Interactive **S**tructure **A**nalysis of **A**morphous and **C**rystalline **S**ystems

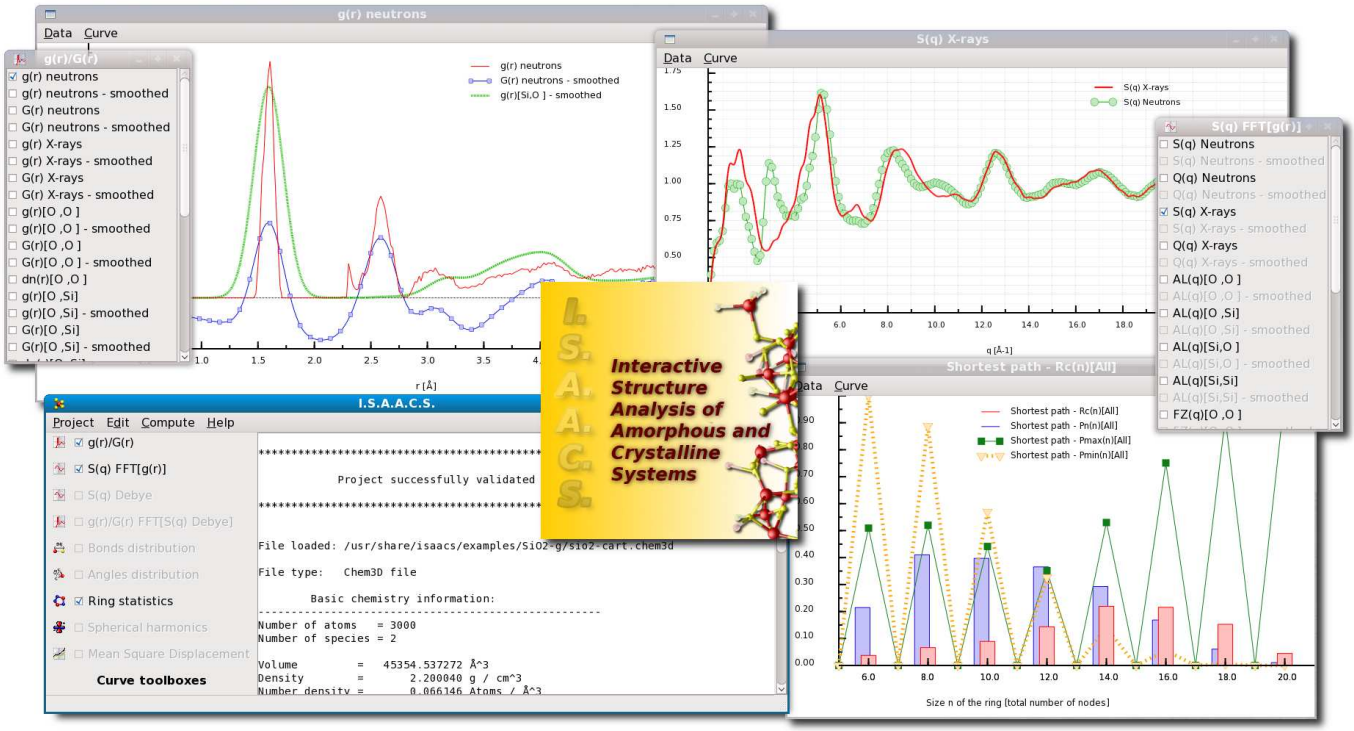
I.S.A.A.C.S.

http://isaacs.sourceforge.net/

http://people.cst.cmich.edu/petko1vg/isaacs/

Published in:

J. Appl. Cryst.(2010). **43** , 181–185



User manual I.S.A.A.C.S. v2.8

Sébastien LE ROUX sebastien.leroux@ipcms.unistra.fr

Valeri PETKOV petko1vg@cmich.edu

Contents

List of Figures

List of Tables

Introduction

I.S.A.A.C.S. **I**nteractive **S**tructure **A**nalysis of **A**morphous and **C**rystalline **S**ystems is a cross-platform program developed to analyze the structural characteristics of three-dimensional models built by computer simulations. The models may have any degree of periodicity (i.e. crystallinity) and local symmetry. The following structural information is computed from the models: total and partial radial distribution and structure factors for X-ray or neutron scattering, coordination numbers, bond angle and near atomic neighbor distributions, bond valence sums, ring statistics and spherical harmonics invariants. The information may be visualized conveniently and stored for further use.

An article describing the I.S.A.A.C.S. program and its features has been published in the *Journal of Applied Crystallography*.
Users who consider to use I.S.A.A.C.S. for research purposes should refer to this publication:

S. Le Roux and V. Petkov. *J. Appl. Cryst.*, **43**:181-185 (2010).

Programming framework

I.S.A.A.C.S. is developed in C, FORTRAN90 and GTK+ [1] for the **G**raphical **U**ser **I**nterface. The C part of the code is used as a binding to wrap the GTK+ GUI over the FORTRAN90 core routines used for the calculations. Separating the I.S.A.A.C.S. GUI from the FORTRAN90 routines makes the latter very easy to re-implement in other programs. Basically the FORTRAN90 routines are controlled by the GUI with minor exceptions where GTK+ functions are called from FORTRAN90 routines to update a bar indicating the progress of calculations.

2.1 Supported platforms

The GTK+ library is a highly portable environment which allows I.S.A.A.C.S. to be a cross-platform software. Microsoft Windows (32 bits), Linux (32 and 64 bits), as well as Mac OS X (Intel-based Macintoshes) versions of the program are available.

2.2 The I.S.A.A.C.S. Project File format

I.S.A.A.C.S. program uses an intuitive format for a project file (see Tab. 2.1) which contains all parameters needed to set up a calculation of structural characteristics of a 3D model. The structure of the file follows the XML coding [2] and allows to store detailed information about the system to be analyzed: chemical composition, chemical and physical properties of each atomic species (e.g. atomic weight and x-ray/neutron scattering amplitudes), size of the model box, atomic coordinates type (e.g. Cartesian or fractional), time series properties, description of the bonding between atoms. The information is provided by the user and may be re-used/modified during I.S.A.A.C.S. execution. The '*.ipf' file illustrated in table [Tab. 2.1] shows the parameters needed to run calculations for a 3D structure model of silica glass.

