result so that the coefficients sum to 1. An example of this for the  $\text{GaPO}_4$  example is given below.

Table B.1: Configuration information

| Atom | Z  | Concentration |
|------|----|---------------|
| Ga   | 31 | $1/6$         |
| P    | 15 | $1/6$         |
| O    | 6  | $4/6$         |

The coefficients for this system are calculated as follows:

$$(\sum_i c_i Z_i)^2 = c_{\text{Ga}}^2 Z_{\text{Ga}}^2 + 2(c_{\text{Ga}} Z_{\text{Ga}} c_{\text{P}} Z_{\text{P}}) + 2(c_{\text{Ga}} Z_{\text{Ga}} c_{\text{O}} Z_{\text{O}}) + c_{\text{P}}^2 Z_{\text{P}}^2 + 2(c_{\text{P}} Z_{\text{P}} c_{\text{O}} Z_{\text{O}}) + c_{\text{O}}^2 Z_{\text{O}}^2$$

Table B.2: Coefficients

| Ga      | P       | O       |    |
|---------|---------|---------|----|
| 26.6944 | 25.8333 | 55.1111 | Ga |
|         | 6.2500  | 26.6667 | P  |
|         |         | 28.4444 | O  |

To use these coefficients with `RMCPprofile` you need to divide by the total so that they add up to 1. Therefore the final values for the `NEUTRON_COEFFICIENTS` line are as shown below:

Table B.3: Normalised coefficients

| Ga     | P      | O      |    |
|--------|--------|--------|----|
| 0.1580 | 0.1529 | 0.3261 | Ga |
|        | 0.0370 | 0.1577 | P  |
|        |        | 0.1683 | O  |

## Appendix C

# Preparation of EXAFS data

Preliminary analyses of EXAFS data (the raw data are included in the tutorial/ex\_7/Exp directory, 'nb\_sralnb.001' – Nb as the absorbing atom, 'sr\_san.001' – Sr as the absorbing atom) are needed to obtain accurate values of the energy shifts,  $E_0$ , and to subtract the background from the data, as following:

1. Use Athena (or similar) to extract EXAFS from the absorption spectra.
2. For each experimental EXAFS build an appropriate cluster(s) around the absorbing atom. An average structure provides a good starting model. Examples of public domain software that can be used to generate these clusters based on the space group and atomic positions include Atoms/Artemis. For structures with the same absorber species located in non-equivalent crystallographic positions, separate clusters have to be generated for each of these sites. These non-equivalent sites must be designated using distinct atom types in the \*.cfg file. The contributions of these clusters to the total EXAFS signal are set proportional to the respective site occupancies.
3. Perform self-consistent FEFF calculations of the scattering paths in a separate folder for each cluster. In the FEFF input file, set the amplitude-reduction parameter  $S_0^2$  to a value less than 0.1 to let FEFF estimate this parameter from the atomic overlap integrals. Identify scattering paths to be used in the fit.
4. Use Artemis (or similar) to fit the experimental EXAFS using selected scattering paths. For all paths, set  $S_0^2$  ('amp' parameter in the GDS section) to a value calculated by FEFF. (FEFF8.20 provides satisfactorily estimates for  $S_0^2$ ), which are specified in the header of the chi.dat output file. The fitted value of  $S_0^2$  may deviate from the theoretical due, for example, to inhomogeneous sample; significant deviations likely indicate a problem and should be considered seriously. If multiple experimental EXAFS datasets are available, a simultaneous fit is recommended. Each dataset should be assigned a single value of  $E_0$ .
5. Return to the EXAFS data-reduction software (e.g. Athena) and adjust  $E_0$  to the fitted value to convert EXAFS oscillations from the energy- to k-space. Repeat the fit in Artemis using these modified data and obtain a new value of  $E_0$ . Repeat this procedure iteratively until  $E_0 < 0.5$  eV. This process minimizes systematic errors caused by the incorrect choice of  $E_0$ .
6. In Artemis, fit the background, if possible, and save the experimental data and the background as chi(k).

7. Use any spreadsheet software (e.g. Excel) to subtract the background from the experimental data. The background subtraction option in the current version of Artemis appears to work incorrectly.
8. Save the background-subtracted experimental data in the following format:
  - 1st line – number of experimental points;
  - 2nd line – title;
  - Subsequent lines –  $k$  ( $\text{\AA}^{-1}$ ) and  $\chi(k)$  values in xy format.

# Appendix D

## EXAFS technical background

### D.1 Single Scattering

Single scattering (two-leg path) contributions to EXAFS for the  $i^{th}$  absorber and  $j^{th}$  scatterer are calculated according to a formula given in the FEFF manual:

$$\chi_{ij}^{(1)} = \frac{S_i^2 \mathcal{R}(k) |F_j(\pi, k)|}{kr_{ij}^2} \sin(2kr_{ij} + 2\delta_i(k) - l\pi + \phi_j(\pi, k)) \exp(-2r_{ij}/\lambda(k)) \quad (D.1)$$

where  $S_i^2$  is the amplitude reduction factor,  $\mathcal{R}(k)$  is the total central-atom loss factor,  $r_{ij}$  is the interatomic distance,  $k$  is the photoelectron wave number,  $F_j(\pi, k) = |F_j(\pi, k)| \exp(i\phi_j(\pi, k))$  is the complex backscattering amplitude,  $[2\delta_i(k) - l\pi]$  is the total scattering phase shift for the absorbing atom, and  $\lambda(k)$  is the photoelectron mean free path.

Functions  $S_i^2$ ,  $\mathcal{R}(k)$ ,  $2\delta_i(k) - l\pi$ ,  $\lambda(k)$ ,  $|F_j(\pi, k)|$  and  $\phi_j(\pi, k)$  are calculated by FEFF. The utility program **exafs\_inter.exe** is used to interpolate these functions to k-mesh of the experimental data sets and store the parameters in the **\*.nsc1** file.

### D.2 Double and Triple Scattering

The exact equation for a double-scattering (i.e. a three-leg path, Fig. D.1) contribution that involves an absorber  $i$  and scatters  $j$  and  $n$  is expressed as:

$$\chi_{ijn}^{(2)} = \text{Im} \mathcal{C}(\theta_{nij}) \frac{S_i^2 \mathcal{R} F_j(\theta_{ijn}, k) F_n(\theta_{jni}, k)}{kr_{ij} r_{jn} r_{ni}} \exp(i(2kr_{eff} + 2\delta_i(k) - l\pi) - 2r_{eff}/\lambda(k)) \quad (D.2)$$

Here,  $r_{eff} = 1/2(r_{ij} + r_{jn} + r_{ni})$  is the effective scattering-path length,  $F_j(\theta_{ijn}, k) = |F_j(\theta_{ijn}, k)| \exp(i\phi_j(\theta_{ijn}, k))$  is the complex scattering amplitude for an atom  $j$  and the scattering angle  $\theta_{ijn}$ , and  $\mathcal{C}(\theta_{nij})$  is the angle-dependent parameter (in the plane wave approximation  $\mathcal{C}(\theta_{nij}) = \cos(\theta_{nij})$ ).

The FEFF code provides two real functions as an output:

$$F^{effective}(k) = \frac{|\mathcal{C}(\theta_{nij}) F_j(\theta_{ijn}, k) F_n(\theta_{jni}, k)|}{r_{ij} r_{jn} r_{ni}} r_{eff}^2 \quad (D.3)$$

$$\phi^{effective}(k) = \text{arg} \mathcal{C}(\theta_{nij}) F_j(\theta_{ijn}, k) F_n(\theta_{jni}, k) \quad (D.4)$$



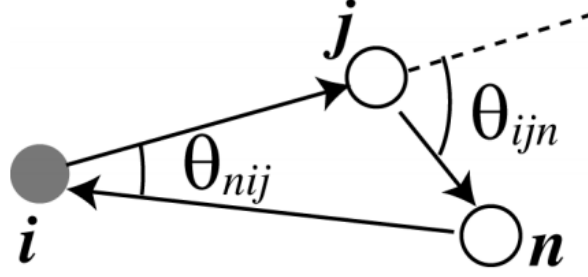


Figure D.1: A schematic representation of double scattering in a three-leg path. The scattering process involves an absorber  $i$  and scatterers  $j$  and  $n$ .