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- I.S.A.A.C.S. v1.1 XML file -->
<isaacs-xml>
 <!-- Format and file of the configuration(s) -->
 <data>
  <type>Chem3D file</type>
  <file>/home/leroux/Desktop/sio2.chem3d</file>
 </data>
 <!-- Chemistry information -->
 <chemistry>
  <atoms>3000</atoms>
  <species number="2">
   <label id="0">O </label>
   <label id="1">Si</label>
  </species>
  <element symbol="O ">
   <name>Oxygen          </name>
   <z>8</z>
   <mass>16.000000</mass>
   <rad>0</rad>
   <radius>0.660000</radius>
   <nscatt>5.803000</nscatt>
   <xscatt>8.000000</xscatt>
  </element>
  <element symbol="Si">
   <name>Silicon         </name>
   <z>14</z>
   <mass>28.090000</mass>
   <rad>0</rad>
   <radius>1.110000</radius>
   <nscatt>4.153000</nscatt>
   <xscatt>14.000000</xscatt>
  </element>
 </chemistry>
 <!-- Box information -->
 <box>
  <edges>
   <a>35.662100</a>
   <b>35.662100</b>
   <c>35.662100</c>
  </edges>
  <angles>
   <alpha>90.000000</alpha>
   <beta>90.000000</beta>
   <gamma>90.000000</gamma>
  </angles>
  <vectors>
   <a.x>35.662100</a.x>
   <a.y>0.000000</a.y>
   <a.z>0.000000</a.z>
   <b.x>0.000000</b.x>
   <b.y>35.662100</b.y>
   <b.z>0.000000</b.z>
   <c.x>0.000000</c.x>
   <c.y>0.000000</c.y>
   <c.z>35.662100</c.z>
  </vectors>
 </box>
 <!-- PBC information -->
 <pbc>
  <apply>TRUE</apply>
  <fractional>FALSE</fractional>
  <fractype>0</fractype>
 </pbc>
 <!-- Bonds information -->
 <cutoffs>
  <total>2.184304</total>
  <partials>
   <O-O>2.808390</O-O>
   <O-Si>2.184304</O-Si>
   <Si-O>2.184304</Si-O>
   <Si-Si>3.432477</Si-Si>
  </partials>
 </cutoffs>
 <!-- Time series -->
 <time-series>
  <dt>2.500000</dt>
  <unit>t [fs]</unit>
  <ndt>20</ndt>
 </time-series>
 <!-- Apply project -->
 <project>TRUE</project>
</isaacs-xml>
```

Table 2.1 *Example of I.S.A.A.C.S. project file '*.ipf' in XML format for glassy silica.*

Features

The main interface of the I.S.A.A.C.S. program [Fig. 3.1-a] gives access to different menus:

- The 'Project menu' [Fig. 3.1-b] is used to read and write **ipf** files [Tab. 2.1] as well as to import/export coordinates of atoms from the analyzed structure models [Sec. 3.1].
- The 'Edit menu' [Fig. 3.1-c] is used to define the properties of the structural characteristics of a system to be studied.
- The 'Compute menu' [Fig. 3.1-d] is used to run the calculations [Sec. 3.2].
- The 'Help menu' [Fig. 3.1-e] is used to access the documentation provided to help the users.

3.1 Data Inputs and Outputs

The current version of I.S.A.A.C.S. can import 3D structure models in five different formats [Tab. 3.1]:

1. XYZ [3]
2. Chem3D [4]
3. PDB (Protein Data Bank)
4. CPMD trajectory [5] *
5. VASP trajectory [6] **

Table 3.1 *Model structure files read by I.S.A.A.C.S..*

* atomic units are assumed in the case of CPMD trajectories.
* and ** require to enter extra parameters through interactive dialog boxes:

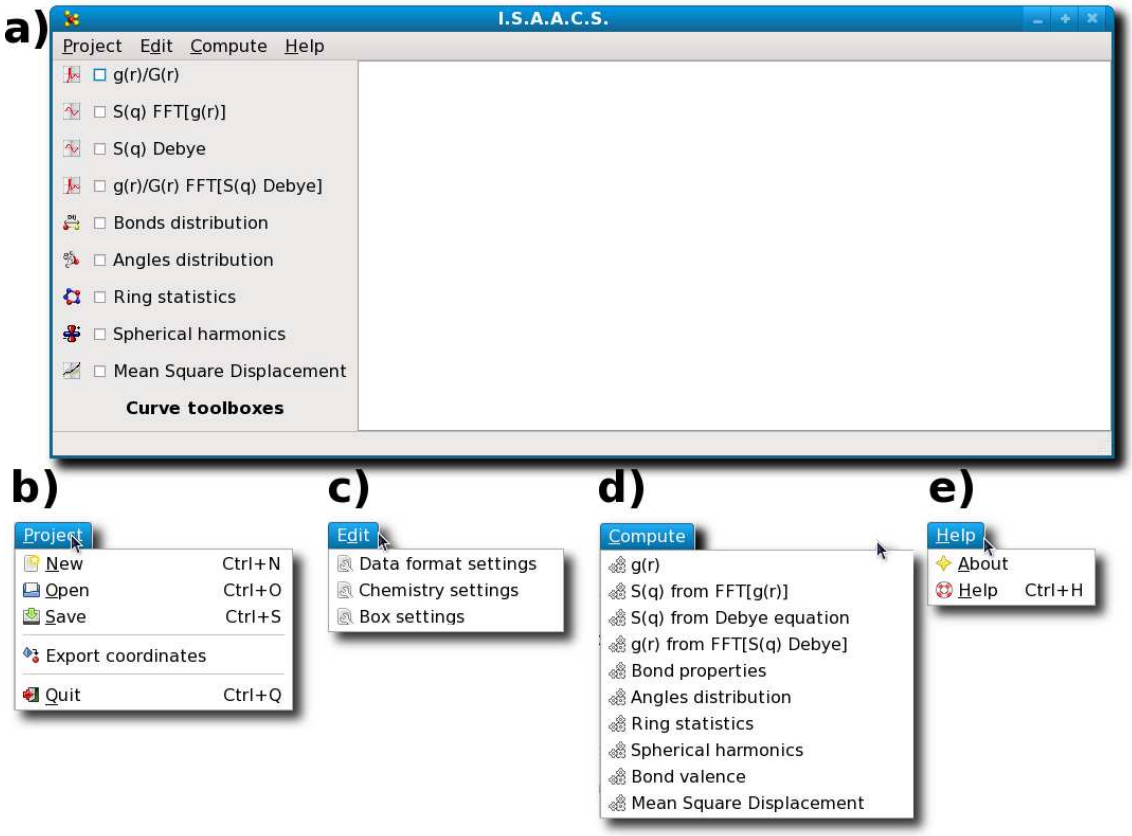


Figure 3.1 *Main interface of the I.S.A.A.C.S. program*

- Total number of atoms
- Number of chemical species
- Label and number of each atomic species

the input order, label and number of each atomic species will be the one assumed when reading the coordinates from the trajectory file.

The coordinates of atoms from a 3D model can be in any of the widely used formats listed in table [Tab. 3.2]:

I.S.A.A.C.S. can output an already imported 3D model in XYZ format [3] in either of the atomic coordinates formats presented in table [Tab. 3.2].

1. Cartesians
2. Atomic units (in input for CPMD trajectory files only)
3. Fractional with the center of the model box at (0,0,0)
4. Fractional with one of the corners of the model box at (0,0,0)

Table 3.2 *Formats of atomic coordinates imported and exported by I.S.A.A.C.S..*

3.2 What can be computed by I.S.A.A.C.S. ?

I.S.A.A.C.S. can compute the following important structural characteristics of a 3D structure model:

- Radial distribution functions g(r) (RDFs) [7] including °:
  - Total RDFs for neutrons and X-rays.
  - Partial RDFs.
  - Bhatia-Thornton RDFs [8]

  ° Radial distribution functions can be computed by i) direct real space calculation and/or ii) Fourier transforming of the structure factor calculated using the Debye formalism [9]
- Structure factors S(q) [9] including °°:
  - Total structure factors S(q) for neutrons and X-rays.
  - Total Q(q) [9, 10] for neutrons and X-rays.
  - Partial S(q):
    - Faber-Ziman [11] partial S(q)
    - Ashcroft-Langreth [12–14] partial S(q)
    - Bhatia-Thornton [15] partial S(q)

  °° Structure factors can be computed by i) Fourier transforming of the radial distribution functions and/or ii) using the Debye formalism [9]
- Interatomic bond properties
  - Coordination numbers
  - Atomic near neighbor distribution
  - Fraction of links between tetrahedra
  - Fraction of tetrahedral units
  - Bond lengths distribution for the first coordination sphere

- Distribution of Bond angles
- Distribution of Dihedral angles
- Ring statistics
  According several definitions:
  - All closed paths (no rules)
  - King's rings [16, 17]
  - Guttman's rings [18]
  - Primitive rings [19, 20] (or Irreducible [21])
  - Strong rings [19, 20]

  Also included are options for:
  - Possibility to look only for ABAB rings
  - Possibility to exclude rings with homopolar bonds (A-A or B-B) from the analysis

  Ring statistics is presented according to the R.I.N.G.S. method [22].
- Spherical harmonics invariant, Q_l, as local atomic ordering symmetry identifiers [23]
  - Average Q_l for each chemical species
  - Average Q_l for a user specified structural unit
- Bond valence sums [24–26]
  - Average bond valence for each chemical species
  - Average bond valence for a user specified structural unit
- Mean Square Displacement of atoms (MSD)
  - Atomic species MSD
  - Directional MSD (x, y, z, xy, xz, yz)
  - Drift of the center of mass

Running I.S.A.A.C.S.

A set of structural characteristics for a 3D model of silica glass computed by I.S.A.A.C.S. is shown below as an example of the program utilization and output. The model has been constructed by reverse Monte Carlo simulations guided by high-energy x-ray diffraction data for silica glass [27]. It consists of 2000 oxygen and 1000 silicon atoms inside a box of dimensions 35.6621 Å. The model is available in the example files distributed with the program, details are provided on the web site at:
http://isaacs.sourceforge.net/ex.html
http://people.cst.cmich.edu/petko1vg/isaacs/ex.html

4.1 Set up a calculation

When starting I.S.A.A.C.S. the 'Compute menu' [Fig. 3.1-d] is not activated. It is indeed mandatory to import a structure model before being able to run any calculation on it.
Thus the user has the choice:

- To open an existing project, i.e. an '*.ipf' I.S.A.A.C.S. project file using The 'Project menu' => *O*pen [Fig. 3.1-b] button.
- To create a new project using the 'Project menu' => *N*ew [Fig. 3.1-b] button.

4.1.1 The 'Project settings' windows

The *O*pen [Fig. 3.1-b] and the *N*ew [Fig. 3.1-b] buttons will open the 'Project settings' window [Fig. 4.1] which allows to import and set up a structure model for the calculation. The different tabs of the 'Project settings' window are also accessible through the 'Edit menu' [Fig. 3.1-c].

In the 'Data format settings' tab [Fig. 4.1] the user can select the format of the file with the atomic coordinates [Fig. 4.2] and then can open that file.

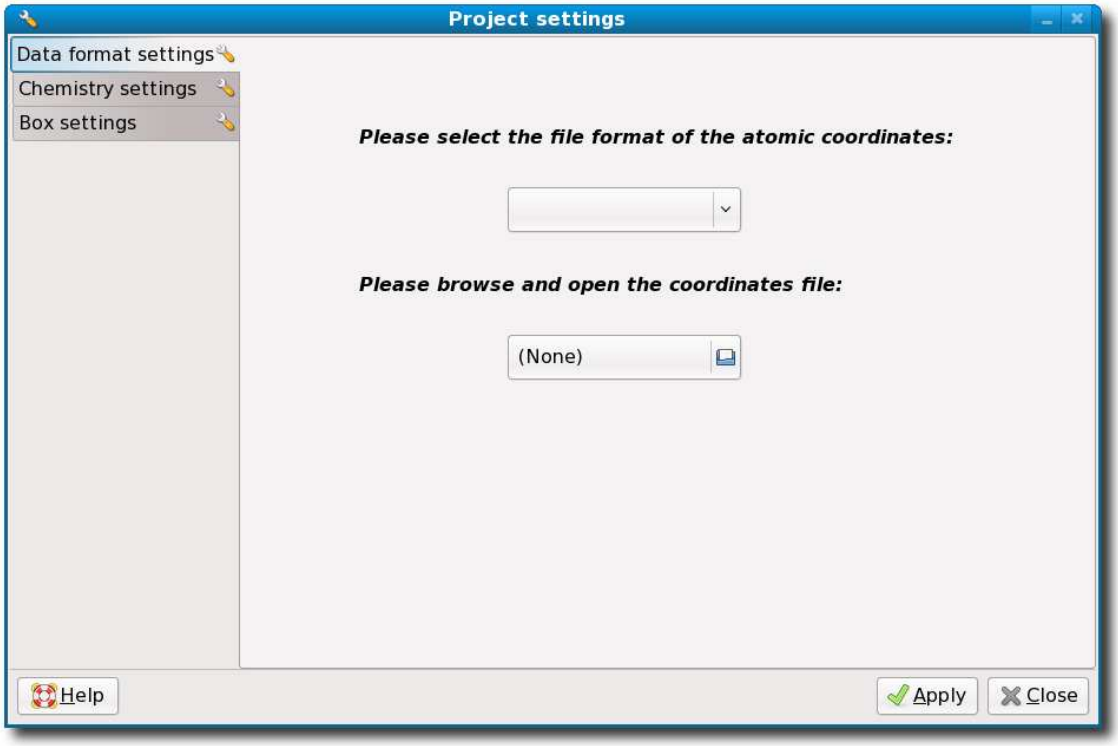


Figure 4.1 *The 'Project settings' window of the I.S.A.A.C.S. program*

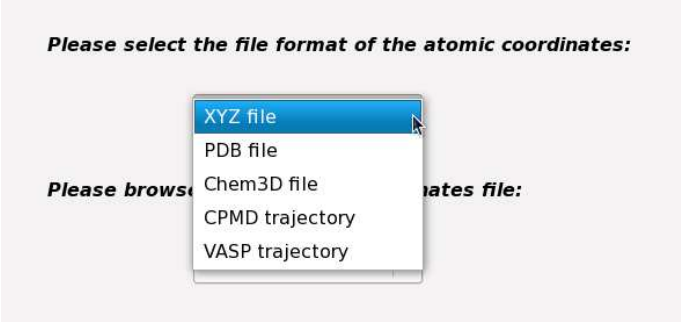


Figure 4.2 *Selection of the file format of atomic coordinates in the I.S.A.A.C.S. program*

If the structure model file is read successfully then the user can adjust the different parameters required to run calculation using the different tabs of the 'Project settings' windows [Fig. 4.1].

The 'Chemistry settings' tab of the 'Project settings' window [Fig. 4.3] allows to access the information regarding the chemistry of the structure model and to check and/or modify the properties of each chemical species.

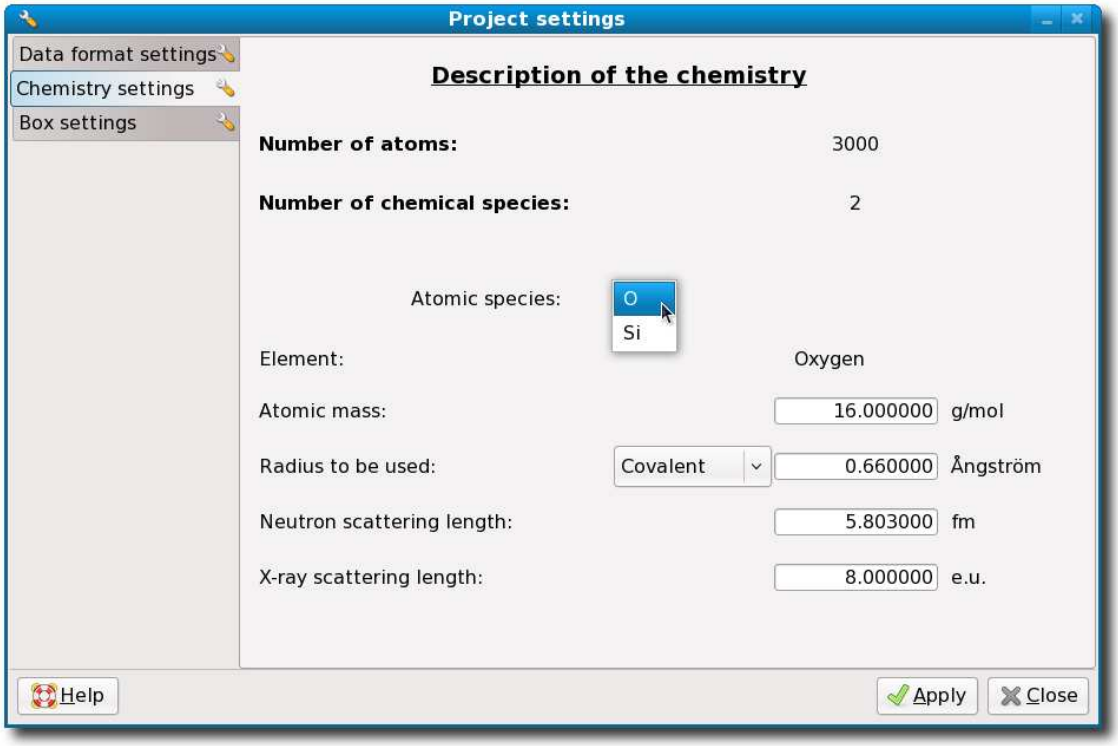


Figure 4.3 *The 'Chemistry settings' tab of the 'Project settings' window in the I.S.A.A.C.S. program*

The user has the freedom to input his own data, however it is also possible to use the database provided by the I.S.A.A.C.S. program using for example one of the different atomic radii (see appendix [A. A] for details).

The 'Box settings' tab of the 'Project settings' window [Fig. 4.4] allows to access the information regarding the periodicity of the system: description of the simulation box, the type (i.e. Cartesian or fractional) of the coordinates and the periodic boundary conditions.

The simulation box is described using the 'A, B, C, α, β, γ' set of parameters [Fig. 4.4], or, alliteratively, by the coordinates of the edges of the simulation box using the 'Lattice vector properties' window [Fig. 4.5] accessible from the 'Box settings' tab of the 'Project settings'

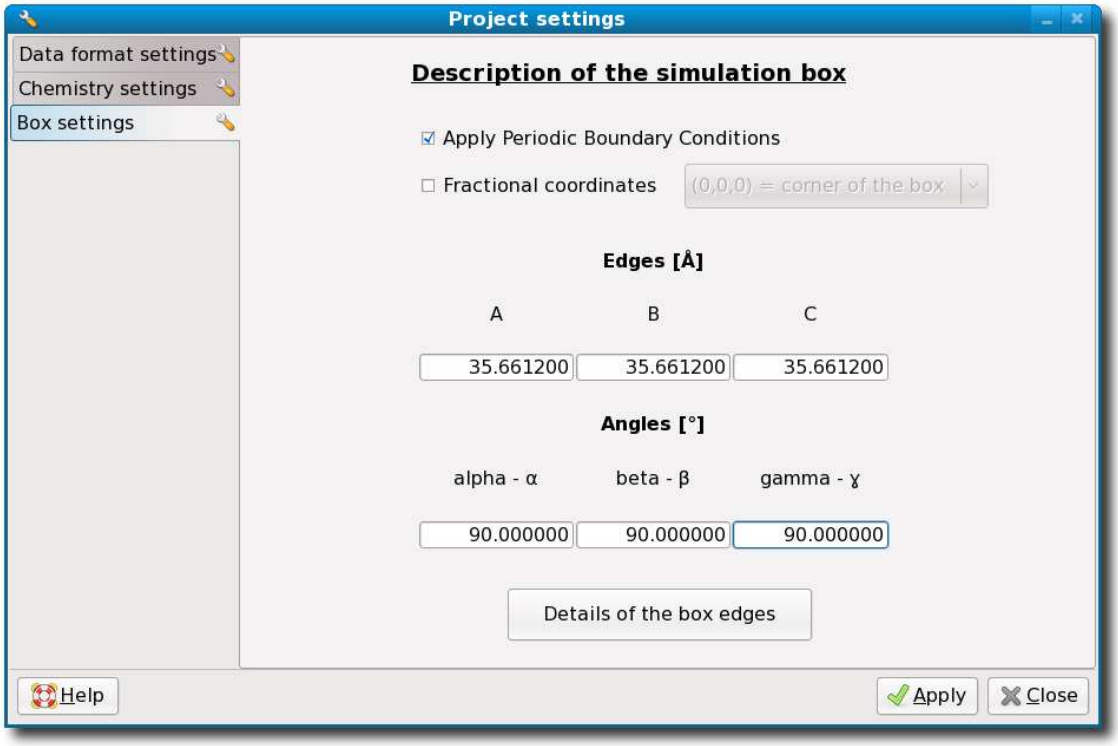


Figure 4.4 *The 'Box settings' tab of the 'Project settings' window in the I.S.A.A.C.S. program*

windows through the 'Details of the box edges' button [Fig. 4.4].

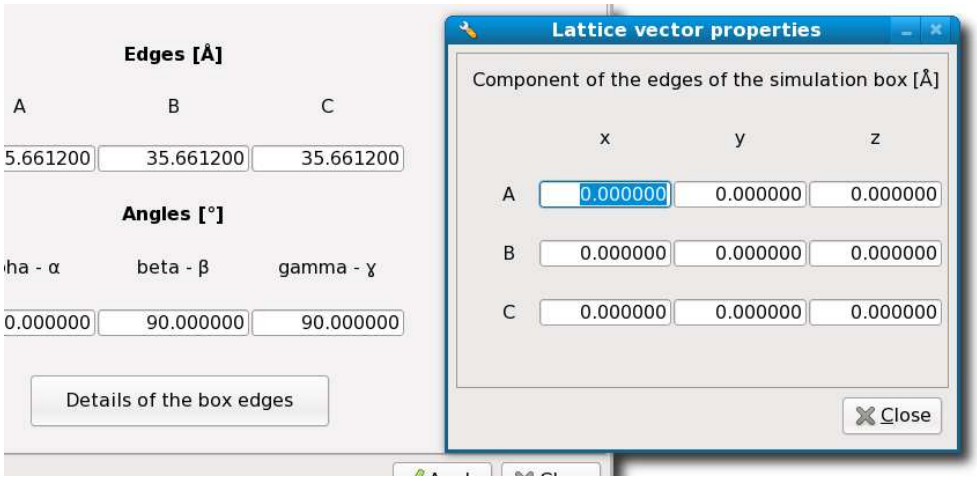


Figure 4.5 *The 'Lattice vector properties' window in the I.S.A.A.C.S. program*

4.1.2 Help to set up a project

A basic help [Fig. 4.6] describing the information of each tab of the 'Project settings' window can be opened using the 'Help' button of the 'Project settings' window [Fig. 4.1, 4.3, 4.4].

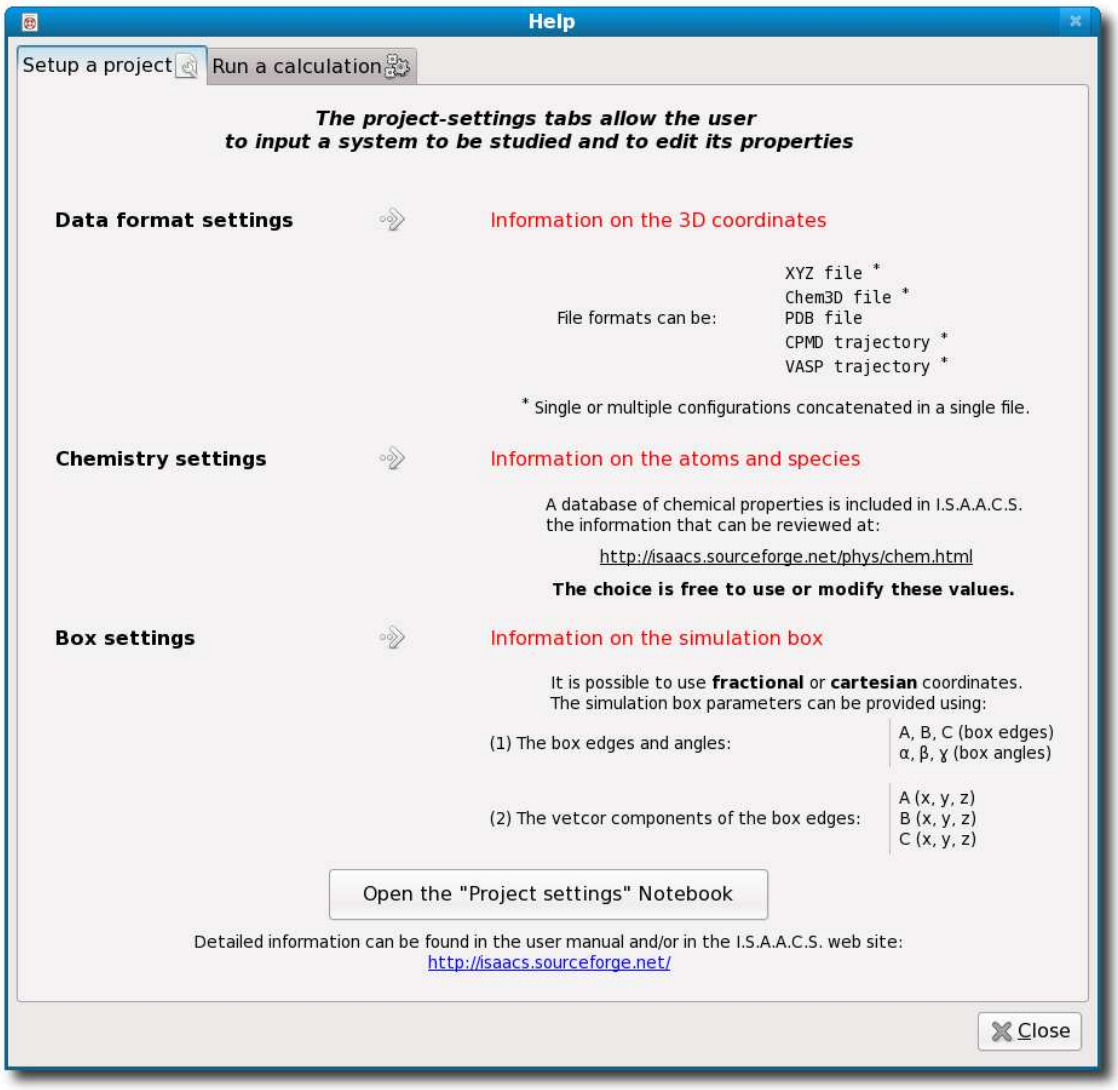


Figure 4.6 *The 'Help' menu of the 'Project settings' window in the I.S.A.A.C.S. program*

This help window [Fig. 4.6] is also accessible from '*H*elp' menu of the main I.S.A.A.C.S. window [Fig. 3.1-e].

4.2 Running a calculation

The first step to run a calculation is to validate the project by clicking on the 'Apply' button of the 'Project settings' window [Fig. 4.1, 4.3, 4.4]. First the program verifies that calculations can be done using the information provided by the user in the fields of the different tabs of the 'Project settings' window.
If the project is validated then calculation buttons in the 'Compute menu' [Fig. 4.7] of main window of the program become active and some basic information on the system is displayed in the main I.S.A.A.C.S. window [Fig. 4.8]:

- Name of the file with the atomic coordinates
- Format of that file
- Basic chemistry information
  - Total number of atoms
  - Number of chemical species
  - Volume and density of the model system
  - Free volume
  - Empirical chemical formula
  - Concentration, fraction and number density of each chemical species

Please notice that all calculations are not immediately accessible after validating the project. Indeed some of these calculations require options to be applied and/or others calculation to be run prior:

- The 'g(r)', and 'S(q)' calculations require to enter parameters to describe the size of the simulation box.
- The 'S(q) from FFT[g(r)]' calculation requires the 'g(r)' calculation to be completed.
- The 'g(r) from FFT[S(q) Debye]' calculation requires the 'S(q) from Debye Eq.' calculation to be completed.
- The 'Spherical harmonics' and 'Bond valence' calculations require the 'Bond properties' calculation to be completed.
- The 'Mean Square Displacement' calculation requires to have multiple atomic configurations to be present in the atomic configuration data file.

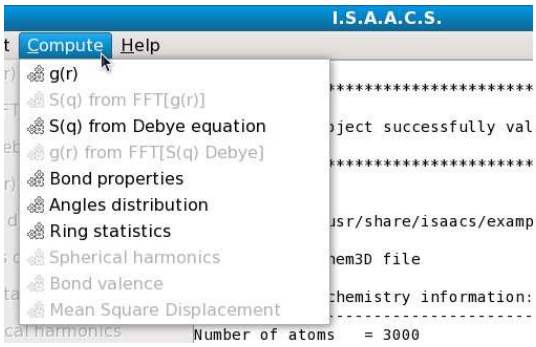


Figure 4.7 *The 'Compute' menu with some calculations activated in the I.S.A.A.C.S. program*

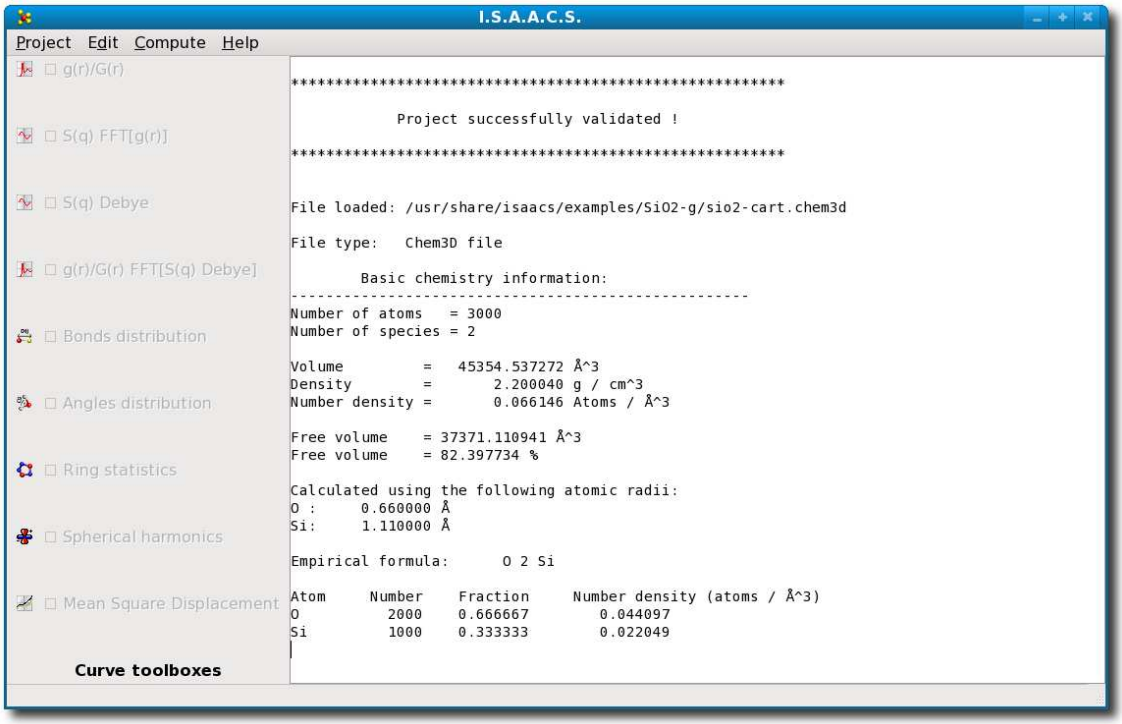


Figure 4.8 *The main interface of the I.S.A.A.C.S. program after the validation of the project*

To compute each of the characteristics shown below the relevant dialog/menu boxes have been used to set up the values of the program parameters required to run a structure analysis. For example, figure [Fig. 4.9] presents the dialog boxes used to control the calculations of the radial distribution function [Fig. 4.9-a] and the ring statistics [Fig. 4.9-b] for silica glass.
For example once the 3D structure is read and validated by I.S.A.A.C.S. the only parameter required to compute the RDF is the number of steps in real space, δr [Fig. 4.9-a] (after the calculation an optional smooth of the results is also possible). In the second case [Fig. 4.9-b] a few more control parameters (definition of a ring, chemical species used to initiate the search, maximum size of a ring, maximum number of rings per node and description of the chemical bonds) are needed.

To guide the calculation sequence I.S.A.A.C.S. requires users to supply control parameters. Information about these parameters can be obtained through a help button in the dialog box [Fig. 4.9-a,b] or through the help menu of the main window of the program [Fig. 3.1-e] that will open the help window of the program [Fig. 4.10].

The next pages illustrate the different dialog boxes used to control the calculations as well as the help associated to these boxes and displayed in the 'Help' window of the I.S.A.A.C.S. program.

Important: The web sites dedicated to the I.S.A.A.C.S. program provide a detailed theoretical background of the structural analysis proposed in the software:
http://isaacs.sourceforge.net/
http://people.cst.cmich.edu/petko1vg/isaacs/

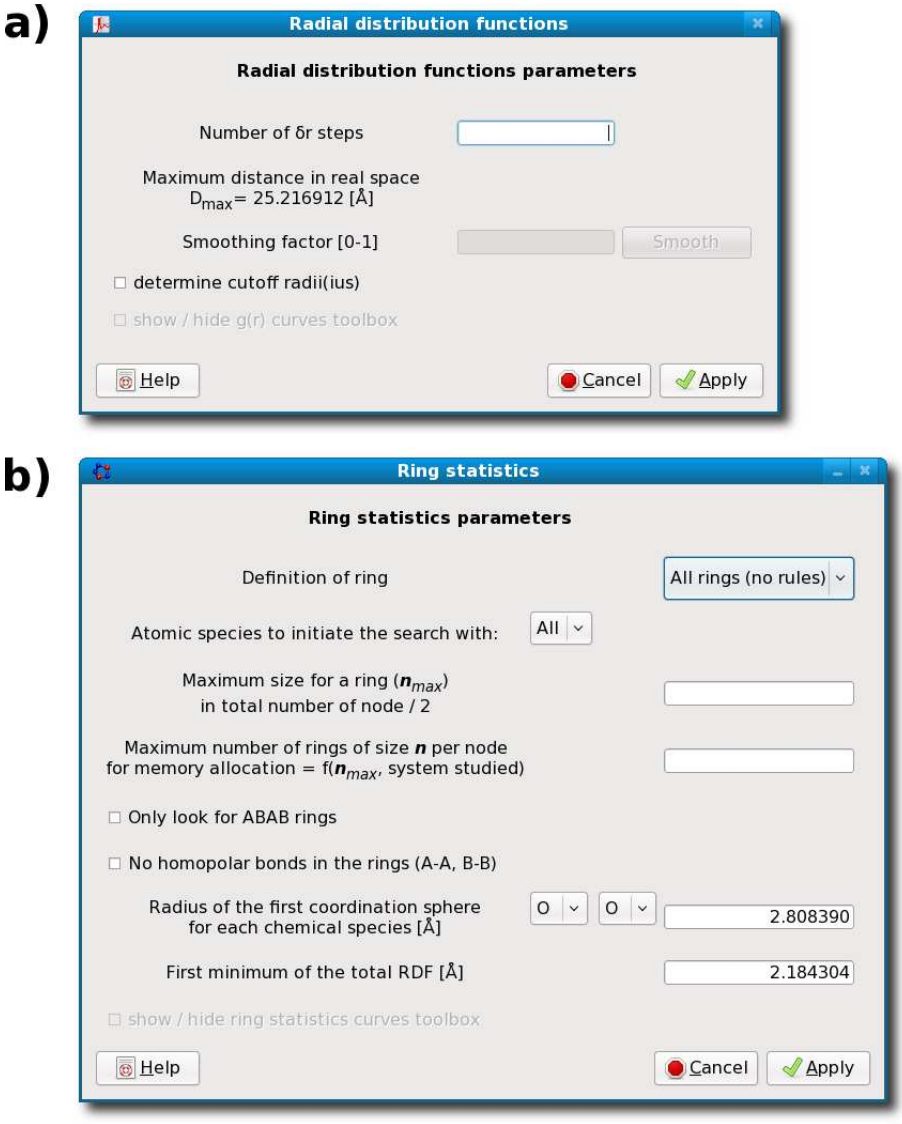


Figure 4.9 *Dialog boxes controlling a) the calculation of radial distribution functions and b) ring statistics in the I.S.A.A.C.S. program.*

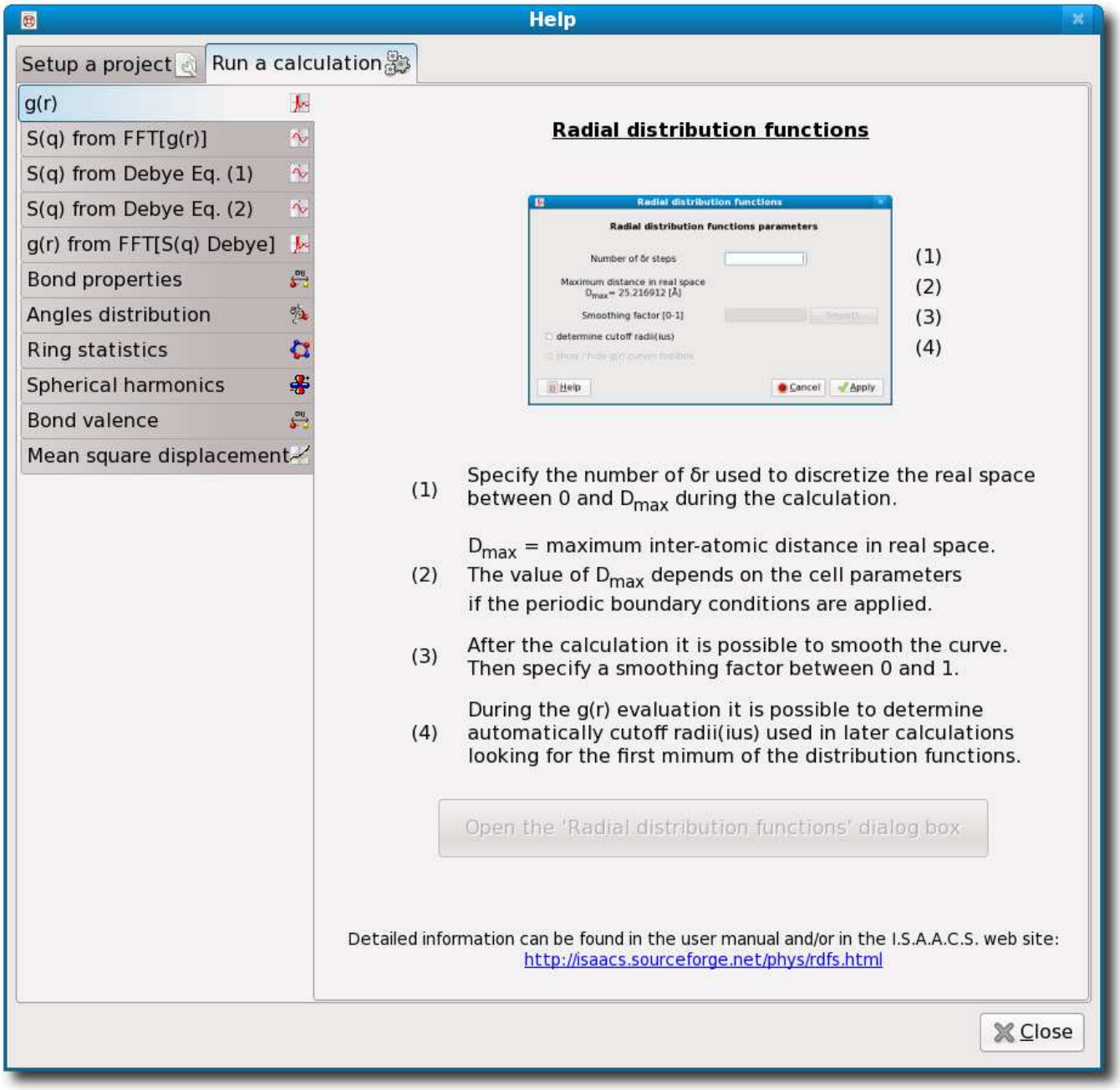


Figure 4.10 *The help window offers a user support to the different calculations accessible in the I.S.A.A.C.S. program. The 1st tab provides information regarding the fields of the dialog box controlling the calculation of the radial distribution functions.*

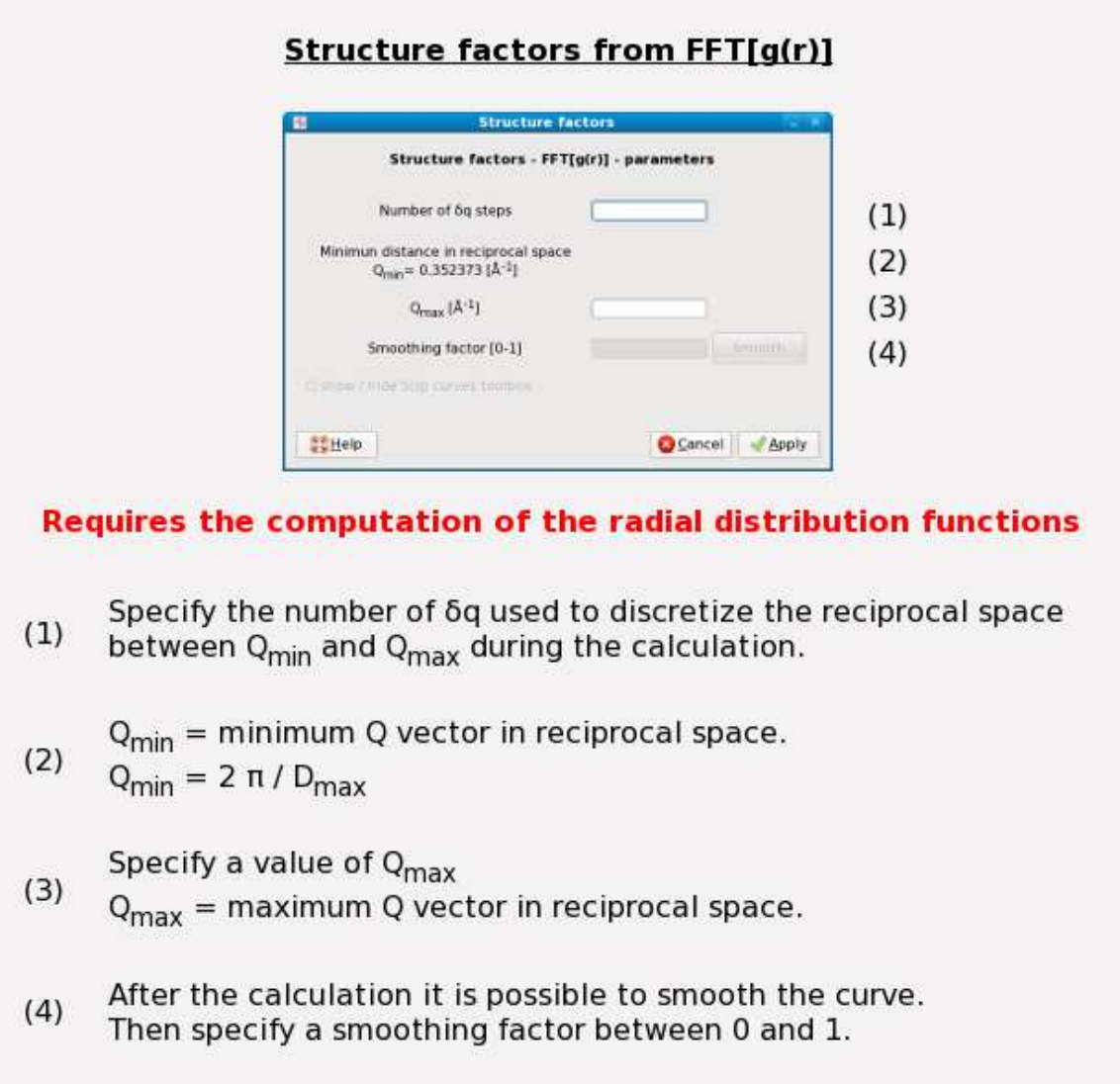


Figure 4.11 *The 2nd tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the structure factor from the Fourier transform of the radial distribution functions in the I.S.A.A.C.S. program.*

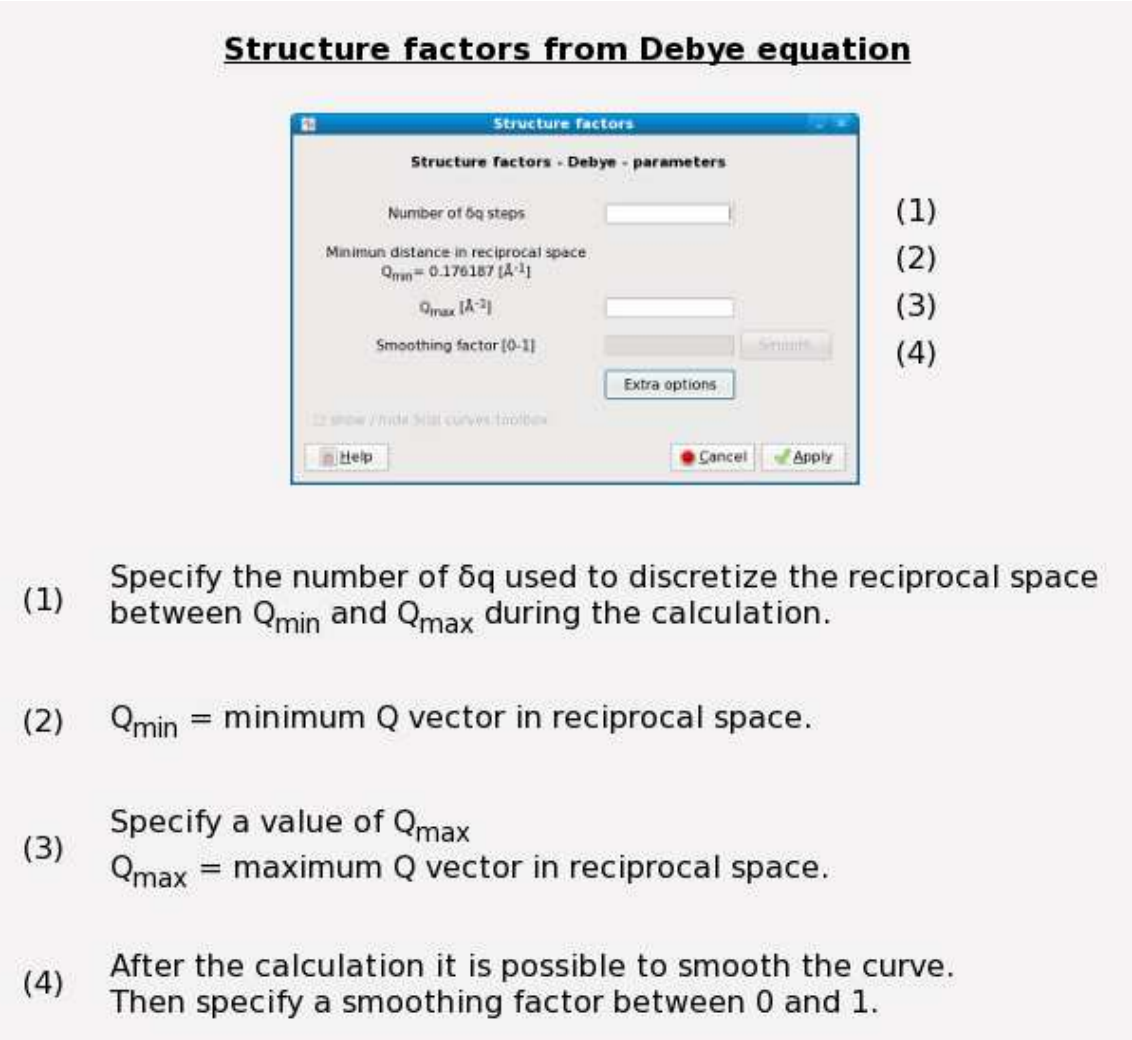


Figure 4.12 *The 3rd tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the structure factor from the Debye equation in the I.S.A.A.C.S. program.*

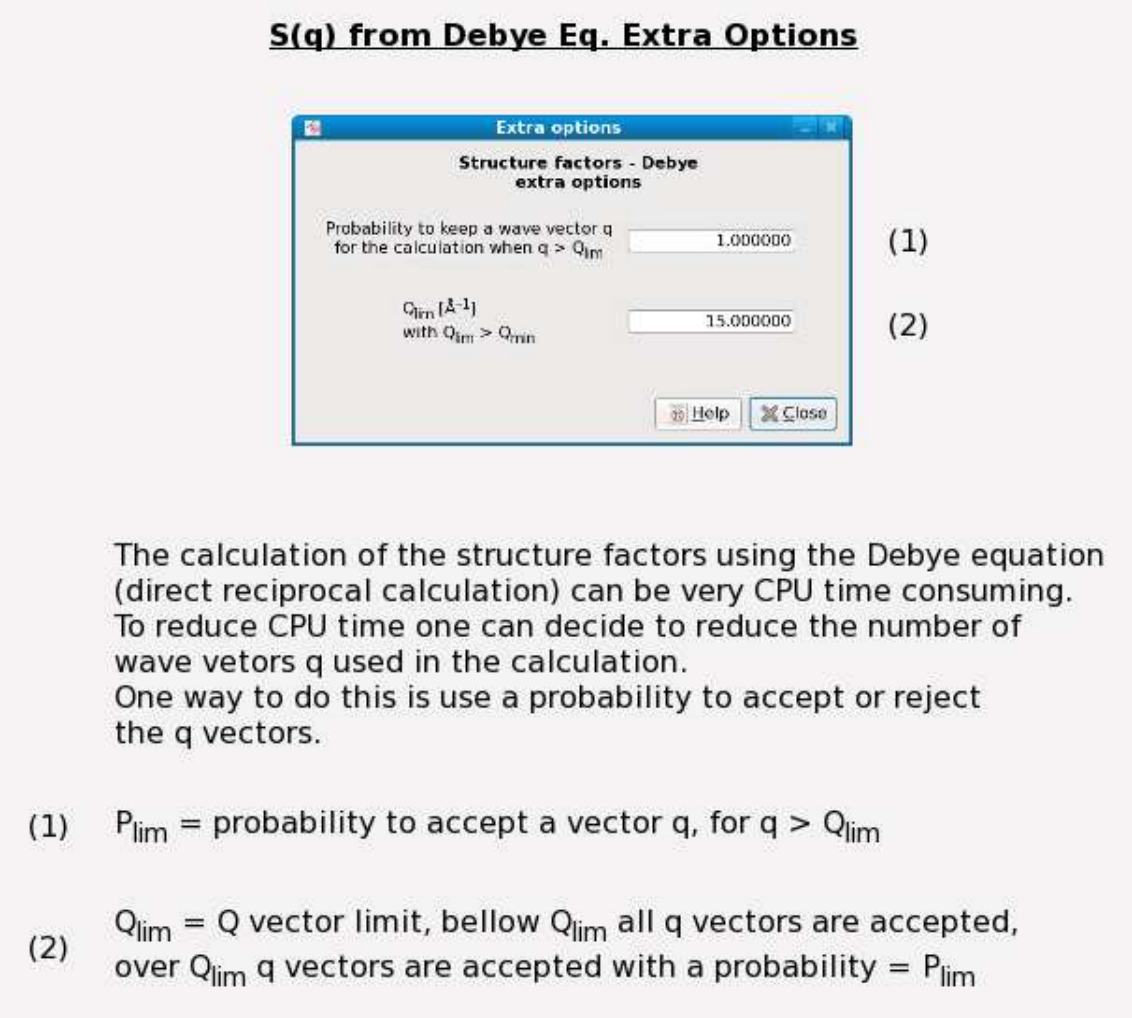


Figure 4.13 *The 4th tab of the help window provides information regarding the fields of the dialog box controlling the extra options for the calculation of the structure factor from the Debye equation in the I.S.A.A.C.S. program.*

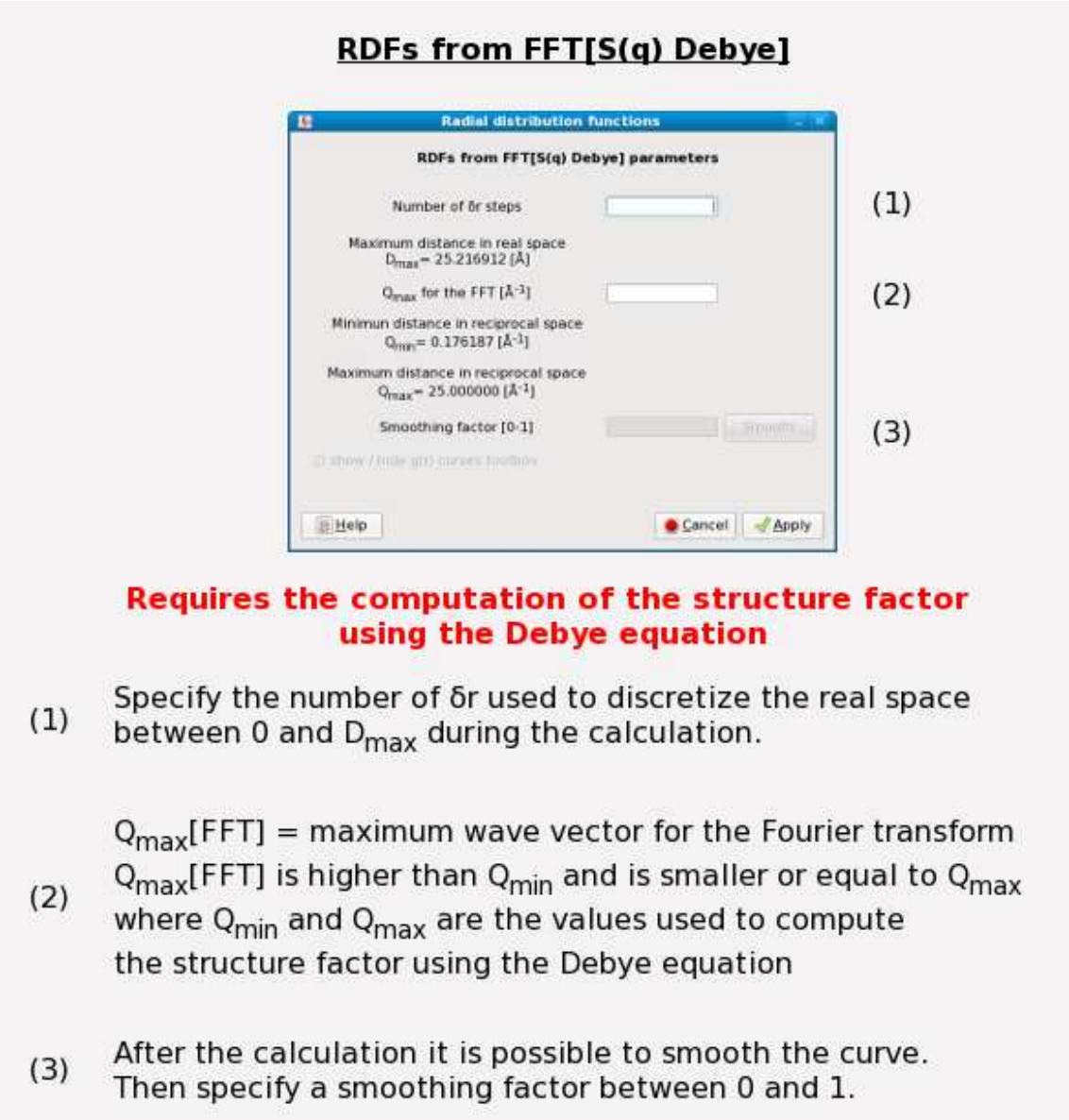


Figure 4.14 *The 5th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the radial distribution functions from the Fourier transform of the structure factor calculated using the Debye equation.*

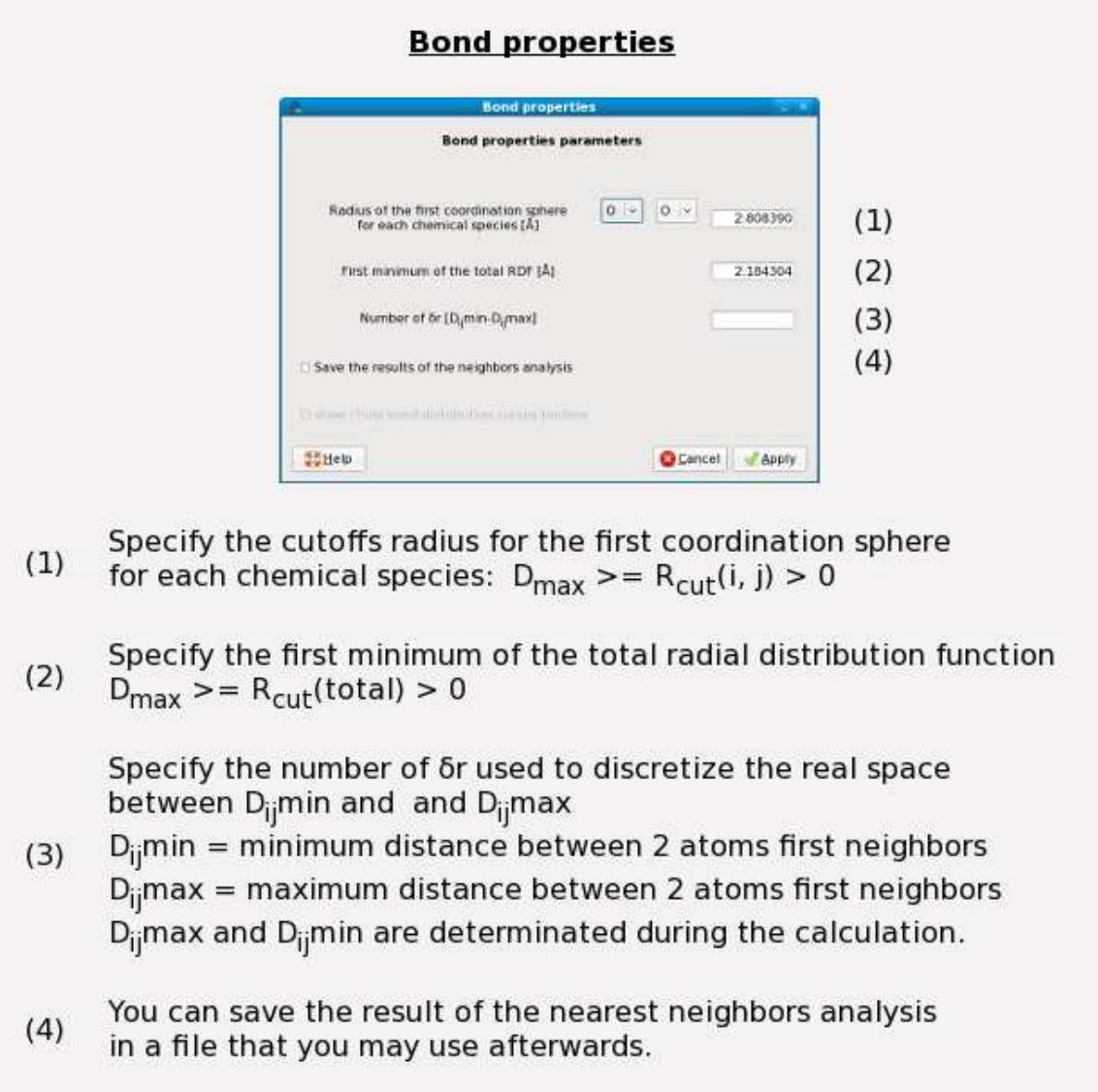


Figure 4.15 *The 6th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the bond properties in the I.S.A.A.C.S. program.*

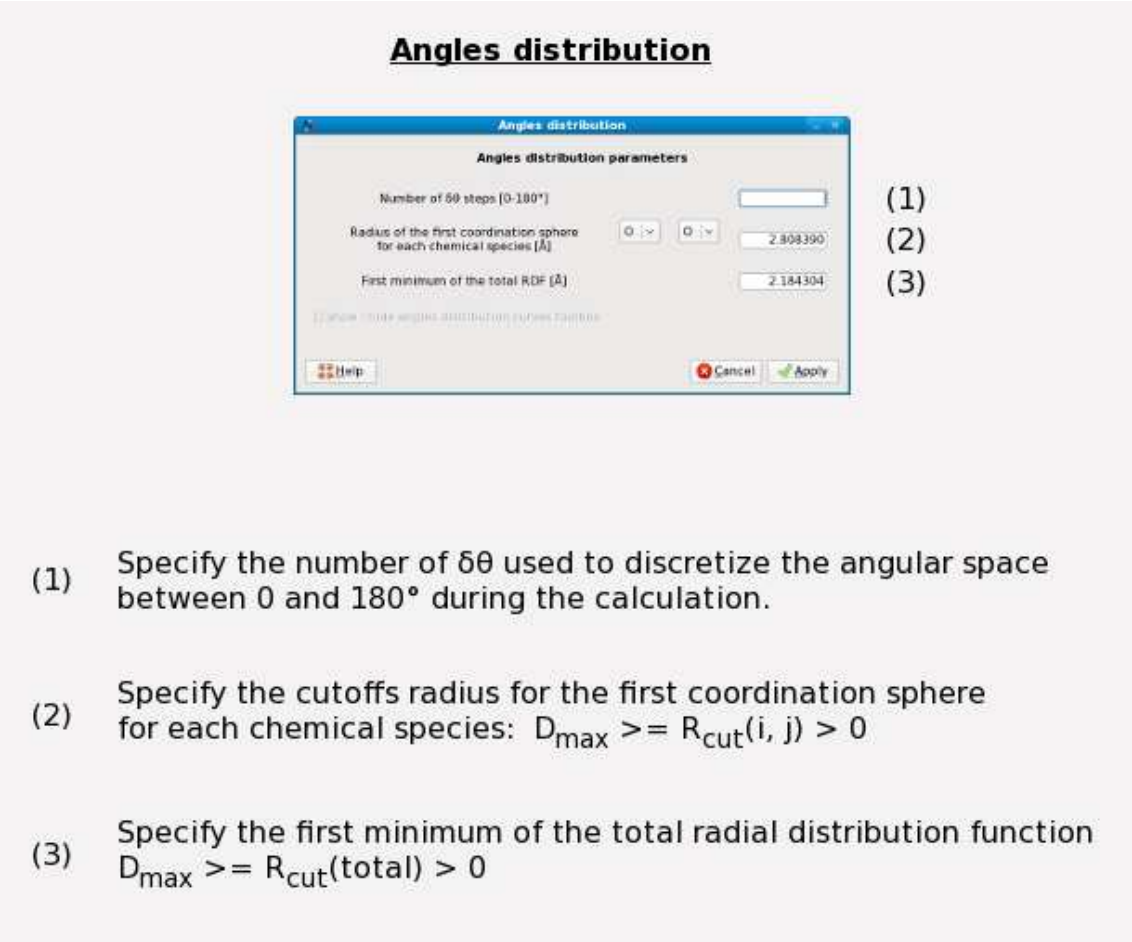


Figure 4.16 *The 7th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the angles distributions in the I.S.A.A.C.S. program.*

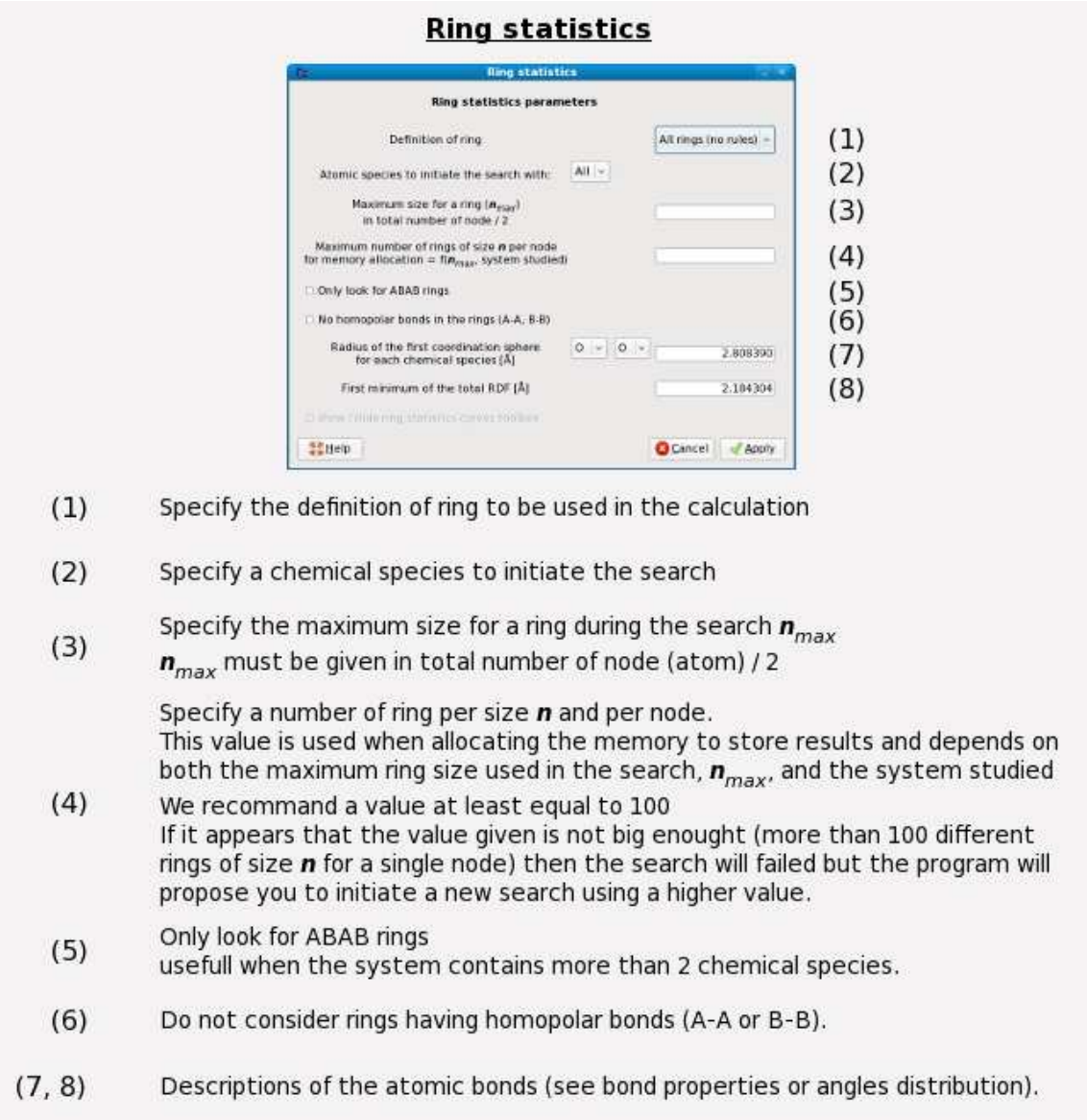


Figure 4.17 *The 8th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the ring statistics in the I.S.A.A.C.S. program.*

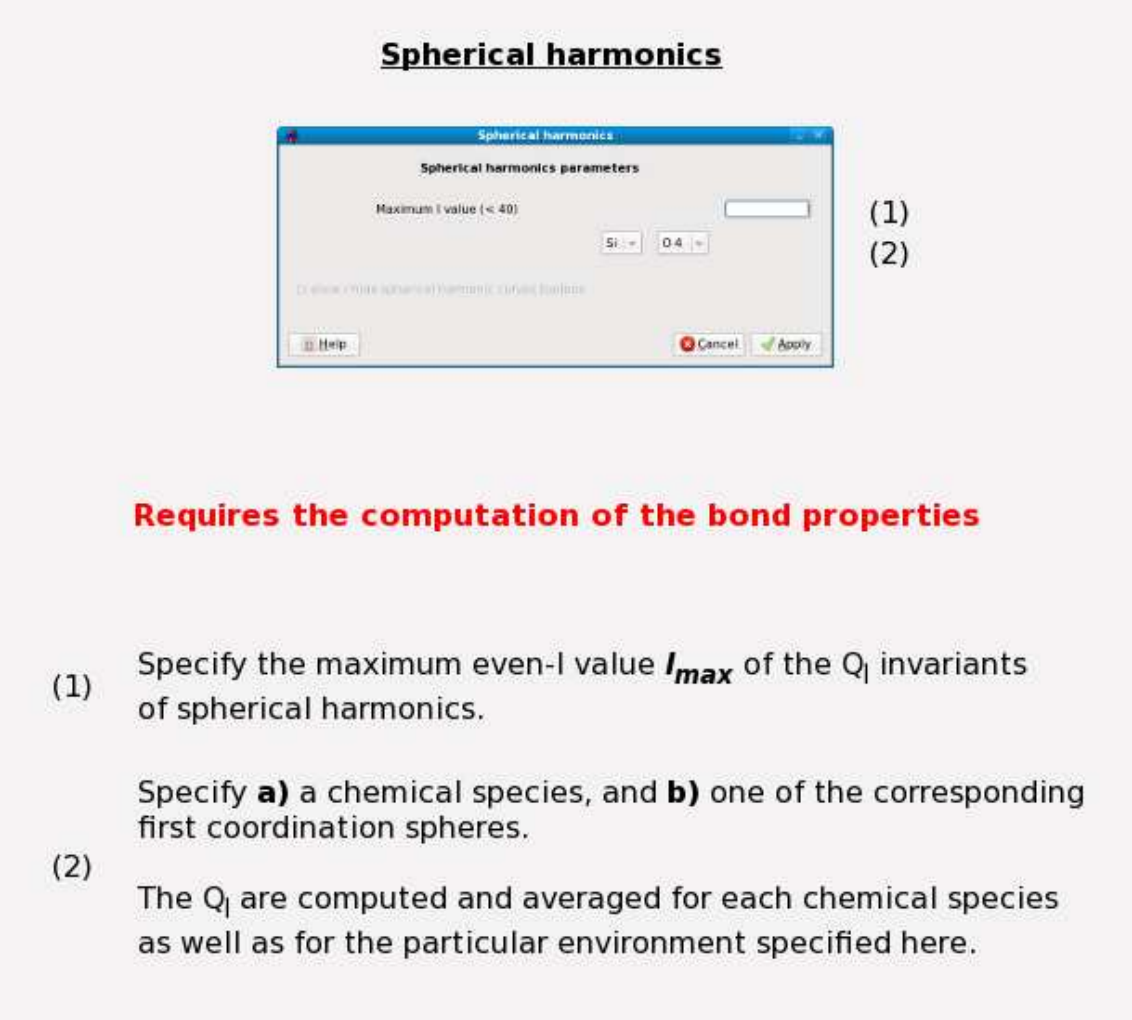


Figure 4.18 *The 9th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the local order parameters from the invariants of spherical harmonics in the I.S.A.A.C.S. program.*

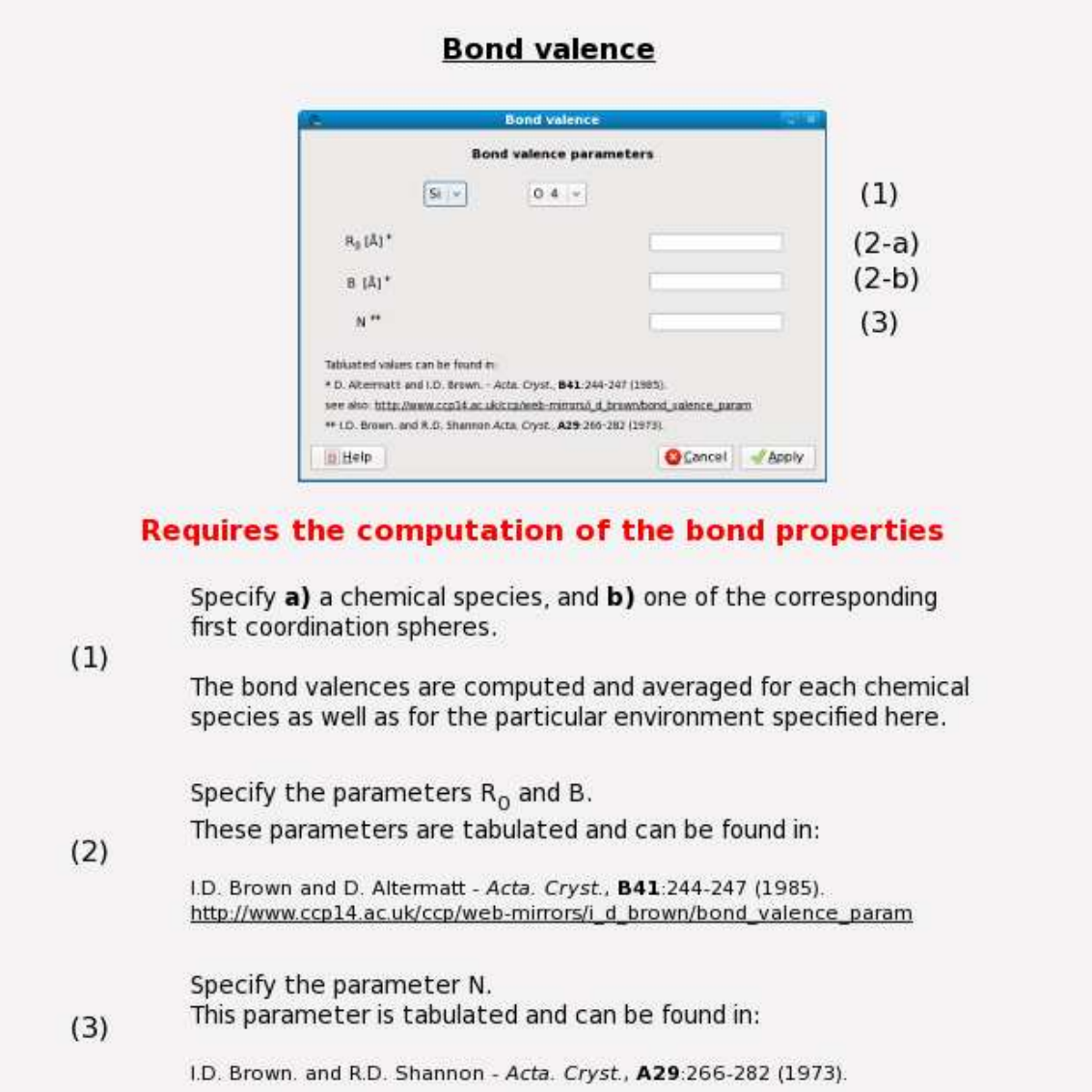


Figure 4.19 *The 10th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the bond valence sums in the I.S.A.A.C.S. program.*

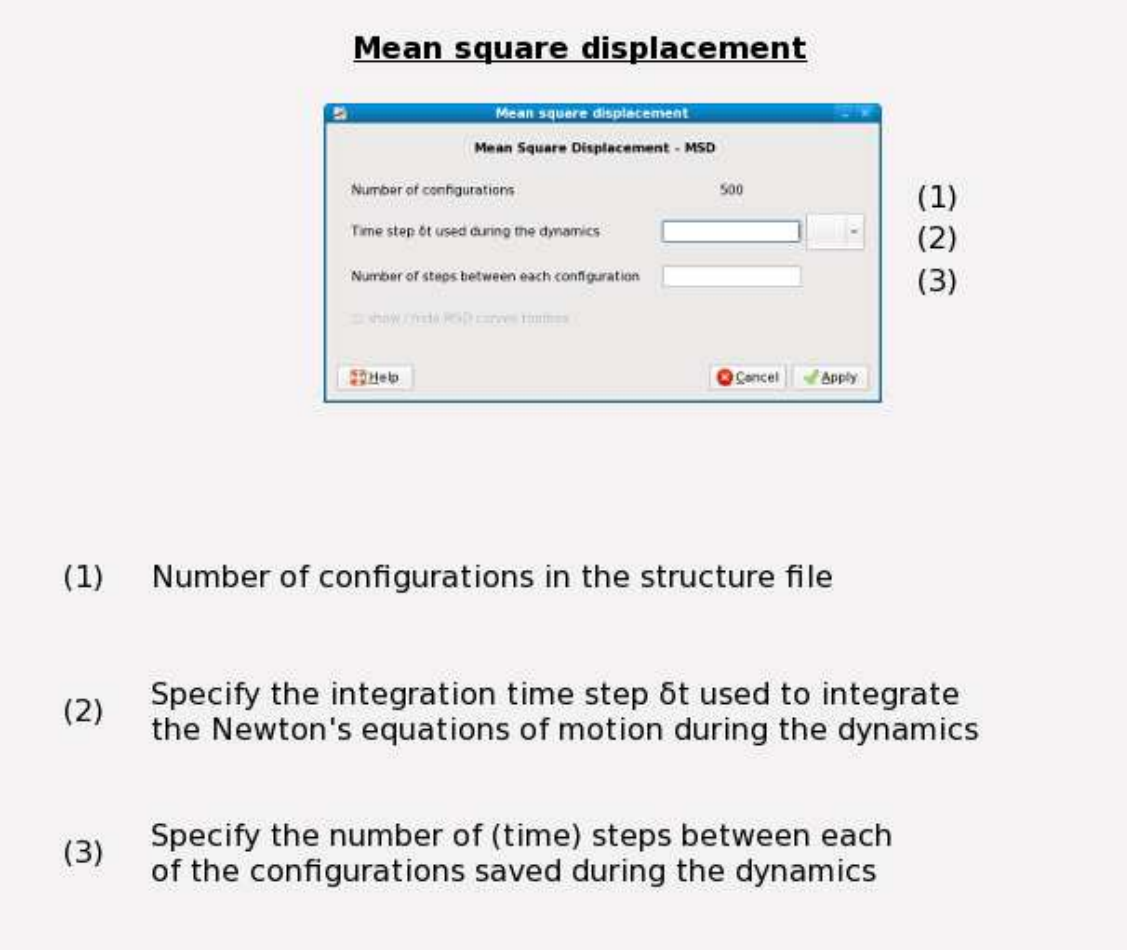


Figure 4.20 *The 11th tab of the help window provides information regarding the fields of the dialog box controlling the calculation of the mean square displacement in the I.S.A.A.C.S. program.*

4.3 Visualisation of the results of the calculations

When a particular structural characteristic is computed it can be directly displayed in the main I.S.A.A.C.S. windows [Fig. 3.1]. In addition the visualization mode of most of the computed characteristics can be controlled via interactive menus such as the one presented in figure [Fig. 4.21-a].
When a button in an interaction menu is activated [Fig. 4.21-a] the corresponding result is instantaneously displayed as a smooth curve or a histogram [Fig. 4.21-b, 4.22, 4.23, 4.24] depending on the nature of the computed structural characteristic.

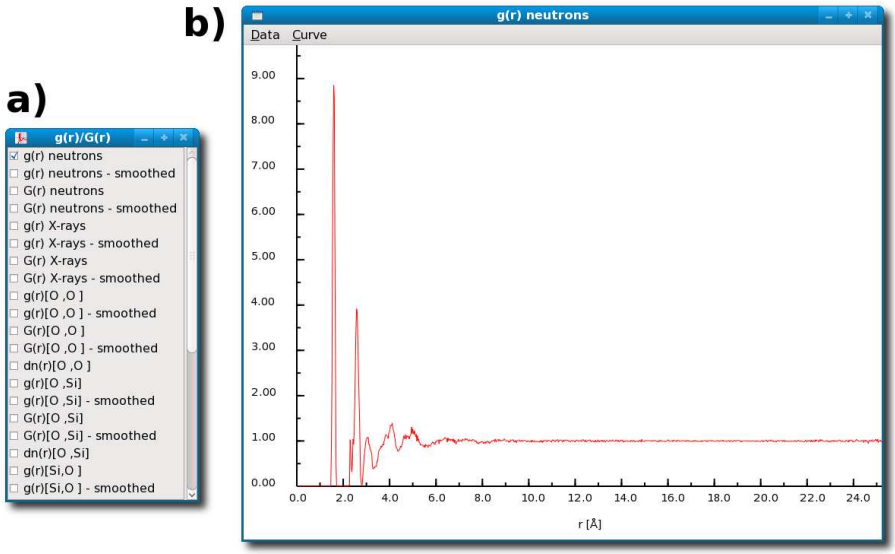


Figure 4.21 *Results of the calculation of the radial distribution functions for glassy SiO_2 by the I.S.A.A.C.S. program: the interaction box a) allows to display the computed RDFs, b) shows the total radial distribution function obtained for SiO_2 glass in case of neutrons diffraction. The RDF shows a first sharp peak position at 1.6 Å reflecting the presence of well defined $Si(O)_4$ tetrahedra in the glass. The second RDF peak reflects the correlations between O atoms sitting on the vertices of those tetrahedra.*

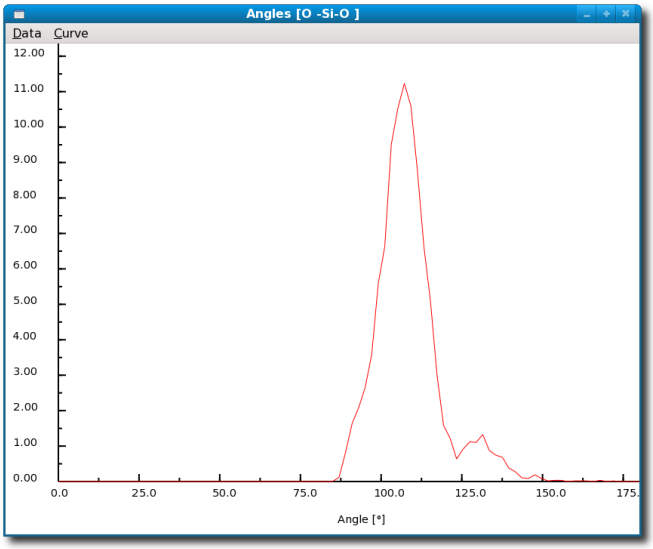


Figure 4.22 *Results from the calculation of the angle distribution for glassy SiO_2 by the I.S.A.A.C.S. program: the figure represents the distribution of the bond angles (O-Si-O) computed and immediately displayed using I.S.A.A.C.S.. In SiO_2 glass the O-Si-O angle distribution peaks at about 109° as may be expected for tetrahedral Si-O coordination.*

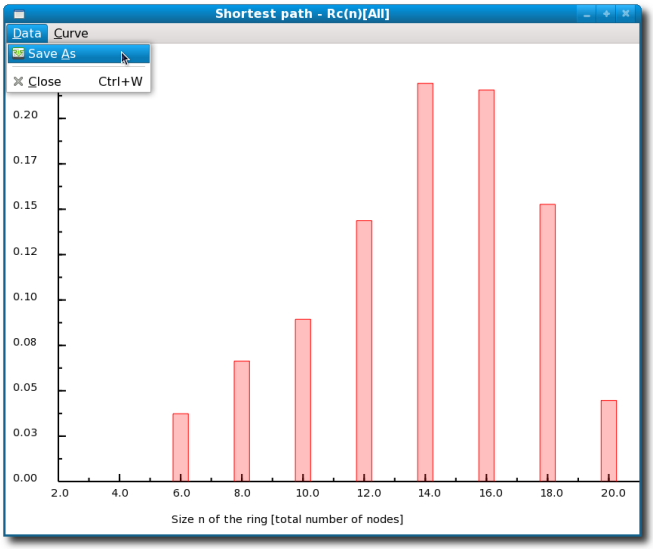


Figure 4.23 *Ring statistics for silica glass as computed by I.S.A.A.C.S.These results present the average number of rings per atom in the simulation box for a maximum ring size fixed to 20 atoms. Note if the number of rings is normalized per $Si(O)_4$ unit and not per atom the "ring size" would drop by a factor of two, i.e. the distribution would peak around rings made of seven $Si(O)_4$ units.*

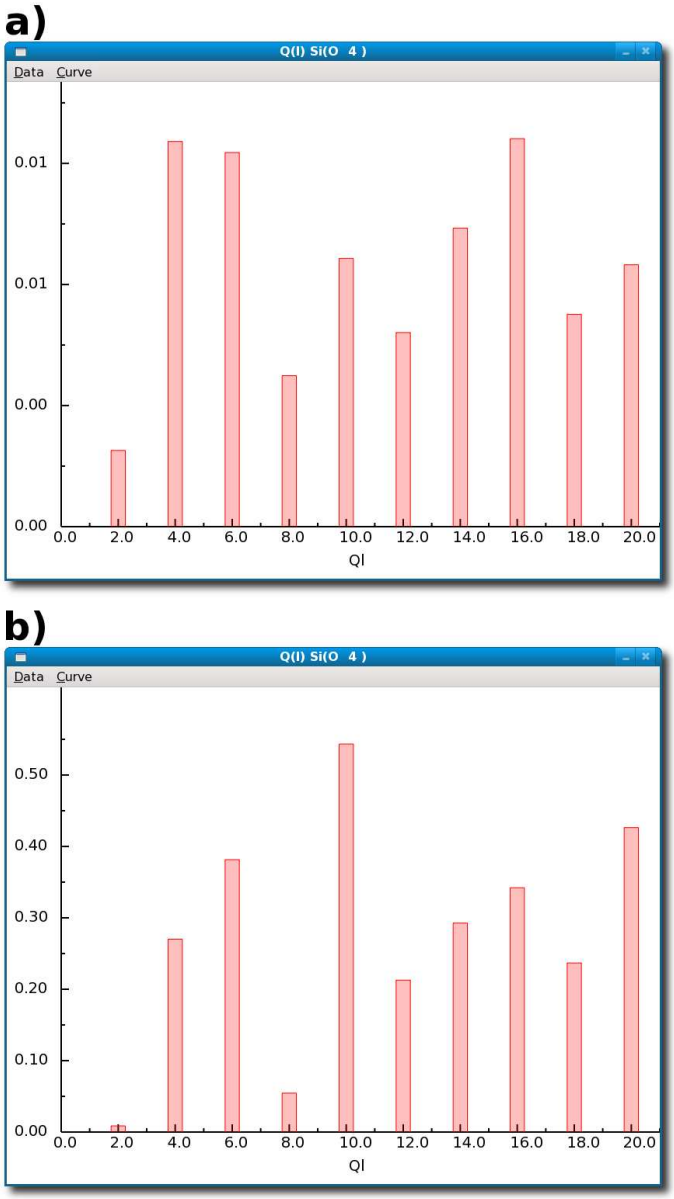


Figure 4.24 *Results of the calculation of the spherical harmonics as local order parameters for SiO_2 glass by the I.S.A.A.C.S. program: a) average Q_l computed for the $Si(O)_4$ environments (distorted tetrahedra) in glassy SiO_2, compared with b) average Q_l computed for the $Si(O)_4$ environments (ideal tetrahedra) in crystalline quartz-α. The two sets of Ql's follow a similar trend showing the similarity between the structural units in crystalline and glassy SiO_2. The differences are mostly quantitative and are due to the fact that the Si-O tetrahedra are somewhat distorted in the glass.*

4.4 Data plot edition

Since version 2.0, I.S.A.A.C.S. offers a curves editing tool which allows to configure the layout of the graphs showing result from the calculations. This data plot editing tool is accessible through the 'Curve' menu of any graph window [Fig. 4.25].

Figure 4.25 *The 'Curve' menu of the graph window in the I.S.A.A.C.S. program.*

Using the data plot editing tool [Fig. 4.26] it is possible to configure the layout of the selected graph, as well as the layout each data set which happen to be plotted on the graph [Fig. 4.26-a)], it is also possible to configure X and Y axis layout and/or position [Fig. 4.26-b)]. Furthermore depending on the calculation several data sets can be selected and plotted together with the main data set of the active window [Fig. 4.26-c)].
Data plot edition in the I.S.A.A.C.S. program is illustrated with the examples in [Fig. 4.27] and [Fig. 4.28].

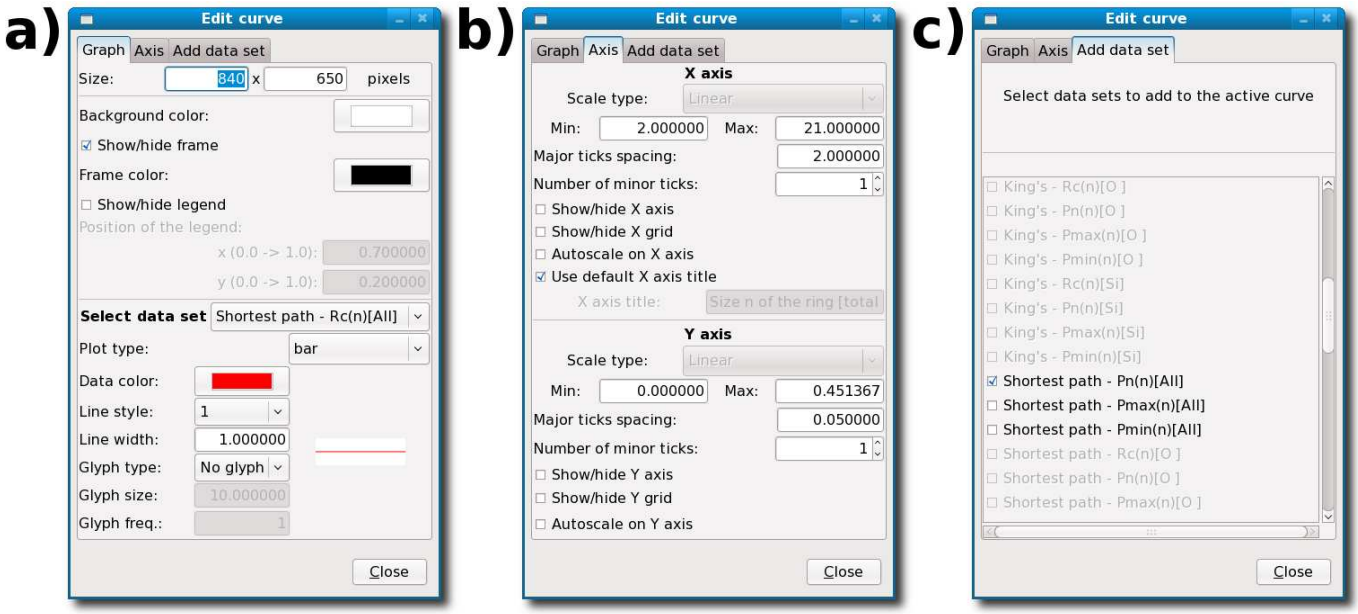


Figure 4.26 *The data plot editing tool box in the I.S.A.A.C.S. program.*

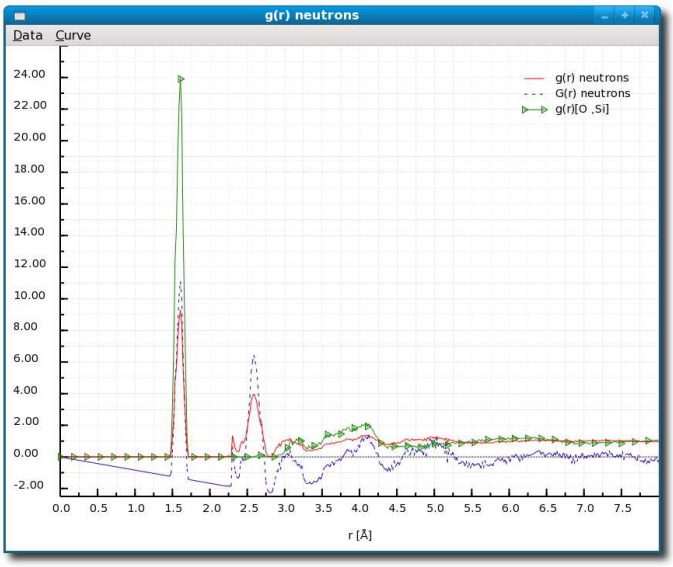


Figure 4.27 *Summary of the results of the calculation of the radial distribution functions for glassy SiO_2 by the I.S.A.A.C.S. program: total neutron $g(r)$ and $G(r)$ distribution functions (see Sec. 5.2 and 5.3 for details) as well as partial O-Si distribution function are displayed on the same graph, layout, legend and axis scales are configured*

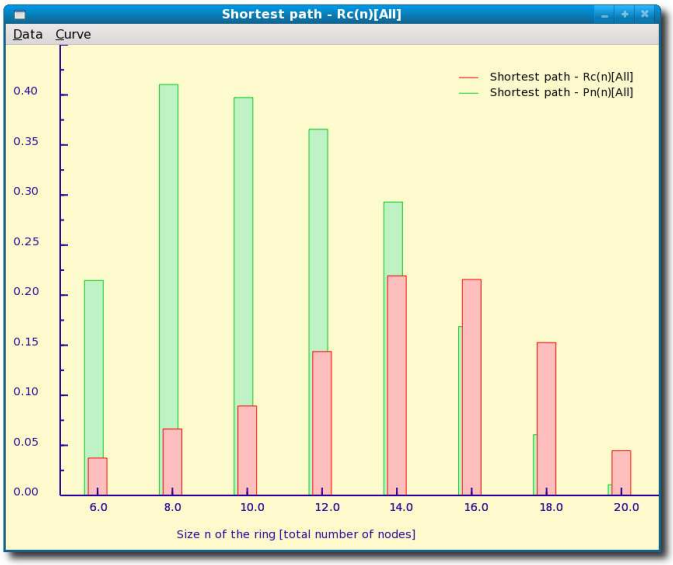


Figure 4.28 *Summary of the results of ring statistics for silica glass as computed by I.S.A.A.C.S.. R_c and P_n for shortest path analysis (see Sec. 5.5 for details) are displayed on the same graph, layout, legend and axis scales are configured.*

4.5 Exporting an image

Thanks to the 'Curve' menu [Fig. 4.25], plots can be exported from every graph window in the I.S.A.A.C.S. program. Images can be saved in the 'PNG' *Portable Network Graphics*, 'PDF' *Portable Document File* and 'SVG' *Scalable Vector Graphics formats*.

4.6 Saving the data

Results computed by I.S.A.A.C.S. can be easily saved using the standard copy and paste method (for the results presented in the main I.S.A.A.C.S. window) or using the 'Data menu' [Fig. 4.29]. Also the user has the possibility to export data either in a raw ASCII format (simple two

Figure 4.29 *The 'Data' menu of the graph window in the I.S.A.A.C.S. program.*

columns file with x and y) or in the Xmgrace format which can be used immediately in the Grace WYSIWYG 2D plotting tool [28] for a further analysis.

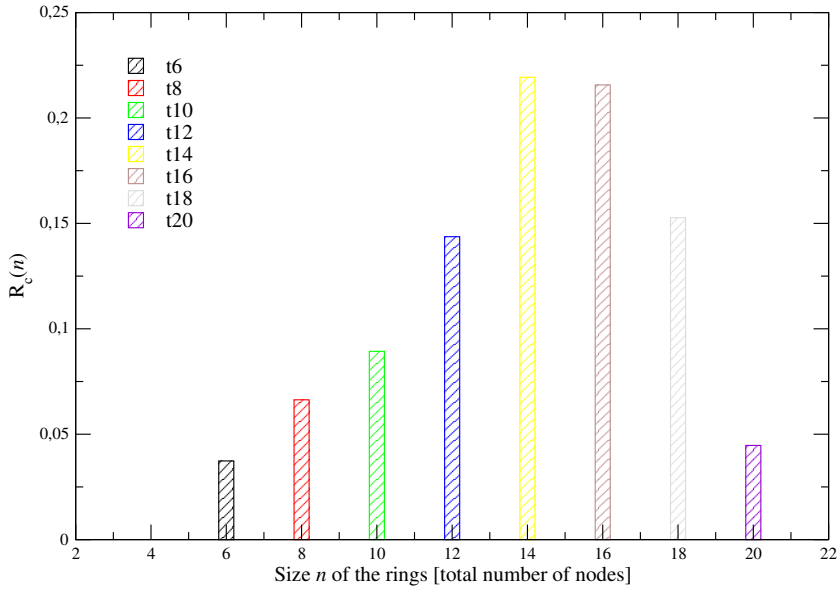


Figure 4.30 *Illustration of the utilization of the Xmgrace export filter in the I.S.A.A.C.S. program.*

Figure [Fig. 4.30] shows the data immediately saved from the results presented in figure [Fig. 4.23] using the Xmgrace export filter of the I.S.A.A.C.S. program.
Note that if more than one data sets are presented on the same graph window, then all data sets will be written in the same file when saving the data. Thus, for a particular calculation, if all data sets are added to the graph window using the data plot editing tool [Fig. 4.26], then all the data result of this analysis can be saved at once. This is true for both ASCII and Xmgrace file formats.

The physics in I.S.A.A.C.S.

5.1 The periodic boundary conditions

Taking into account the finite size of model/simulation box is crucial to computing correctly many of the structural characteristics (e.g. ring statistics) of the system being studied.
The importance of the finite size of model box can be illustrated using a 1 dm^3 edged cube of water (1 L) at room temperature. This cube contains approximately 3.3×10^{25} water molecules, each of them can be considered as a sphere having a diameter of 2.8 Å. Following this scheme surface interactions can affect up to 10 layers of spheres (water molecules) far from the surface of the model cubic box. In this case the number of water molecules exposed to the surface is about 2×10^{19}, which is a small fraction of the total numebr of molecules in the model.
Currently structure models often contain somewhere from 1 thousand to several thousands of molecules/atoms. As a result a very substantial fraction of them will be influenced by the finite size of the simulation/model box. The problems is solved by applying the so-called **P**eriodic **B**oundary **C**onditions "PBC" which means surrounding the simulation box with its translational images in the 3 directions of space, as illustrated below. Users of I.S.A.A.C.S. should take special care that their model boxes are inhearently periodic so that when the periodic boundary conditions are applied the structural characteristics computed are not compromized.