These functions can be substituted into Eq. D.1 to calculate an EXAFS signal as implemented, for example, in the IFEFFIT/Artemis software. Similarly, in the present RMC calculations, the contribution of double scattering to EXAFS is calculated using the approximate formula:

$$\chi_{ijn}^{(2)} = \frac{S_i^2 \mathcal{R} F_{jn}^{eff(2)}(\theta, k)}{k r_{ij} r_{jn} r_{ni}} \sin(2k r_{eff} + 2\delta_i(k) - l\pi + \phi_{jn}^{eff(2)}(\theta, k)) \exp(-2r_{eff}/\lambda(k)) \quad (D.5)$$

where the effective amplitude factor is

$$F_{jn}^{eff(2)}(\theta, k) = \frac{|\mathcal{C}(\theta_{nij}) F_j(\theta_{ijn}, k) F_n(\theta_{jni}, k)|}{r_{0eff}^2} r_{0ij} r_{0jn} r_{0ni} \quad (D.6)$$

and the effective amplitude factor is

$$\phi_{jn}^{eff(2)}(\theta, k) = \arg(\mathcal{C}(\theta_{nij}) F_j(\theta_{ijn}, k) F_n(\theta_{jni}, k)) \quad (D.7)$$

Here  $r_{0ij}$  is the distance between the atoms  $i$  and  $j$  in the cluster used in FEFF calculations.

Similarly, the contribution of the triple-scattering paths (Fig. D.2) that involve either one or two scatterers are calculated according to the approximate formula:

$$\chi_{ijn}^{(3)} = \frac{S_i^2 \mathcal{R} F_{jn}^{eff(3)}(\theta_{ijn}, k)}{k r_{ij}^2 r_{jn}^2} \sin(2k r_{eff} + 2\delta_i(k) - l\pi + \phi_{jn}^{eff(3)}(\theta_{ijn}, k)) \exp(-2r_{eff}/\lambda(k)) \quad (D.8)$$

$$r_{eff} = r_{ij} + r_{jn} \text{ (Fig. D.2 a)}$$

$$\chi_{ij}^{(3)} = \frac{S_i^2 \mathcal{R} F_j^{eff(3)}(\pi, k)}{k r_{ij}^4} \sin(2k r_{eff} + 2\delta_i(k) - l\pi + \phi_j^{eff(3)}(\pi, k)) \exp(-2r_{eff}/\lambda(k)) \quad (D.9)$$

$$r_{eff} = 2r_{ij} \text{ (Fig. D.2 b)}$$

$$\chi_{ijn}^{(3)} = \frac{S_i^2 \mathcal{R} F_{jn}^{eff(3)}(\theta_{nij}, k)}{k r_{ij}^2 r_{in}^2} \sin(2k r_{eff} + 2\delta_i(k) - l\pi + \phi_{jn}^{eff(3)}(\theta_{nij}, k)) \exp(-2r_{eff}/\lambda(k)) \quad (D.10)$$

$$r_{eff} = r_{ij} + r_{in} \text{ (Fig. D.2 c)}$$

Triple-scattering paths that involve three different scatterers are neglected in the present RMC refinements because of their small effect on the EXAFS signal.

Calculations of the effective amplitude factors and phase corrections are time consuming. Therefore, these characteristics are calculated prior to RMC refinements on the appropriate  $k$  and  $\theta$  meshes. During the refinements, the scattering amplitudes for the intermediate values of  $\theta$  are calculated using linear interpolation.

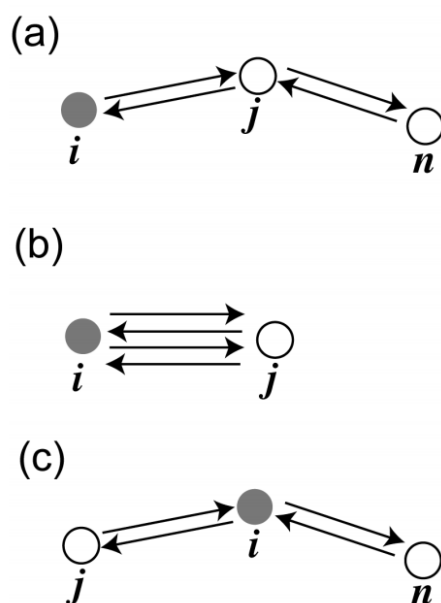


Figure D.2: Schematic representations of triple scattering paths that involve two scatterers (a,c) and a single scatterer (b).