Figure [Fig. 5.1] illustrates the principle of the periodic boundary conditions that can be used[1] in I.S.A.A.C.S.: a particle which goes out from the simulation box by one side is reintroduced in the box by the opposite side (in the 3 dimensions of space). The maximum inter-atomic distance rcut which is taken into account in the calculations is therefore equal to the half of the edge of the simulation box:

$$r_{cut} = L/2 \tag{5.1}$$

The surface/finite model size effects would therefore be small, if any. In general, the larger the simulation box and the number of molecules/atoms in it, the smaller the surface/size effects will be.

[1] Please note that the use of PBC is not mandatory, isolated molecules can be studied using I.S.A.A.C.S.

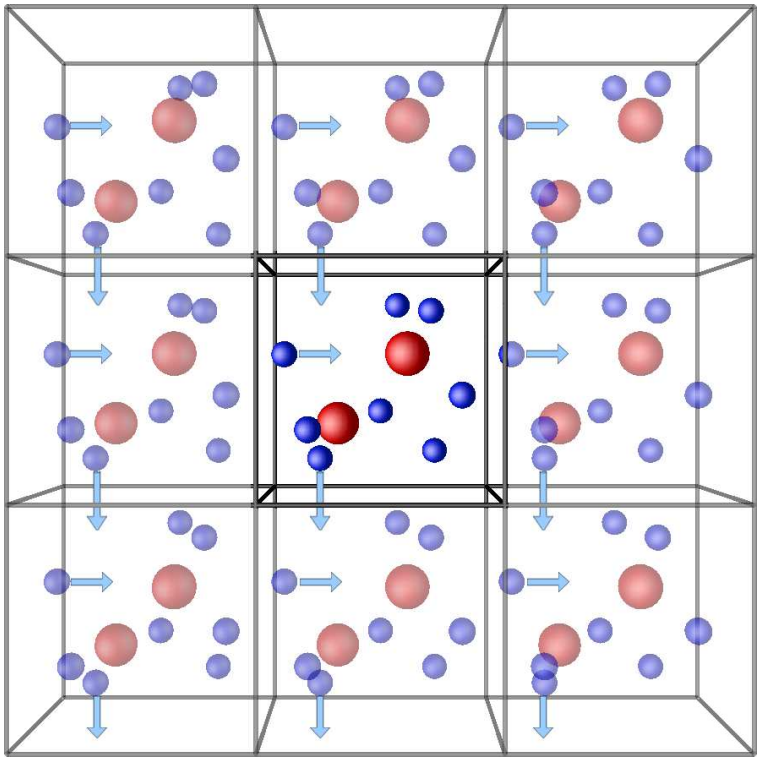


Figure 5.1 *Schematic representation of the idea of periodic boundary conditions.*

5.2 Radial distribution functions fundamentals

The Radial Distribution Function, R.D.F. , g(r), also called pair distribution function or pair correlation function, is an important structural characteristic, therefore computed by I.S.A.A.C.S. Considering a homogeneous distribution of the atoms/molecules in space, the $g(r)$ represents the probability to find an atom in a shell dr at the distance r of another atom chosen as a reference point [Fig. 5.2]. By dividing the physical space/model volume into shells dr [Fig. 5.2] it is possible to compute the number of atoms $dn(r)$ at a distance between r and $r+dr$ from a given atom:

$$dn(r) = \frac{N}{V} g(r)\, 4\pi\, r^2\, dr \tag{5.2}$$

where N represents the total number of atoms, V the model volume and where $g(r)$ is the radial distribution function. In this notation the volume of the shell of thickness dr is approximated:

$$\left(V_{\text{shell}} = \frac{4}{3}\pi(r+dr)^3 - \frac{4}{3}\pi r^3 \simeq 4\pi\, r^2\, dr \right) \tag{5.3}$$

When more than one chemical species are present the so-called partial radial distribution functions $g_{\alpha\beta}(r)$ may be computed :

$$g_{\alpha\beta}(r) = \frac{dn_{\alpha\beta}(r)}{4\pi r^2\, dr\, \rho_\alpha} \qquad \text{with} \qquad \rho_\alpha = \frac{V}{N_\alpha} = \frac{V}{N \times c_\alpha} \tag{5.4}$$

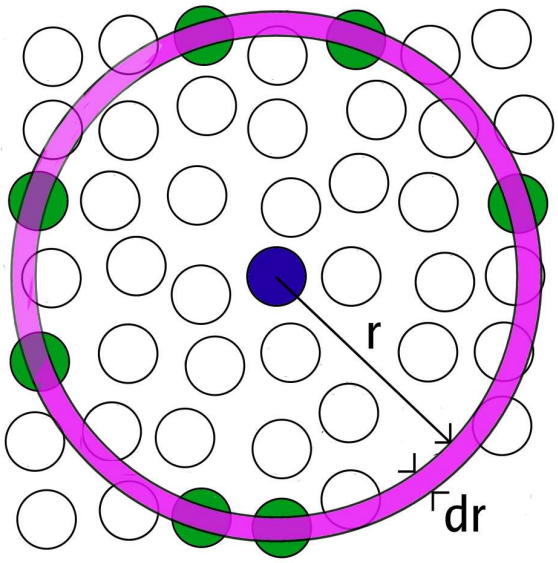


Figure 5.2 *Space discretization for the evaluation of the radial distribution function.*

where c_α represents the concentration of atomic species α.
These functions give the density probability for an atom of the α species to have a neighbor of the β species at a given distance r. The example features GeS_2 glass.

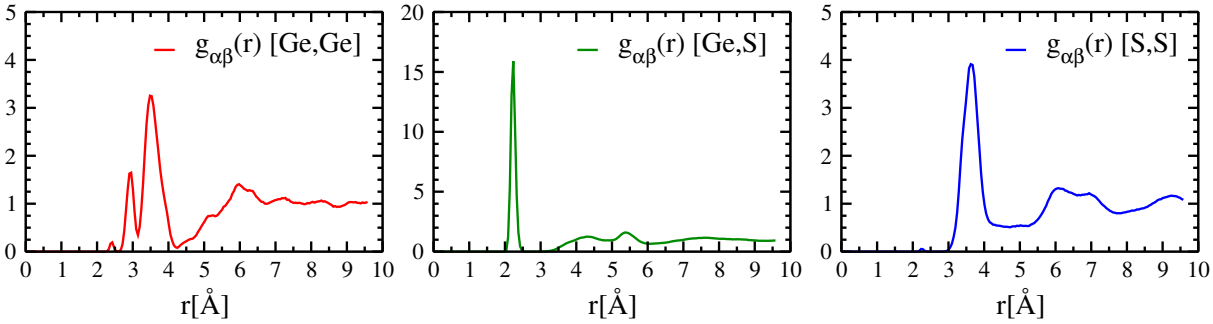


Figure 5.3 *Partial radial distribution functions of glassy GeS_2 at 300 K.*

Figure [Fig 5.3] shows the partial radial distribution functions for GeS_2 glass at 300 K. The total RDF of a system is a weighterd sum of the respective partial RDFs, with the weights depend on the relative concentration and x-ray/neutron scattering amplitudes of the chemical species involved.
It is also possible to use the reduced $\mathbf{G}_{\alpha\beta}(r)$ partial distribution functions defined as:

$$\mathbf{G}_{\alpha\beta}(r) \;=\; 4\pi r\rho_0\left(g_{\alpha\beta}(r)\;-1\right) \tag{5.5}$$

I.S.A.A.C.S. gives access to the partial $g_{\alpha\beta}(r)$ and $\mathbf{G}_{\alpha\beta}(r)$ distribution functions. Two methods are available to compute the radial distribution functions:

- The standard real space calculation typical to analyze 3-dimensional models
- The experiment-like calculation using the Fourier transform of the structure factor obtained using the Debye equation (see Sec. 5.3 for details).

5.3 Neutrons and X-rays scattering

Model static structure factors $S(q)$ may be compared to experimental scattering data and that is why are useful structural characteristics computed by I.S.A.A.C.S.Thereafter we describe the theoretical background of $S(q)$s computed by I.S.A.A.C.S.

5.3.1 Total scattering - Debye approach

Neutron or X-ray scattering static structure factor $S(q)$ is defined as:

$$S(q) = \frac{1}{N}\sum_{j,k} b_j b_k \left\langle e^{iq[\mathbf{r}_j - \mathbf{r}_k]} \right\rangle \tag{5.6}$$

where b_j et $\mathbf{r}_j$ represent respectively the neutron or X-ray scattering length, and the position of the atom j. N is the total number of atoms in the system studied.
To take into account the inherent/volume averaging of scattering experiments it is necessary to sum all possible orientations of the wave vector q compared to the vector $\mathbf{r}_j - \mathbf{r}_k$. This average on the orientations of the q vector leads to the famous Debye's equation:

$$S(q) = \frac{1}{N}\sum_{j,k} b_j b_k \frac{\sin(q|\mathbf{r}_j - \mathbf{r}_k|)}{q|\mathbf{r}_j - \mathbf{r}_k|} \tag{5.7}$$

Nevertheless the instantaneous individual atomic contributions introduced by this equation [Eq. 5.7] are not easy to interpret. It is more interesting to express these contributions using the formalism of radial distribution functions [Sec. 5.2].
In order to achieve this goal it is first necessary to split the self-atomic contribution ($j = k$), from the contribution between distinct atoms:

$$S(q) = \sum_j c_j b_j^2 + \underbrace{\frac{1}{N}\sum_{j\neq k} b_j b_k \frac{\sin(q|\mathbf{r}_j - \mathbf{r}_k|)}{q|\mathbf{r}_j - \mathbf{r}_k|}}_{I(q)} \tag{5.8}$$

with $c_j = \dfrac{N_j}{N}$.
$4\pi \sum_j c_j b_j^2$ represents the total scattering cross section of the material.

The function $I(q)$ which describes the interaction between distinct atoms is related to the radial distribution functions through a Fourier transformation:

$$I(q) \;=\; 4\pi\rho \int_0^{\infty} dr\, r^2\, \frac{\sin qr}{qr}\, G(r) \tag{5.9}$$

where the function $G(r)$ is defined using the partial radial distribution functions [Eq. 5.4]:

$$G(r) \;=\; \sum_{\alpha,\beta} c_\alpha b_\alpha\, c_\beta b_\beta\, (g_{\alpha\beta}(r) - 1) \tag{5.10}$$

where $c_\alpha = \dfrac{N_\alpha}{N}$ and b_α represents the neutron or X-ray scattering length of species α.
$G(r)$ approaches - $-\sum_{\alpha,\beta} c_\alpha b_\alpha\, c_\beta b_\beta$ for $r = 0$, and 0 for $r \to \infty$.
Usually the self-contributions are substracted from equation [Eq. 5.8] and the structure factor is normalized using the relation:

$$S(q) \;-\; 1 \;=\; \frac{I(q)}{\langle b^2 \rangle} \quad \text{avec} \quad \langle b^2 \rangle = \left(\sum_\alpha c_\alpha b_\alpha \right)^2 \tag{5.11}$$

It is therefore possible to write the structure factor [Eq. 5.7] in a more standard way:

$$S(q) \;=\; 1 \;+\; 4\pi\rho \int_0^{\infty} dr\, r^2\, \frac{\sin qr}{qr} (\mathbf{g}(r) - 1) \tag{5.12}$$

where $\mathbf{g}(r)$ (the radial distribution function) is defined as:

$$\mathbf{g}(r) \;=\; \frac{\sum_{\alpha,\beta} c_\alpha b_\alpha\, c_\beta b_\beta\, g_{\alpha\beta}(r)}{\langle b^2 \rangle} \tag{5.13}$$

In the case of a single atomic species system the normalization allows to obtain values of $S(q)$ and $\mathbf{g}(r)$ which are independent of the scattering factor/length and therefore independent of the measurement technique. In most cases, however, the total $S(q)$ and $\mathbf{g}(r)$ are combinations of the partial functions weighted using the scattering factor and therefore depend on the measurement technique (Neutron, X-rays ...) used or simulated.

Figure [Fig. 5.4] presents a comparison bewteen the calculations of the total neutron structure factor done using the Debye relation [Eq. 5.7] and the pair correlation functions [Eq. 5.12]. The material studied is a sample of glassy GeS_2 at 300 K obtained using ab-initio molecular dynamics. In several cases the structure factor $S(q)$ and the radial distribution function $\mathbf{g}(r)$ [Eq. 5.13] can be compared to experimental data. To simplify the comparison I.S.A.A.C.S. computes several radial distribution funcrions used in practice suh as $G(r)$ defined [Eq. 5.10], the differential correlation function $D(r)$, $\mathbf{G}(r)$, and the total correlation function $T(r)$ defined

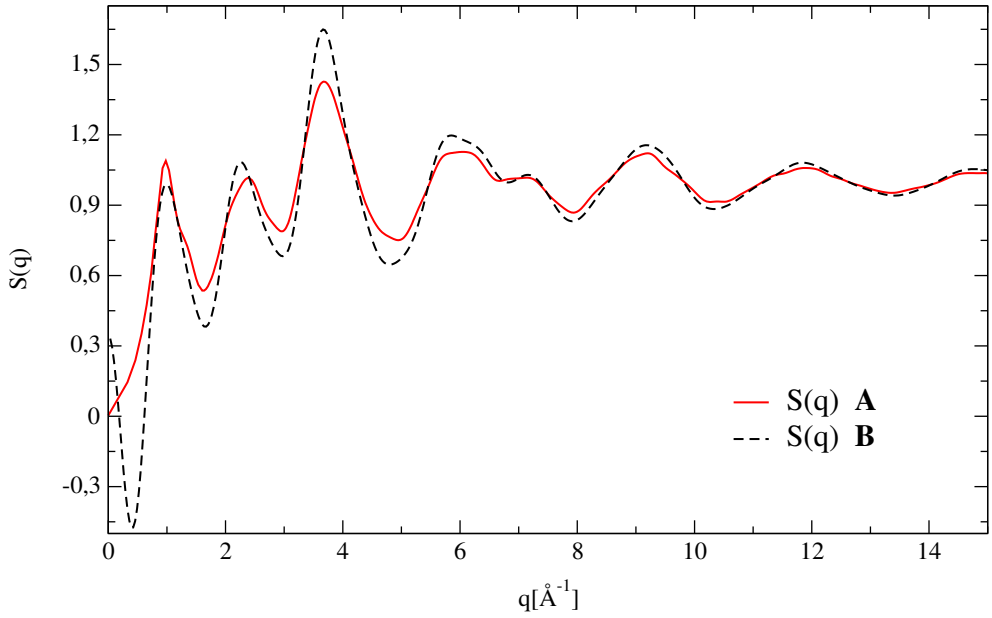


Figure 5.4 *Total neutron structure factor for glassy GeS_2 at 300 K* - **A** *Evaluation using the atomic correlations [Eq. 5.7],* **B** *Evaluation using the pair correlation functions [Eq. 5.12].*

by:

$$\begin{aligned} D(r) &= 4\pi r\rho\, G(r) \\ \mathbf{G}(r) &= \frac{D(r)}{\langle b^2 \rangle} \\ T(r) &= D(r) + 4\pi r\rho\, \langle b^2 \rangle \end{aligned} \tag{5.14}$$

$\mathbf{g}(r)$ equals zero for $r = 0$ and approaches 1 for $r \to \infty$.
$D(r)$ equals zero for $r = 0$ and approaches 0 for $r \to \infty$.
$\mathbf{G}(r)$ equals zero for $r = 0$ and approaches 0 for $r \to \infty$.
$T(r)$ equals zero for $r = 0$ and approaches ∞ for $r \to \infty$.
This set of functions for a model of GeS_2 glass (at 300 K) obtained using ab-intio molecular dynamics is presented in figure [Fig. 5.5].

I.S.A.A.C.S. can compute, for the case of x-ray or neutrons, the following functions:

- $S(q)$ and $Q(q) = q[S(q) - 1.0]$ [9, 10] computed using the Debye equation
- $S(q)$ and $Q(q) = q[S(q) - 1.0]$ [9, 10] computed using the Fourier transform of the $\mathbf{g}(r)$
- $\mathbf{g}(r)$ and $\mathbf{G}(r)$ computed using the standard real space calculation
- $\mathbf{g}(r)$ and $\mathbf{G}(r)$ computed using the Fourier transform of the structure factor calculated using the Debye equation

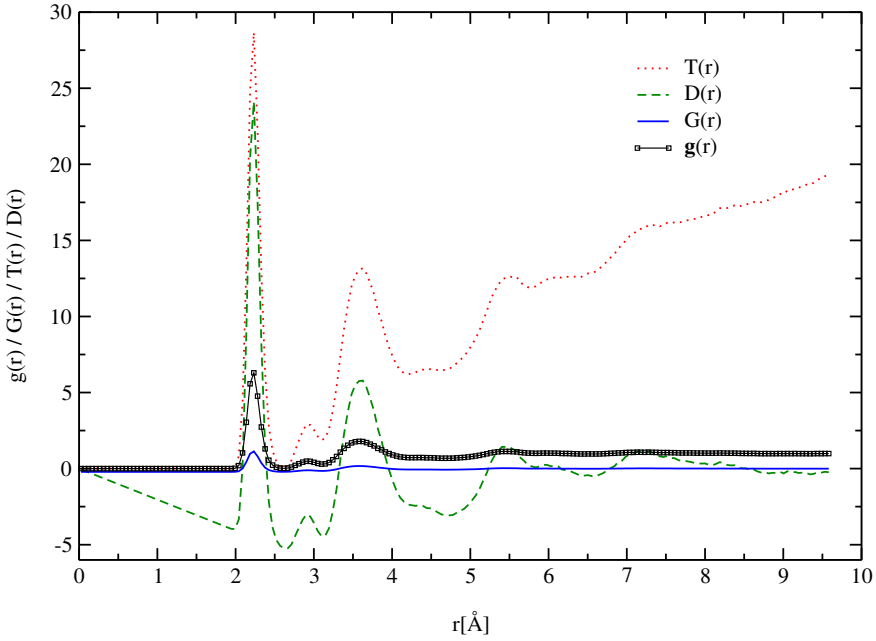


Figure 5.5 *Exemple of various distribution functions neutron-weighted in glassy GeS_2 at 300 K.*

5.3.2 Partial structure factors

There are a few, somewhat different definitions of partials $S(q)$ used in practice, and computed by I.S.A.A.C.S.

5.3.2.1 Faber-Ziman definition/formalism

One way used to define the partial structure factors has been proposed by Faber and Ziman [11]. In this approach the structure factor is represented by the correlations between the different chemical species. To describe the correlation between the α and the β chemical species the partial structure factor $S^{FZ}_{\alpha\beta}(q)$ is defined by:

$$S^{FZ}_{\alpha\beta}(q) \;=\; 1 \;+\; 4\pi\rho \int_0^\infty dr\, r^2\, \frac{\sin qr}{qr}\, \left(g_{\alpha\beta}(r) - 1\right) \tag{5.15}$$

where the $g_{\alpha\beta}(r)$ are the partial radial distribution functions [Eq. 5.4].
The total structure factor is then obtained by the relation:

$$S(q) \;=\; \sum_{\alpha,\beta} c_\alpha b_\alpha\, c_\beta b_\beta\, \left[S^{FZ}_{\alpha\beta}(q) \;-\; 1\right] \tag{5.16}$$

5.3.2.2 Ashcroft-Langreth definition/formalism

In a similar approach, based on the correlation between the chemical species, and developped by Ashcroft et Langreth [12–14], the partial structure factors $S^{AL}_{\alpha\beta}(q)$ are defined by:

$$S^{AL}_{\alpha\beta}(q) \;=\; \delta_{\alpha\beta} \;+\; 4\pi\rho\left(c_\alpha c_\beta\right)^{1/2} \int_0^\infty dr\, r^2\, \frac{\sin qr}{qr}\, \left(g_{\alpha\beta}(r)-1\right) \tag{5.17}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta, $c_\alpha = \dfrac{N_\alpha}{N}$, and the $g_{\alpha\beta}(r)$ are the partial radial distribution functions [Eq. 5.4].
Then the total structure factor can be calculated using:

$$S(q) \;=\; \frac{\displaystyle\sum_{\alpha,\beta} b_\alpha b_\beta\, \left(c_\alpha c_\beta\right)^{1/2}\, \left[S^{AL}_{\alpha\beta}(q) \;+\; 1\right]}{\displaystyle\sum_{\alpha} c_\alpha b_\alpha^2} \tag{5.18}$$

5.3.2.3 Bhatia-Thornton definition/formalism

In this approach, used in the case of binary systems AB_x [15] only, the total structure factor $S(q)$ can be express as the weighted sum of 3 partial structure factors:

$$S(q) = \frac{\langle b\rangle^2 S_{NN}(q) + 2\langle b\rangle(b_A - b_B)S_{NC}(q) + (b_A - b_B)^2 S_{CC}(q) - (c_A b_A^2 + c_B b_B^2)}{\langle b\rangle^2} \;+\; 1 \tag{5.19}$$

where $\langle b\rangle = c_A b_A + c_B b_B$, with c_A and b_A reprensenting respectively the concentration and the scattering length of species A.
$S_{NN}(q)$, $S_{NC}(q)$ and $S_{CC}(q)$ represent combinaisons of the partial structure factors calculated using the Faber-Ziman formalism and weighted using the concentrations of the 2 chemical species:

$$S_{NN}(q) = \sum_{A=1}^{2}\sum_{B=1}^{2} c_A c_B S^{FZ}_{AB}(q) \tag{5.20}$$

$$S_{NC}(q) = c_A c_B \times \left[\; c_A \times \left(S^{FZ}_{AA}(q) - S^{FZ}_{AB}(q)\right) \;-\; c_B \times \left(S^{FZ}_{BB}(q) - S^{FZ}_{AB}(q)\right) \;\right] \tag{5.21}$$

$$S_{CC}(q) = c_A c_B \times \left[1 + c_A c_B \times \left[\sum_{A=1}^{2}\sum_{B\neq A}^{2} \left(S^{FZ}_{AA}(q) - S^{FZ}_{AB}(q)\right)\right]\right] \tag{5.22}$$

- $S_{NN}(q)$ is the Number-Number partial structure factor.
  Its Fourier transform allows to obtain a global description of the structure of the solid, ie. of the repartition of the experimental scattering centers, or atomic nuclei, positions. The nature of the chemical species spread in the scattering centers is not considered. Furthermore if $b_A = b_B$ then $S_{NN}(q) = S(q)$.

- $S_{CC}(q)$ is the Concentration-Concentration partial structure factor.
  Its Fourier transform allows to obtain an idea of the distribution of the chemical species over the scattering centers described using the $S_{NN}(q)$. Therefore the $S_{CC}(q)$ describes the chemical order in the material. In the case of an ideal binary mixture of 2 chemical species A and B[2], $S_{CC}(q)$ is constant and equal to $c_A c_B$. In the case of an ordered chemical mixture (chemical species with distinct diameters, and with heteropolar and homopolar chemical bonds) it is possible to link the variations of the $S_{CC}(q)$ to the product of the concentrations of the 2 chemical species of the mixture:
  - $S_{CC}(q) = c_A c_B$: radom distribution.
  - $S_{CC}(q) > c_A c_B$: homopolar atomic correlations (A-A, B-B) prefered.
  - $S_{CC}(q) < c_A c_B$: heterpolar atomic correlations (A-B) prefered.
  - $\langle b \rangle = 0$: $S_{CC}(q) = S(q)$.
- $S_{NC}(q)$ is the Number-Concentration partial structure factor.
  Its Fourier transform allows to obtain a correlation between the scattering centers and their occupation by a given chemical species. The more the chemical species related partial structure factors are different ($S_{AA}(q) \neq S_{BB}(q)$) and the more the oscillations are important in the $S_{NC}(q)$. In the case of an ideal mixture $S_{NC}(q) = 0$, and all the information about the structure of the system is given by the $S_{NN}(q)$.

If we consider the binary mixture as an ionic mixture then it is possible to calculate the Charge-Charge $S_{ZZ}(q)$ and the Number-Charge $S_{NZ}(q)$ partial structure factors using the Concentration-Concentration $S_{CC}(q)$ and the Number-Concentration $S_{NC}(q)$:

$$S_{ZZ}(q) = \frac{S_{CC}(q)}{c_A c_B} \qquad \textit{and} \qquad S_{NZ}(q) = \frac{S_{NC}(q)}{c_B / Z_A} \tag{5.23}$$

c_A et Z_A represent the concentration and the charge of the chemical species A, the global neutrality of the system must be respected therefore $c_A Z_A + c_B Z_B = 0$.

Figure [Fig. 5.6] illustrates, and allows to compare, the partial structure factors of glassy GeS_2 at 300 K calculated in the different formalisms Faber-Ziman [11], Ashcroft-Langreth [12–14], and Bhatia-Thornton [15].

I.S.A.A.C.S. can compute the following partial structure factors:

- Faber-Ziman $S^{FZ}_{\alpha\beta}(q)$
- Ashcroft-Langreth $S^{AL}_{\alpha\beta}(q)$
- Bhatia-Thornton $S_{NN}(q)$, $S_{NC}(q)$, $S_{CC}(q)$ and $S_{ZZ}(q)$

[2]Particles that can be described using spheres of the same diameter and occupying the same molar volume, subject to the same thermal constrains, in a mixture where the substitution energy of a partcile by another is equal to zero.

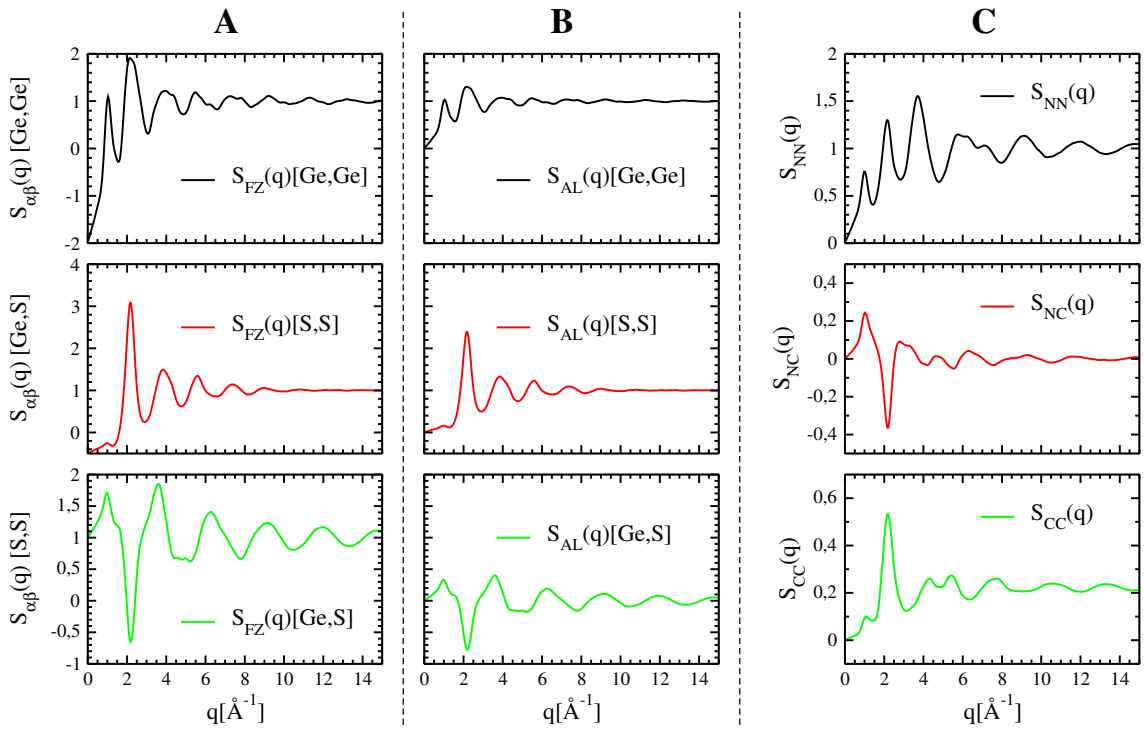


Figure 5.6 *Partial structure factors of glassy GeS_2 at 300 K.* **A** *Faber-Ziman [11],* **B** *Ashcroft-Langreth [12–14] et* **C** *Bhatia-Thornton [15].*

5.4 Local atomic coordination properties

Several properties related to the atomic bonds and angles between them can be computed using I.S.A.A.C.S. In I.S.A.A.C.S. the existence or the absence of a bond between two atoms i of species α and j of species β is determined by the analysis of the partial $g_{\alpha\beta}(r)$ and total $g(r)$ radial distribution functions. Precisely the program will consider that a bond exists if the interatomic distance D_{ij} is smaller than both the cutoff given to desribe the maximum distance for first neighbor atoms between the species α and β, $Rcut_{\alpha\beta}$ (often the first minimum of the partial radial distribution function $g_{\alpha\beta}(r)$), and the first minimum of the total radial distribution function, $Rcut_{tot}$.
I.S.A.A.C.S. allows the user to specify both $Rcut_{\alpha\beta}$ and $Rcut_{tot}$ to choose an appropriate definition of the atomic bonds to described the system under study. When atomic bonds in a model are defined properly other structural characteristics can be evaluated, as follows:

5.4.1 Average first coordination numbers

I.S.A.A.C.S. computes total as well as partials coordinations numbers.

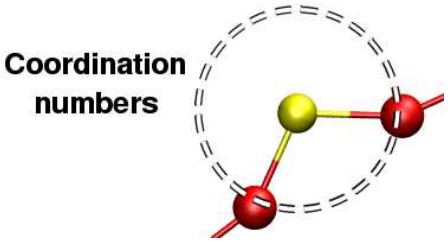


Figure 5.7 *Coordination numbers.*

5.4.2 Individual atomic neighbor analysis

I.S.A.A.C.S. computes the fraction of each type of first coordination spheres occurring in the model. The presence of of structural deffects can lead to a wide number of local structural environments, figure [Fig. 5.8] illustrates the differents first coordination spheres that can be found in a GeS_2 glass.

5.4.3 Proportion of tetrahedral links and units in the structure model

Often the structure of a material is represented using building blocks. One of the the most frecuently occuring building blocks are tetrahedra. Figure [Fig. 5.9] shows a model of GeS_2 materials using GeS_4 tetrahedra as building blocks.

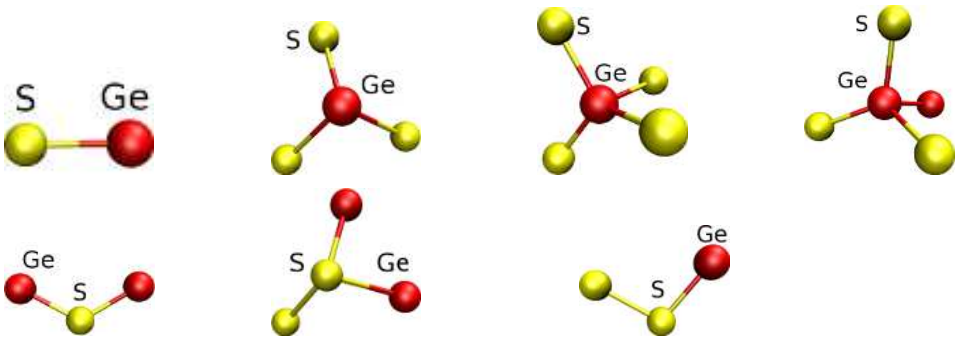


Figure 5.8 *Illustration of several coordination spheres that can be found in glassy GeS_2.*

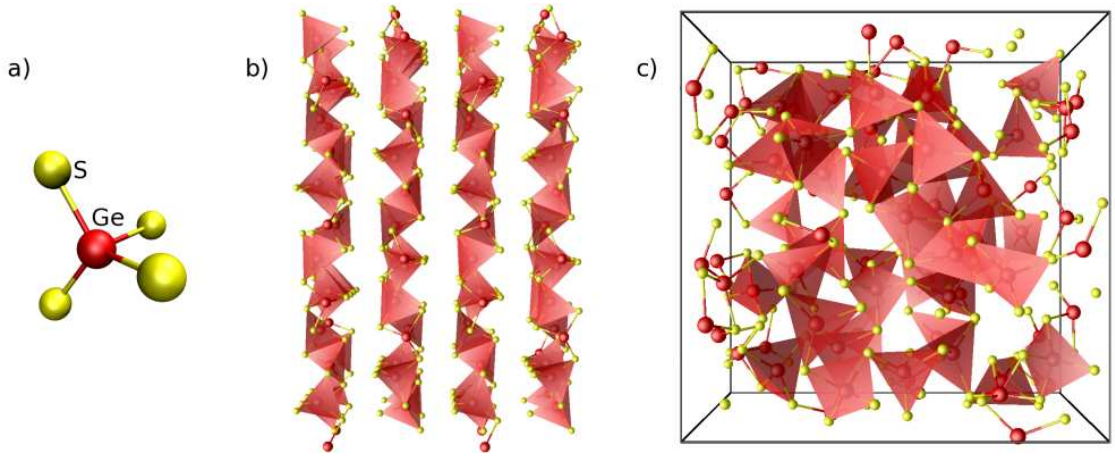


Figure 5.9 *Illustration of the presence of GeS_4 tetrahedra in the GeS_2 material's family. a) GeS_4 tetrahedra, representations b) of the α-GeS_2 crystal and c) of the GeS_2 glass using tetrahedra.*

I.S.A.A.C.S. computes the fraction of the differents tetrahedra in materials, the distinction between these tetrahedra being made on the nature of the connection between each of them. Tetrahedra can be linked either by corners or edges [Fig. 5.10], I.S.A.A.C.S. computes the fraction of atoms forming tetrahedra as well as to the fraction of linked tetrahedra.

5.4.4 Distribution of bond lengths for the first coordination sphere

I.S.A.A.C.S. gives access to the bond length distribution between first neighbor atoms [Fig. 5.11].

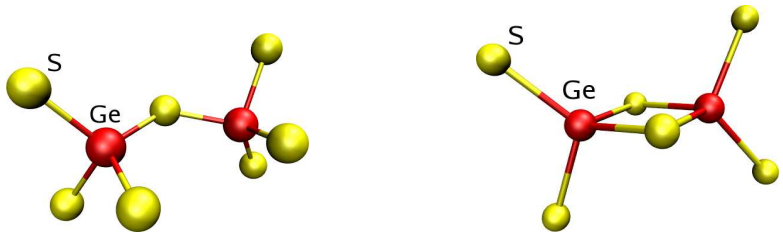


Figure 5.10 *Corner sharing (left) and edge sharing (right) tetrahedra.*

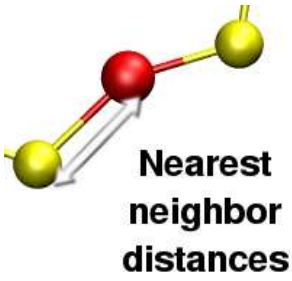


Figure 5.11 *Nearest neighbor distances distribution.*

5.4.5 Angles distribution

Using I.S.A.A.C.S. it is very easy to compute bond angles and dihedral angles [Fig. 5.12] distributions:

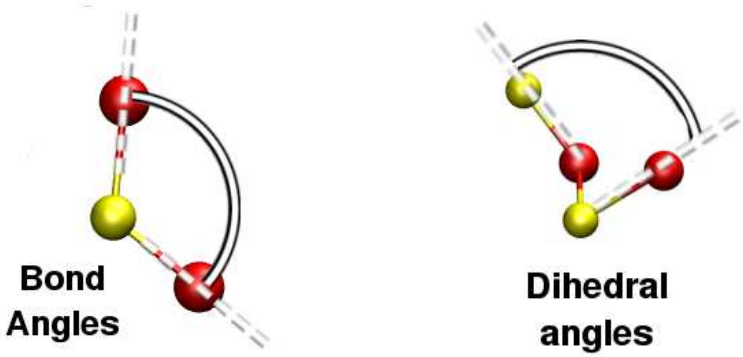


Figure 5.12 *Bond angles (left) and diehdral angles (right).*

5.5 Ring statistics

The analysis of the topology of network-type structure models (liquid, crystalline or amorphous systems) is often based on the part of the structural information which can be represented in the graph theory using nodes for the atoms and links for the bonds. The absence or the existence of a link between two nodes is determined by the analysis of the total and partial radial distribution functions of the system.
In such a network a series of nodes and links connected sequentially without overlap is called a path. Following this definition a ring is therefore simply a closed path. If we study thoroughly a specific node of this network we see that this node can be involved in numerous rings. Each of these rings is characterized by its size and can be classified based upon the relations between the nodes and the links which constitute it.

5.5.1 Size of the rings

There are two possibilities for the numbering of rings. On the one hand, one can use the total number of nodes of the ring, therefore a N-membered ring is a ring containing N nodes. One the other hand, one can use the number of *network forming* nodes (ex: Si atoms in SiO_2 and Ge atoms in GeS_2 which are the atoms of highest coordination in these materials) an N-membered ring is therefore a ring containing $2\times$N nodes. For crystals and SiO_2-like glasses the second definition is usually applied. Nevertheless the first method has to be used in the case of chalcogenide liquids and glasses in order to count rings with homopolar bonds (ex: Ge-Ge and S-S bonds in GeS_2) - See Section 5.5.5 for further details.
From a theoretical point of view it is possible to obtain an estimate for the ring of maximum size that could exist in a network. This theoretical maximum size will depend on the properties of the system studied as well as on the definition of a ring.

5.5.2 Definitions

5.5.2.1 King's shortest paths criterion

The first way to define a ring has been given by Shirley V. King [16] (and later by Franzblau [17]). In order to study the connectivity of glassy SiO_2 she defines a ring as the shortest path between two of the nearest neighbors of a given node [Fig. 5.13].
In the case of the King's criterion one can calculate the maximum number of different ring sizes, $NS_{max}(KSP)$, which can be found using the atom **At** to initiate the search:

$$NS_{max}(KSP) = \frac{Nc(\mathbf{At}) \times (Nc(\mathbf{At}) - 1)}{2} \tag{5.24}$$

where $N_c(\mathbf{At})$ is the number of neighbors of atom **At**. $NS_{max}(KSP)$ represents the number of ring sizes found if all couples of neighbors of atom **At** are connected together with paths of different sizes.
It is also possible to calculate the theoretical maximum size, $TMS(KSP)$, of a King's shortest

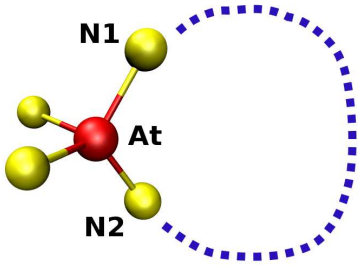


Figure 5.13 *King's criterion in the ring statistics: a ring represents the shortest path between two of the nearest neighbors (***N1** *and* **N2***) of a given node (***At***).*

path ring in the network using:

$$TMS(KSP) = 2 \times (D_{max} - 2) \times (Nc_{max} - 2) + 2 \times D_{max} \quad (5.25)$$

where D_{max} is the longest distance, in number of chemical bonds, separating two atoms in the network, and Nc_{max} represents the average number of neighbors of the chemical species of higher coordination. If used when looking for rings, periodic boundary conditions have to be taken into account to calculate D_{max}. The relation [Eq. 5.25] is illustrated in figure [Fig. 5.16-2].

5.5.2.2 Guttman's shortest paths criterion

A later definition of ring was proposed by Guttman [18], who defines a ring as the shortest path which comes back to a given node (or atom) from one of its nearest neighbors [Fig. 5.14].