### D.3 Double and Triple Scattering In the Nearly-Collinear Chains

Double- and triple-scattering paths that yield significant contributions to EXAFS involve nearly ( $\pm 30^\circ$ ) collinear chains (Figs. D.1, D.2, when atoms *i*, *j* and *n* are nearly located in one line) containing the intervening atom (i.e. atom *j* in Figs. D.1, D.2) in the first coordination shell of the absorber. In principle, scattering amplitudes and phase shifts depend on all three angles within the atomic triangle formed by the absorber and the scatterers. However, the present code assumes that for the nearly collinear chains these parameters are determined entirely by the  $\theta_{ijn}$  angle; this simplifying, yet sufficiently accurate, assumption was introduced to speed up the calculations. The program `exafs_inter.exe` is used to calculate the amplitudes and phase shifts as a function of  $\theta_{ijn}$  on a  $\theta$ -mesh selected by the user, and write the resulting data for all double and triple scattering paths into `*.n1n2n3c2` and `*.n1n2n3c3` files respectively.

### D.4 Double and Triple Scattering In the Nearest Coordination Spheres

Another important geometry for double scattering involves triangular paths with the scatterers located in the 1st and 2nd coordinations shells around the absorber (Fig. D.1). In this case, the effective amplitude factors and phase corrections are determined primarily by the  $\theta_{nij}$  angle. Again, the program `exafs_inter.exe` is used to calculate the amplitudes and phase shifts as a function of  $\theta_{nij}$  on the  $\theta$ -mesh selected by the user and store the resulting data for this type of double-scattering paths in the file `*.n1n2n3s2`. The amplitude and phase parameters for the triple-scattering paths shown in Fig. D.2 b are stored in the file `*.nsc3`. Similar tables for the triple-scattering paths shown in the Fig. D.2 c are stored in the file `*.n1n2n3s2`.

## Appendix E

# Format of The Input and Output Files for `scat_abs` and `exafs_inter` routines

For the situation of RMC fits without swap moves, an example of the `scat_abs` input file is shown below:

```
snao
2 NESPECTRA
Nb
4.2          0.0  4.2  0.0  4.2  5.0  4.2  0.0
Sr
4.2          0.0  4.2  0.0  4.2  0.0  4.2  0.0
```

The first line provides the stem name of the `rmc6f` file. The second line specifies the number of experimental EXAFS datasets to be included in the fit (e.g., Nb and Sr). This line is followed by the NSPECTRA 2-line blocks in arbitrary order.

The first line in each block specifies a chemical symbol of the absorbing atom. The second line contains `ntypes` pairs of real numbers, where `ntypes` is the number of atomic types in the configuration file. The first number in the pair is the maximum distance separating the corresponding scattering atoms from the absorber. The second number is the maximum scattering angle (in degrees) to search for an intervening atom in chain-scattering paths that involve these absorber-scatterer pairs. If this number is set to 0.0, no search is performed.

NOTE: For the `*.rmc6f` format, the number of types of atoms is determined automatically from the list of chemical species in the `.dat` file.

For the situation of fits with swap moves (in this example, Nb and Al will be swapped, whereas only the Nb EXAFS is available), an example of the `scat_abs` input file is shown below:

```
snao
3 NESPECTRASWAP
Nb
4.2          0.0  4.2  0.0  4.2  5.0  4.2  0.0
Al
4.2          0.0  4.2  0.0  4.2  5.0  4.2  0.0
Sr
4.2          0.0  4.2  0.0  4.2  0.0  4.2  0.0
```

Again, the first line provides the stem name of the rmc6f file. The second line specifies the number of experimental EXAFS datasets to be included in the fit (e.g., Nb and Sr) plus the number of atom types that will be involved in the swap (Al) in case their experimental spectrum is missing. This line is followed by the NSPECTRASWAP 2-like line blocks in arbitrary order. That is, a block for Al is included even though its spectrum is unavailable.