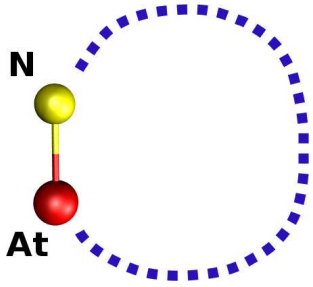


Figure 5.14 *Guttman's criterion in the ring statistics: a ring represents the shortest path which comes back to a given node (***At***) from one of its nearest neighbors (***N***).*

Differences between the King and the Guttman's shortest paths criteria are illustrated in figure [Fig. 5.15].
Like for the King's criterion, with the Guttman's criterion one can calculate the maximum

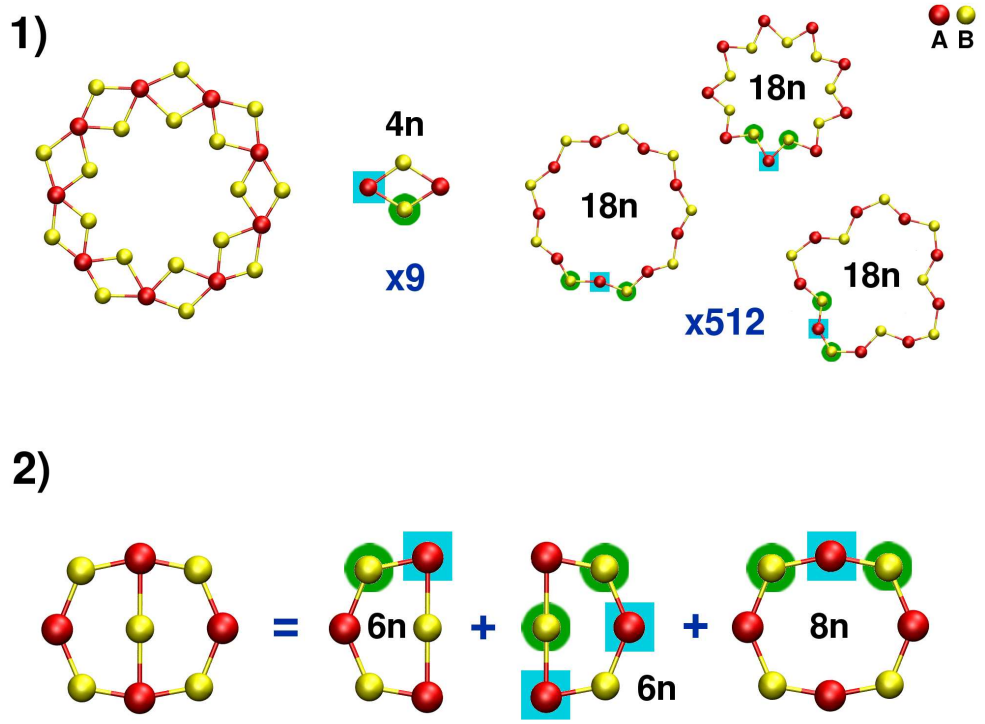


Figure 5.15 *Differences between the King and the Guttman shortest paths criteria for the ring statistics in an AB_2 system. In these two examples the search is initiated from chemical species A (blue square). The nearest neighbor(s) of chemical species B (green circles) are used to continue the analysis. 1) In the first example only rings with 4 nodes are found using the Guttman's criterion, whereas rings with 18 nodes are also found using the King's criterion (2^9 rings with 18 nodes). 2) In the second example the King's shortest path criterion allows to find the ring with 8 nodes ignored by the Guttman's criterion which is only able to find the rings with 6 nodes.*

number of different ring sizes, $NS_{max}(GSP)$, which can be found using the atom **At** to initiate the search:

$$NS_{max}(GSP) = N_c(\mathbf{At}) - 1 \tag{5.26}$$

where $N_c(\mathbf{At})$ is the number of neighbors of atom **At**. $NS_{max}(GSP)$ represents the number of ring sizes found if the neighbors of atom **At** are connected together with paths of different sizes. It is also possible to calculate the **T**heoretical **M**aximum **S**ize, $TMS(GSP)$, of a Guttman's ring in the network using:

$$TMS(GSP) = 2 \times D_{max} \tag{5.27}$$

where D_{max} represents the longest distance, in number of chemical bonds, separating two atoms in the network. If used when looking for rings, periodic boundary conditions have to be taken into account to calculate D_{max}. The relation [Eq. 5.27] is illustrated in figure [Fig. 5.16-1].

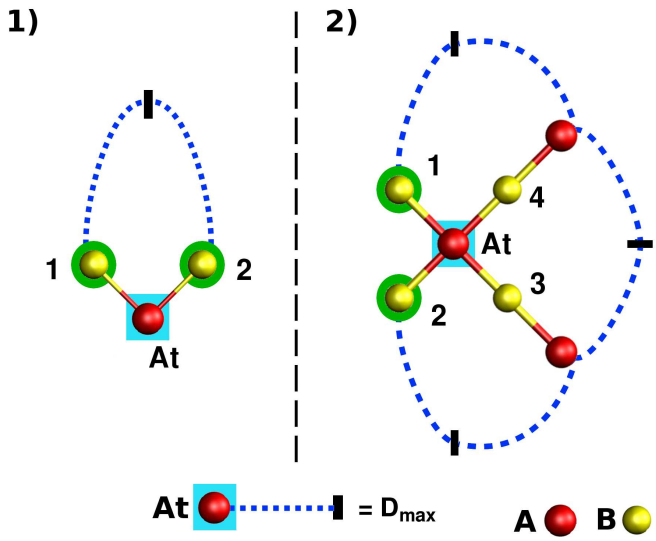


Figure 5.16 *Theoretical maximum size of the rings for an AB_2 system ($Nc_{max} = Nc_A = 4$) and using: 1) the Guttman's criterion, 2) the King's criterion. The theoretical maximum size represent the longest distance between two nearest neighbors 1 and 2 (green circles) of the atom* **At** *used to initiate the search (blue square).*

Since the introduction of the King's and the Guttman's criteria other definitions of rings have been proposed. These definitions are based on the properties of the rings to be decomposed into the sum of smaller rings.

5.5.2.3 The primitive rings criterion

A ring is primitive [19, 20] (or Irreducible [21]) if it can not be decomposed into two smaller rings [Fig. 5.17].

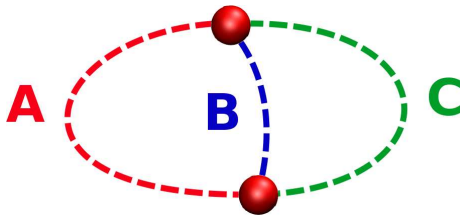


Figure 5.17 *Primitive rings in the ring statistics: the 'AC' ring defined by the sum of the A and the C paths is primitive only if there is no B path shorter than A and shorter than C which allows to decompose the 'AC' ring into two smaller rings 'AB' and 'AC'.*

The primitive rings analysis between the paths in figure [Fig. 5.17] may lead to 3 results depending on the relations between the paths A, B, and C:

- If paths A, B, and C have the same length: A = B = C then the rings 'AB', 'AC' and 'BC' are primitives.
- If the relation between the paths is like ? =? <? (ex: A = B < C) then 1 smaller ring ('AB') and 2 bigger rings ('AC' and 'BC') exist. None of these rings can be decomposed into the sum of two smaller rings therefore the 3 rings are again primitives.
- If the relation between the path is like ? <? =? (ex: A < B = C) or ? <? <? (ex: A < B < C) then a shortest path exists (A). It will be possible to decompose the ring ('BC') built without this shortest path into the sum of 2 smaller rings ('AB' and 'AC'), therefore this ring will not be primitive.

5.5.2.4 The strong rings criterion

The strong rings [19, 20] are defined by extending the definition of primitive rings. A ring is strong if it can not be decomposed into a sum of smaller rings whatever this sum is, ie. whatever the number of paths in the decomposition is.

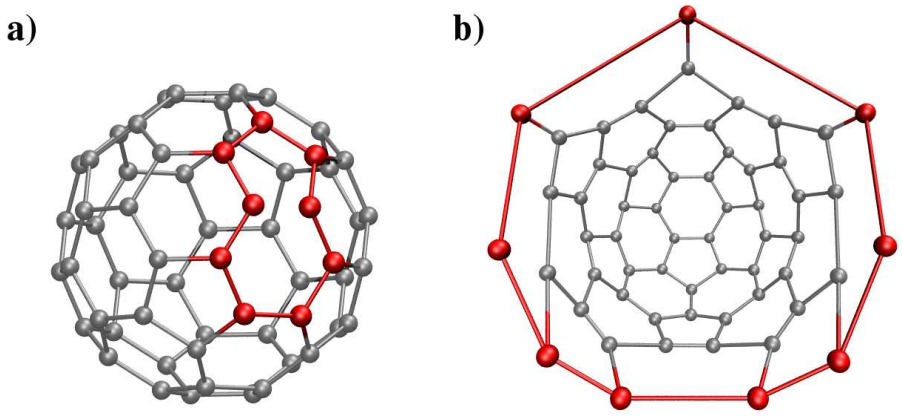


Figure 5.18 *Strong rings in the ring statistics:* **a)** *the 9-carbon-atoms ring created after breaking a C-C bond in a Buckminster fulleren molecule is a counterexample of strong ring;* **b)** *the combination of shortest rings, 11 5-carbon-atoms rings and 19 6-carbon-atoms rings, appears easily after the deformation of the* C_{60} *molecule.*

By definition the strong rings are also primitives, therefore to search for strong rings can be summed as to find the strong rings among the primitive rings. This technique is limited to relatively simple cases, like crystals or structures such as the one illustrated in figure [Fig. 5.18]. On the one hand the CPU time needed to complete such an analysis for amorphous systems is very important. On the other hand it is not possible to search for strong rings using the same search depth than for other types of rings. The strong ring analysis is indeed diverging which makes it very complex to implement for amorphous materials.

In the case of primitive rings like in the case of strong rings, there is no theoretical maximum size of rings in the network.

5.5.3 Description of a network using ring statistics - existing tools

Ring statistics are mainly used to obtain a snapshot of the connectivity of a network. Thereby the better the snapshot will be, the better the description and the understanding of the properties of the material will be. In the literature many papers present studies of materials using ring statistics. In these studies either the number of **R**ings per **N**ode 'R_N' [29, 30] or the number of **R**ings per **C**ell 'R_C' [31–33] are given as a result of the analysis. The first (R_N) is calculated for one node by counting all the rings correponding to the property we are looking for (King's, Guttman's, primitive or strong ring criterion). The second (R_C) is calculated by counting all the different rings corresponding at least once (at least for one node) to the property we are looking for (King's, shortest path, primitive or strong ring criterion). The values of R_N and R_C are often reduced to the number of nodes of the networks. Furthermore the results are presented according to each size of rings.
An example is proposed with a very simple network illustrated in figure [Fig. 5.19]. This network is composed of 10 nodes, arbitrary of the same chemical species, and 7 bonds. Furthermore it is clear that in this network there are 1 ring with 3 nodes and 1 ring with 4 nodes.

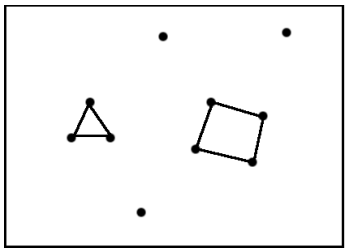


Figure 5.19 *A very simple network.*

It is easy to calculate R_N and R_C for the network in figure [Fig. 5.19] (n = number of nodes):

n	$R_N(n)$
3	1/10
4	1/10

n	$R_C(n)$
3	3/10
4	4/10

In the literature the values of R_N and R_C are usually given separately [29–33].
Nevertheless these two properties are not sufficient in order to describe a network using rings. A simple example is proposed in figure [Fig. 5.20]. The two networks [Fig. 5.20-a] and [Fig. 5.20-b] do have very similar compositions with 10 nodes and 7 links but they are clearly different. Nevertheless the previous definitions of rings per cell and rings per node even taken together will lead to the same description for these two different networks [Tab. 5.1].

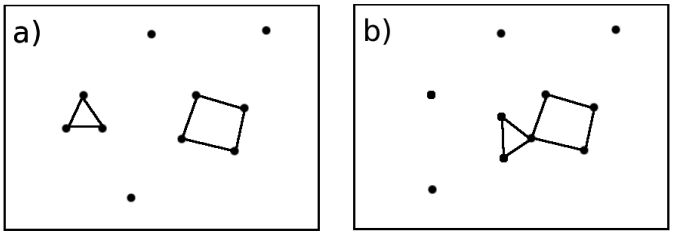


Figure 5.20 *Two simple networks having very close compositions: 10 nodes and 7 links.*

n	$R_N(n)$
3	1/10
4	1/10

n	$R_N(n)$
3	3/10
4	4/10

Table 5.1 *R_N and R_C calculated for the networks illustrated in figure [Fig. 5.20].*

In both cases a) and b) there are 1 ring with 3 nodes and 1 ring with 4 nodes. It has to be noticed that these two rings have properties which correspond to each of the definitions introduced previously (King, Guttman, primitive and strong). Thus none of these definitions is able to help to distinguish between these two networks. Therefore eventhough these simple networks are different, the previous definitions lead to the same description.

Thereby it is justified to wonder about the interpretation of the data presented in the literature for amorphous systems with a much higher complexity.

5.5.4 Rings and connectivity: the new R.I.N.G.S. method implemented in I.S.A.A.C.S.

In the I.S.A.A.C.S. program the results of the ring statistics analysis are outputted following the new R.I.N.G.S. method [22, 34], this method is presented in the next pages.

The first goal of ring statistics is to give a faithful description of the connectivity of a network and to allow to compare this information with others obtained for already existing structures. It is therefore important to find a guideline which allows to establish a distinction and then a comparison between networks studied using ring statistics. We propose thereafter a new method to achieve this goal. First of all we noticed fundamental points that must be considered to get a reliable and transferable method:

1. *The results must be reduced to the* **total** *number of nodes in the network.*
   The nature of the nodes used to initiate the analysis when looking for rings will have a significant influence, therefore it is essential to reduce the results to a value for one node.

Otherwise it would be impossible to compare the results to the ones obtained for systems made of nodes (particles) of different number and/or nature.

2. *Different networks must be distinguishable whatever the method used to define a ring.* Indeed it is essential for the result of the analysis to be trustworthy independently of the method used to define a ring (King, Guttman, primitives, strong). Furthermore this will allow to compare the results of these different ring statistics.

5.5.4.1 Number of rings per cell 'R_C'

We have already introduced this value, which is the first and the easiest way to compare networks using ring statistics.

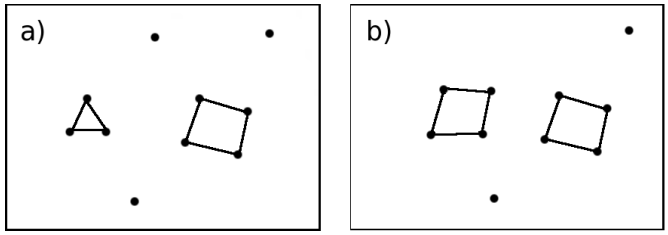


Figure 5.21 *The first comparison element: the total number of rings in the network.*

a)

n	$R_C(n)$
3	1/10
4	1/10

b)

n	$R_C(n)$
3	0/10
4	2/10

Table 5.2 *Number of rings in the simple networks represented in figure [Fig. 5.21].*

In the most simple cases, such as the one represented in figure [Fig. 5.21], the networks can be distinguished using only the number of rings [Tab. 5.2]. Nevertheless in most of the cases other informations are needed to describe accurately the connectivity of the networks.

5.5.4.2 Description of the connectivity: difference between rings and nodes

The second information needed to investigate the properties of a network using rings is the evaluation of the connectivity between rings. Indeed the distribution of the ring sizes gives a first information on the connectivity, nevertheless it can not be exactly evaluated unless one studies how the rings are connected. The impact of the relations between rings, already presented in figure [Fig. 5.20], has been illustrated in detail in figure [Fig. 5.22]. Figure [Fig. 5.22] represents the different possibilities to combine 2 rings with 6 nodes and 1 ring with 4

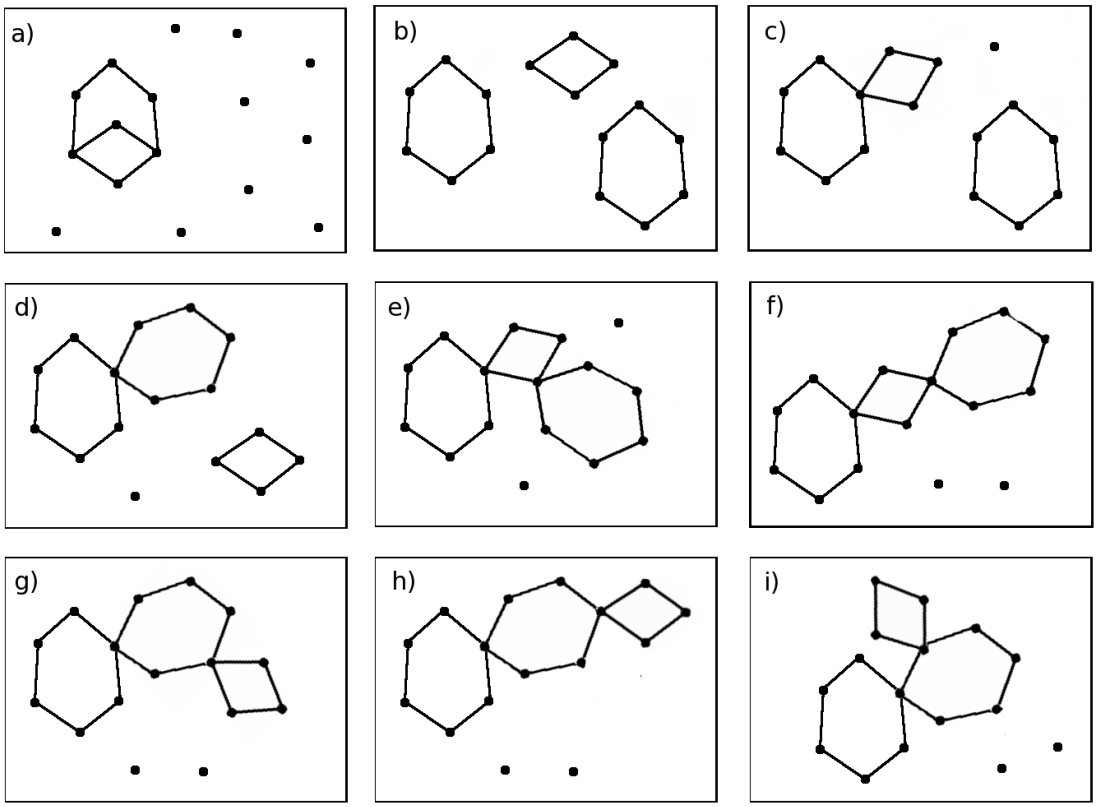


Figure 5.22 *Illustration of the 9 different newtorks with 16 nodes, composed of 2 rings with 6 nodes and 1 ring with 4 nodes.*

nodes in a network composed of 16 nodes. Among the 9 networks presented in figure [Fig. 5.22] none can be distinguished using the R_C value [Tab. 5.3].

Table 5.3 *Number of rings for the different networks presented in figure [Fig. 5.22].*

n	$R_C(n)$
4	1/16
6	2/16

Furthermore it is not possible to distinguish these networks using the R_N value. It seems possible to isolate the case a) [Tab. 5.4] from the other cases b) $\rightarrow$ i) [Tab. 5.4]. Nevertheless the results obtained using the primitive rings criterion are similar for all cases a) $\rightarrow$ i) [Tab. 5.4], this is in contradiction with the second statement [2] proposed in our method.

Before introducing parameters able to distinguish the configurations presented in figure

Case a)	$R_N(n)$	
n	King / Guttman.	Primitive / Strong.
4	4/16	**4/16**
6	10/16	**12/16**

Cases b) → i)	$R_N(n)$
n	All criteria.
4	**4/16**
6	**12/16**

Table 5.4 *Number of rings per node for the networks presented in figure [Fig. 5.22].*

[Fig. 5.22] it is important to wonder about the number of cases to distinguish. From the point of view of the connectivity of the rings, configurations a), b), c) and d) are clearly different. Nevertheless following the same approach configurations e) and f) on the one hand and configurations g), h) and i) on the other hand are identical. A schematic representation [Fig. 5.23] is sufficient to illustrate the similarity of the relations between these networks. The difference between each of these networks does not appear in the connectivity of the rings but in the connectivity of the particles.

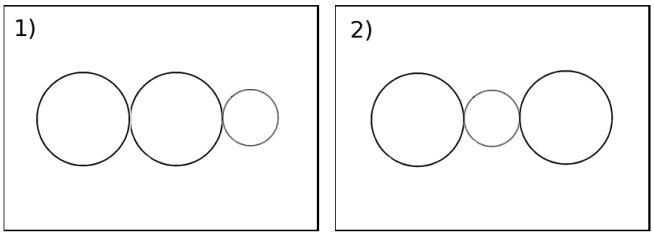


Figure 5.23 *Schematic representation of cases g) → i) (1) and e) → f) (2) illustrated in figure [Fig. 5.22].*

Thus among the networks illustrated in figure [Fig. 5.22] six dispositions of the rings have to be distinguished (a, b, c, d, e, g). The proportions of particles involved, or not involved, in the construction of rings will become an important question.

The new tool defined in our method is able to describe accurately the information still missing on the connectivity. It is a square symmetric matrix of size $(R-r+1)\times(R-r+1)$, where R and r represent respectively the bigger and the smaller size of a ring found when analyzing the network: we have called this matrix the connectivity matrix [Tab. 5.5].
The diagonal elements $P_N(i)$ of this matrix represent the **P**roportion of **N**odes at the origin of at least one ring of size i. And the non-diagonal elements $P_N(i,j)$ represent the **P**roportion of

$$C_{mat} = \begin{bmatrix} P_N(r) & P_N(r+1,r) & \cdots & P_N(R,r) \\ P_N(r,r+1) & \ddots & & P_N(R,r+1) \\ \vdots & & \ddots & \vdots \\ P_N(r,R) & \cdots & \cdots & P_N(R) \end{bmatrix}$$

Table 5.5 *General connectivity matrix.*

Nodes at the origin of ring(s) of size i and j.
The matrix elements have a value ranging between 0 and 1. The lowest and non equal to 0 is of the form $\frac{1}{Nn}$, the highest and non equal to 1 is of the form $\frac{Nn-1}{Nn}$, where Nn represents the number of nodes in the network.

The connectivity matrix of the configurations illustrated in figure [Fig. 5.22] are presented in table [Tab. 5.6]. We see that this matrix allows to distinguish each network whatever the way used to define a ring is. This matrix remains simple for small systems (crystalline or amorphous) or when using a small maximum ring size for the analysis. Nevertheless its reading can be considerably altered when analysing amorphous systems with a high maximum ring size for the analysis.

To simplify the reading and the interpretation of the data contained in this matrix for more complex systems, we chose a similar approach to extract informations on the connectivity between the rings. As a first step we decided to evaluate only the diagonal elements $P_N(n)$ of the general connectivity matrix. Indeed these values allow us to obtain a better view of the connectivity than the standard R_N value.
It is clear [Tab. 5.7] that using $P_N(n)$ improves the separation between the networks illustrated in figure [Fig. 5.22]. Nevertheless $P_N(n)$ does not allow to distinguish each of them. We notice that the distinction between networks is improved [Tab. 5.7] in particular when compared to the one obtain with $R_N(n)$ [Tab. 5.4].

Therefore in a second step we chose to calculate two properties whose definitions are very similar to the one of $P_N(n)$. The first, named $P_{N_{\max}}(n)$, represents the proportion of nodes for which the rings with n nodes are the longest closed paths found using these nodes to initiate the search. The second named, $P_{N_{\min}}(n)$, represents the proportion of nodes for which the rings with n nodes are the shortest closed paths found using these nodes to initiate the search.
The terms *longest* and *shortest path* must be considered carefully to avoid any confusion with the terms used in section [S. 5.5.2] to define the rings. For one node it is possible to find several rings whose properties correspond to the definitions proposed previously (King's, Guttman's, primitive or strong ring criterion). These rings are solutions found when looking for rings using this particular node to initiate the analysis. In order to calculate $P_{N_{\max}}(n)$ and $P_{N_{\min}}(n)$ the longest and the shortest path have to be determined among these different solutions.
$P_{N_{\max}}(n)$ and $P_{N_{\min}}(n)$ have values ranging between 0 and $P_N(n)$. The lowest and non equal to

	King / Guttman.	Primitive / Strong.
Cas a)	$\begin{bmatrix} 4/16 & 2/16 \\ 2/16 & 5/16 \end{bmatrix}$	$\begin{bmatrix} 4/16 & 4/16 \\ 4/16 & 7/16 \end{bmatrix}$

	All criteria.
Case b)	$\begin{bmatrix} 4/16 & 0/16 \\ 0/16 & 12/16 \end{bmatrix}$
Case c)	$\begin{bmatrix} 4/16 & 1/16 \\ 1/16 & 12/16 \end{bmatrix}$
Case d)	$\begin{bmatrix} 4/16 & 0/16 \\ 0/16 & 11/16 \end{bmatrix}$
Case e) $\rightarrow$ f)	$\begin{bmatrix} 4/16 & 2/16 \\ 2/16 & 12/16 \end{bmatrix}$
Case g) $\rightarrow$ i)	$\begin{bmatrix} 4/16 & 1/16 \\ 1/16 & 11/16 \end{bmatrix}$
n= ring with n nodes	$\begin{bmatrix} n4 & n6/n4 \\ n4/n6 & n6 \end{bmatrix}$

Table 5.6 *General connectivity matrix for the networks represented in figure [Fig. 5.22] and studied using the different definitions of rings.*

0 is of the form $\frac{1}{Nn}$, the highest and non equal to 1 is of the form $\frac{Nn-1}{Nn}$, where Nn represents the total number of nodes in the network. For the minimum ring size, s_{min}, existing in the network or found during the search, $P_{N_{min}}(s_{min}) = P_N(s_{min})$. In the same way for the maximum ring size, s_{max}, existing in the network or found during the search, $P_{N_{max}}(s_{max}) = P_N(s_{max})$.

To clarify these informations it is possible to normalize $P_{N_{\max}}(n)$ and $P_{N_{\min}}(n)$ by $P_N(n)$. By reducing these values we obtain, for each size of rings, values independent of the total number of nodes Nn of the system. Then for a considered ring size the values only refer to the number of nodes where the search returns rings of this size:

$$P_{max}(n) = \frac{P_{N_{max}}(n)}{P_N(n)} \quad \text{and} \quad P_{min}(n) = \frac{P_{N_{min}}(n)}{P_N(n)}$$

The normalized terms $P_{\max}(n)$ and $P_{\min}(n)$ have values ranging between 0 and 1. The lowest

	$P_N(n)$	
n	King / Guttman.	Primitive / Strong.
Case a)		
4	4/16	4/16
6	5/16	7/16

	$P_N(n)$
n	All criteria.
Case b) → c)	
4	4/16
6	12/16
Case d)	
4	4/16
6	11/16
Case e) → f)	
4	4/16
6	12/16
Case g) → i)	
4	4/16
6	11/16

Table 5.7 *$P_N(n)$ - Proportion of nodes at the origin of at least one ring of size n for the networks presented in figure [Fig. 5.22].*

and non equal to 0 is of the form $\frac{1}{Nn}$, the highest and non equal to 1 is of the form $\frac{Nn-1}{Nn}$. For the minimum ring size, s_{min}, existing in the network or found during the search, $P_{min}(s_{min}) = 1$. In the same way for the maximum ring size, s_{max}, existing in the network or found during the search, $P_{max}(s_{max}) = 1$.
$P_{\max}(n)$ and $P_{\min}(n)$ give complementary informations to the ones obtained with $R_C(n)$ and $P_N(n)$ in order to distinguish and compare networks using ring statistics. We can illustrate this result by presenting the complete informations obtained with this method [Tab. 5.8] for the networks represented in figure [Fig. 5.22].
$P_{\max}(n)$ and $P_{\min}(n)$ give informations about the connectivity of the rings with each other as a function of their size. If a ring of size n is found using a particular node to initiate the search, $P_{\max}(n)$ gives the probability that this ring is the longest ring which can be found using this node to initiate the search. At the opposite, $P_{\min}(n)$ gives the probability that this ring is the shortest ring which can be found using this node to initiate the search.
Thereafter we will use the terms 'connectivity profile' to designate the results of a ring statistics analysis. This profile is related to the definition of rings used in the search and is made of the 4

King / Guttman.

n	$R_C(n)$	$P_N(n)$	$P_{\max}(n)$	$P_{\min}(n)$
Case a)				
4	1/16	4/16	0.5	1.0
6	2/16	5/16	1.0	0.6

Primitive / Strong.

n	$R_C(n)$	$P_N(n)$	$P_{\max}(n)$	$P_{\min}(n)$
Case a)				
4	1/16	4/16	0.5	1.0
6	2/16	7/16	1.0	3/7

All criteria.

n	$R_C(n)$	$P_N(n)$	$P_{\max}(n)$	$P_{\min}(n)$
Case b)				
4	1/16	4/16	1.0	1.0
6	2/16	12/16	1.0	1.0
Case c)				
4	1/16	4/16	0.75	1.0
6	2/16	12/16	1.0	11/12
Case d)				
4	1/16	4/16	1.0	1.0
6	2/16	11/16	1.0	1.0
Case e) $\rightarrow$ f)				
4	1/16	4/16	0.5	1.0
6	2/16	12/16	1.0	10/12
Case g) $\rightarrow$ i)				
4	1/16	4/16	0.75	1.0
6	2/16	11/16	1.0	10/11

Table 5.8 *Connectivity profiles results of the ring statistics for the networks presented in figure [Fig. 5.22].*

values defined in our method: $R_C(n)$, $P_N(n)$, $P_{\max}(n)$ and $P_{\min}(n)$.

The I.S.A.A.C.S. program provides access to the connectivity profile of the system under study and allows to choose the study the connectivity using all the different methods used to define a ring. Thus King's rings, Guttman's rings, Primitive rings as well as Strong rings analysis are available.

5.5.5 Bond deffects in ring statistics

5.5.5.1 ABAB and BABA rings

The ring statistics of amorphous networks are often focused on finding rings made of a succession of atoms with an alternation of chemical species, called ABAB rings. The most common examples come from the alternation of Si and O atoms (in silica polymorphs) or Ge and S (in GeS_2 polymorphs). These solids are usually built with tetrahedra (SiO_4 or GeS_4) therefore we study the network distribution of tetrahedra.
The ideal technique to setup the analysis of such systems is to choose the atoms of highest coordination to initiate the search, respectively Si in SiO_2 and Ge in GeS_2. In most cases all rings can be found using this method. Nevertheless we can demonstrate that some solutions, so some rings, can be ignored by this analysis. This is highlighted in figure [Fig. 5.24] which represents a cluster of atoms isolated from an AB_2 amorphous network.

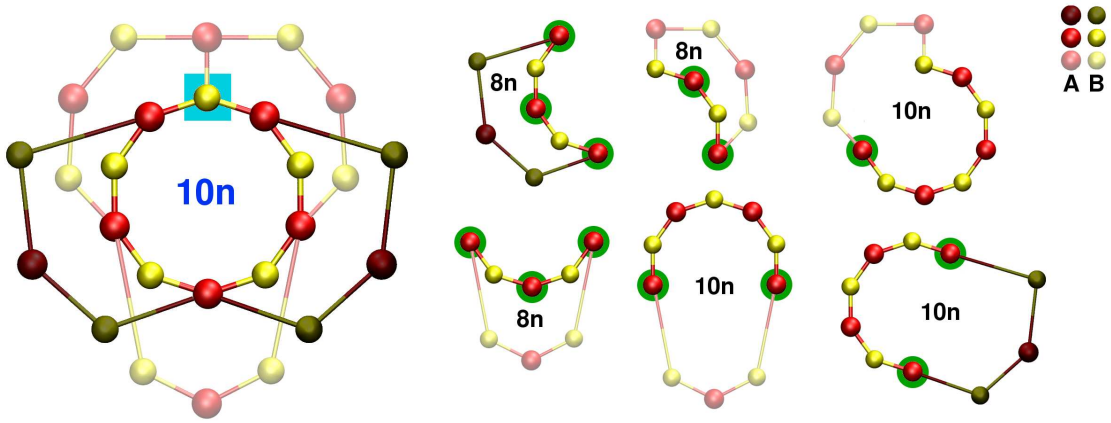


Figure 5.24 *Cluster of atoms isolated from an AB_2 amorphous network. A bond defect is located on an atom of the chemical species B (blue square). When looking for King's shortest paths [S. 5.5.2.1] using the chemical species A to initiate the search the central ring with 10 nodes is ignored. However among the solutions of the analysis (with the initial nodes circled in green) other rings with 10 nodes are found in the network.*