The first line in each block specifies a chemical symbol of the absorbing atom or of the atom that is involved in the swap even though its EXAFS spectrum is missing, like Al in the present example. The second line contains ntypes pairs of real numbers, where ntypes is the number of atomic types in the configuration file. The first number in the pair is the maximum distance separating the corresponding scattering atoms from the absorber. The second number is the maximum scattering angle (in degrees) to search for an intervening atom in chain-scattering paths that involve these absorber-scatterer pairs. If this number is set to 0.0, no search is performed.

An explanation of the **scat\_abs** output files is given as below, first the `scattlist.dat` file:

The first two lines are described in the file itself. These are followed by  $N$  lines, where  $N$  is the number of atoms in the configuration. Each line starts with an entry specifying the number,  $m$ , of scatterers in this line (set to zero for non-absorbing atoms). This entry is followed by  $m$  pairs of numbers. The first number in each pair specifies the index of a scattering atom for the given absorbing atom, while the second number in each pair indicates the index of the intervening atom (if present) between the absorber and the scatterer. If this second number is zero there is no intervening atom.

```
20480 total number of lines 26 maximal number of scatterers in line 26 2049 0 ... 19457 ...  
20 394 0 ... 6538 16778 ... 0 0
```

For the `absorlist.dat` file:

The first two lines are described in the file itself. These are followed by  $N$  lines, where  $N$  is the number of atoms in the configuration. Each line starts with an entry specifying the number of absorbers in this line (set to zero for non-absorbing atoms). This entry is followed by series of numbers specifying the indices of the corresponding absorbing atoms.

```
20480 total number of lines 14 maximal number of absorbers in line 10 2049 2056 ... 5121  
5633 ... 10 2049 2050 ... 2618 3074
```

Format of the output files for the **exafs\_inter** routine is given as following, first for the single scattering:

The first two lines are described in the file itself. These lines are followed by a series of columns. 1st column -  $k$  ( $^{-1}$ ), 2nd column - mean free path of a photoelectron. These columns are followed by  $m$  pairs of columns, where  $m$  is the number of the scattering atom types. The entries in each pair correspond to the scattering amplitude and phase.

Then for the multiple scattering path:

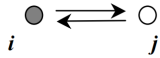
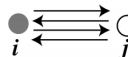
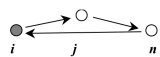
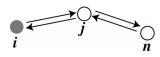
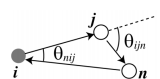

The first several lines are described in the file itself. These are followed by a list of different values of the scattering angle. For each angle, 2 columns are given, which provide values of the real and imaginary parts of the scattering amplitude on the same  $k$ -grid as specified in the single-scattering file.

For the triple-scattering paths of the kind described in the `n1sc3` files, the two columns describe the scattering and amplitudes and phases on the same  $k$ -grid as specified in the single-scattering file.

## Appendix F

# List of Files Required to Run RMCProfile with EXAFS Data

Table F.1: List of Files Required to Run RMCProfile with EXAFS Data

| Filename  | Path pictogram  | File Content  |
|---|---|---|
| *.n <sub>1</sub> sc1                              |   | photoelectron mean free path; moduli and phases of backscattering amplitudes for the single-scattering paths  |
| *.n <sub>1</sub> sc3                              |  | amplitude factors and phase corrections for the triple-scattering paths   |
| *.n <sub>1</sub> n <sub>2</sub> n <sub>3</sub> c2 |  | effective amplitude factors and phase corrections for the double-scattering paths in nearly collinear atomic chains   |
| *.n <sub>1</sub> n <sub>2</sub> n <sub>3</sub> c3 |  | effective amplitude factors and phase corrections for the triple-scattering paths in nearly collinear atomic chains   |
| *.n <sub>1</sub> n <sub>2</sub> n <sub>3</sub> s2 |  | effective amplitude factors and phase corrections for the double-scattering paths having both scattering atoms in the 1 <sup>st</sup> and 2 <sup>nd</sup> coordination shells around the absorber   |
| *.n <sub>1</sub> n <sub>2</sub> n <sub>3</sub> s3 |  | effective amplitude factors and phase corrections for the triple-scattering paths with the two scattering atoms located in the 1 <sup>st</sup> coordination shell around the absorber, and the absorber atom itself acting as a scatterer |
| absorlist.dat                                     |   | list of absorber atoms around each atom in the configuration  |
| scattlist.dat                                     |   | list of all scattering atoms for each absorber  |