We can see that this piece of network is characterized by a bond defect. An atom of the B species appears to be over-coordinated by three atoms of the A species. When looking for rings, using the King's criterion [S. 5.5.2.1] and initiating the search using the A atoms, the central ring with 10 nodes is ignored. Nevertheless other rings with 10 nodes are found and stored as solutions of the analysis. In order to find the central ring the search has to be initiated from the overcoordinated B atom.
By analogy with the terminology ABAB this ring can be called a BABA ring. Indeed the alternation of chemical species is well respected. Therefore it is legitimate to question the relevance of the analysis without this result. In other words we have to check out if this BABA ring is, or

not, an ABAB ring.
The properties of this ring meet the definition and can therefore improve the description of the connectivity of the network. This kind of coordination defect [Fig. 5.24] is uncommon in vitreous silica [29, 35], nevertheless it is frequent in chalcogenide glasses [36, 37].

5.5.5.2 Homopolar bonds

In amorphous materials the homopolar bond defects can have a significant influence on the ring statistics. This is true in particular for AB_2 chalcogenide glasses. Figure [Fig. 5.25] illustrates standard cases that may be encountered when looking for rings in an AB_2 system which contains homopolar bonds.

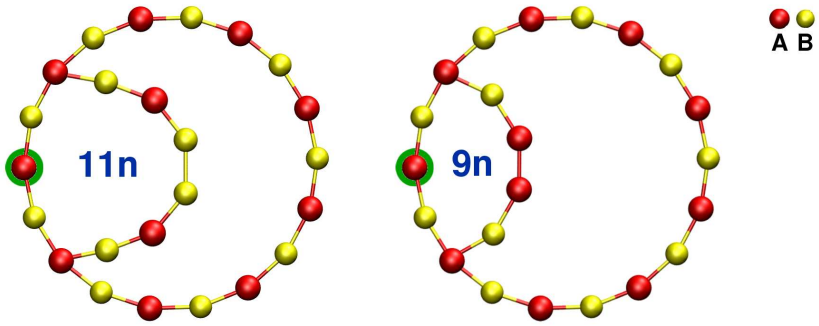


Figure 5.25 *Illustration of the influence of homopolar bonds in ABAB rings: in both examples the smallest rings found when initiating the search using the circled nodes (green color) contain an homopolar bond A-A or B-B.*

The smallest rings found when initiating the search using the circled nodes (green color) are not ABAB rings. Therefore their size must be given using the total number of nodes. In figure [Fig. 5.25] the smallest rings are a ring with 9 nodes and a ring with 11 nodes containing respectively an A-A and a B-B homopolar bond. These rings are significantly smaller than the shortest ABAB ring with 18 nodes that may be found when looking for rings using the same green-circled nodes to initiate the analysis [Fig. 5.25].

The I.S.A.A.C.S. program provides options to take into account or avoid A-B-A-B rings as well as homopolar bonds.

5.5.6 Number of rings not found and that "potentially exist"

One of the first information it is possible to extract from ring statistics, except the number of rings, is the number of rings not found by the analysis. Indeed calculation times do strongly depend on the maximum search depth, ie. the maximum size of a ring. To carry out the analysis this value has to be chosen to get the best possible compromise between CPU time and quality of the description.

Nevertheless whatever this limiting value is, some rings of a size bigger than the maximum search depth may not be found by the analysis. In the King 5.5.2.1 and the Guttman's criteria 5.5.2.2 it is possible to evaluate the number of "potentially not found" rings or rings that "potentially exist".
Thus for a given atom **At** we can consider that a closed path exists and is not found:

1. If the atom **At** has at least 2 nearest neighbors
2. If no closed path is found:
    a- Starting from one neighbor to go back on the considered atom **At** (Guttman's criterion)
    b- Between one couple of neighbors of the atom **At** (King's criterion)
3. If the 2 nearest neighbors of the atom **At** have at least 2 nearest neighbors (to avoid non bridging atoms)

Thus if during the analysis these 3 conditions are full filled (1, 2-a, 3 for the Guttman's criterion, and 1, 2-b, 3 for the King's criterion) then we can say that this analysis has potentially missed a ring between the neighbors of atom **At**. The smaller this number of "potentially" missed rings will be the better this analysis will be and the better the description of the connectivity of the material studied will be. The term "potentially" has been chosen because the method only allows to avoid first neighbor non bridging atoms.

Following this method I.S.A.A.C.S. gives access to the number of rings that "potentially exist" and not found during the analysis.

5.6 Invariants of spherical harmonics as atomic order parameters

Invariants formed from bond spherical harmonics allow to obtain quantitative informations on the local atomic symmetries in materials. The analysis starts by associating a set of spherical harmonics with every bond linking an atom to its nearest neighbors. For a given bond defined by a vector $\vec{r}$ a spherical harmonic may be defined as:

$$Q_{lm}(\vec{r}) \;=\; Y_{lm}\langle\theta(\vec{r}),\psi(\vec{r})\rangle \tag{5.28}$$

where $Y_{lm}(\theta,\psi)$ is the spherical harmonic associated to the bond, θ and ψ are the angular components of the spherical coordinates of the bond which cartesian coordinates are defined by $\vec{r}$.

Because the Q_{lm} for a given l can be scrambled by changing to a rotated coordinate system, it is important to consider rotational invariant combinations, such as [23, 38]:

$$Q_l \;=\; \left[\frac{4\pi}{2l+1}\sum_{m=-l}^{l}\left|\bar{Q}_{lm}\right|^2\right]^{1/2} \tag{5.29}$$

where $\bar{Q}_{lm}$ is defined by:

$$\bar{Q}_{lm} \;=\; \langle Q_{lm}(\vec{r})\rangle \tag{5.30}$$

and represents an average of the $Y_{lm}(\theta,\psi)$ over all $\vec{r}$ vectors in the system whether these vectors belong to the same atomic configuration or not. Just as the angular momentum quantum number, l, is a characteristic quantity of the 'shape' of an atomic orbital, the quantity Q_l is a rotationally invariant characteristic value of the shape/symmetry of a given local atomic configuration (if the average is not taken on all bonds of the system but only within a given configuration) or an average of such values for a set of configurations. Thus it is possible to compare Q_l's computed for well known crystal structures (e.g. FCC, HFC ...) and some local atomic configurations in a material's model. The results of the comparison gives information for the presence/absence of a particular local atomic symmetry.

I.S.A.A.C.S. allows to compute the average Q_l's for each chemical species as well as the average Q_l's for a user specified local atomic coordination.

5.7 Bond valence sums

The bond valence method (or bond valence sum) (not to be mistaken with the valence bond theory in quantum chemistry) is a technique used in coordination chemistry to estimate the oxidation/valence states of atoms.
The basic idea is that the valence V_i of an atom i is the sum of the individual bond valences v_{ij}

of the N_i surrounding atoms:

$$V_i = \sum_{j=1}^{N_i} v_{ij} \tag{5.31}$$

The individual bond valences v_{ij} are calculated using:

$$v_{ij} = e^{\left(\frac{R_0 - R_{ij}}{B}\right)} \tag{5.32}$$

or

$$v_{ij} = \left(\frac{R_{ij}}{R_0}\right)^{-N} \tag{5.33}$$

R_{ij} is the computed bond length between atoms i and j, R_0, B and N are tabulated [24–26].

I.S.A.A.C.S. allows to compute the average bond valence sums for each chemical species as well as the average bond valence sums for a user specified local atomic coordination.

5.8 Mean square displacement of atoms

Atoms in solids, iquids and gases move constantly at any given temperature, i.e. they are subject to a "thermal" displacement from their average positions. This displacement is particularly important in the case of a liquids. Atomic displacement does not follow a simple trajectory: "collisions" with other atoms render atomic trajectories quite complex shaped in space.
The trajectory followed by an atom in a liquid resembles that of a pedestrian random walk. Mathematically this represents a sequence of steps done one after another where each step follows a random direction which does not depend on the one of the previous step (Markov's chain of events).
In the case of a one-dimensional system (straight line) the displacement of the atom will therefore be either a forward step (+) or a backward step (-). Furthermore it will be impossible to predict one or the other direction (forward or backward) since they have an equal probability to occur.
One can conclude that the distance an atom may travel is close to zero. Nevertheless if we choose not to sum the displacements themselves (+/-) but the square of these displacements then we will end up with a non-zero, positive quantity of the total squared distance traveled. Consequently this allows to obtain a better evaluation of the real (square) distance traveled by an atom.
The **M**ean **S**quare **D**isplacement MSD is defined by the relation:

$$MSD(t) = \langle \mathbf{r}^2(t) \rangle = \left\langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \right\rangle \tag{5.34}$$

where $\mathbf{r}_i$(t) is the position of the atom i at the time t, and the $\rangle$ $\langle$ represent an average on the time steps and/or the particles.

However, during the analysis of the results of molecular dynamics simulations it is important to subtract the drift of the center of mass of the simulation box:

$$MSD(t) = \left\langle |\mathbf{r}_i(t) - \mathbf{r}_i(0) - [\mathbf{r}_{cm}(t) - \mathbf{r}_{cm}(0)]|^2 \right\rangle \tag{5.35}$$

where $\mathbf{r}_{cm}$(t) represents the position of the center of mass of the system at the time t.
The MSD also contains information on the diffusion of atoms. If the system is solid (frozen) then MSD "saturate", and the kinetic energy is not sufficient enough to reach a diffusive behavior. Nevertheless if the system is not frozen (e.g. liquid) then the MSD will grow linearly in time. In such a case it is possible to investigate the behavior of the system looking at the slope of the MSD. The slope of the MSD or the so called diffusion constant D is defined by:

$$D = \lim_{t\to\infty} \frac{1}{6t} \langle \mathbf{r}^2(t) \rangle \tag{5.36}$$

I.S.A.A.C.S. provides access to the several MSD related functions:

- MSD for each chemical species with autocorrelation on all the dynamics
- MSD for each chemical species without autocorrelation on all the dynamics (step by step)
- Directional MSD (x, y, z, xy, xz, yz) for each chemical species with autocorrelation on all the dynamics
- Directional MSD (x, y, z, xy, xz, yz) for each chemical species without autocorrelation on all the dynamics (step by step)
- Drift of the center of mass (x, y, z)
- Correction applied to correct the drift of the center of mass in equation [Eq. 5.35] (x, y, z)

The chemical properties database in I.S.A.A.C.S.

A database of chemical/physical properties is included in the I.S.A.A.C.S. program, this appendix presents these properties (atomic radii, x-ray and neutron scattering lengths) as well as the references of the articles from which this information was obtained.
Note that the data presented in this appendix is available for download on the web sites of the program:
http://isaacs.sourceforge.net/phys/chem.html
http://people.cst.cmich.edu/petko1vg/isaacs/phys/chem.html

A.1 Atomic radii

A.1.1 Covalent radii

Figure [Fig. A.1] illustrates the covalent radii used in I.S.A.A.C.S. see [39] for details.

A.1.2 Ionic radii

Figure [Fig. A.2] illustrates the ionic radii used in I.S.A.A.C.S. see [40] for details.

A.1.3 VDW radii

Figure [Fig. A.3] illustrates the Van Der Waals radii used in I.S.A.A.C.S. see [41] for details.

A.1.4 Shannon radii in crystal

Figure [Fig. A.4] illustrates the atomic radii in crystals as compiled by Shannon used in I.S.A.A.C.S. see [42, 43] for details.

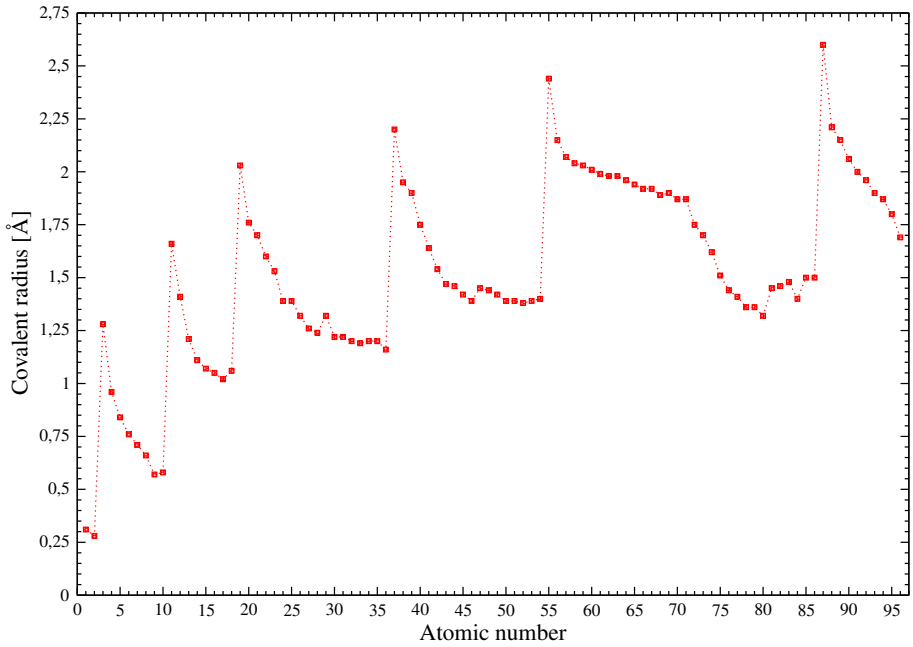


Figure A.1 *Covalent radii used in the I.S.A.A.C.S. program.*

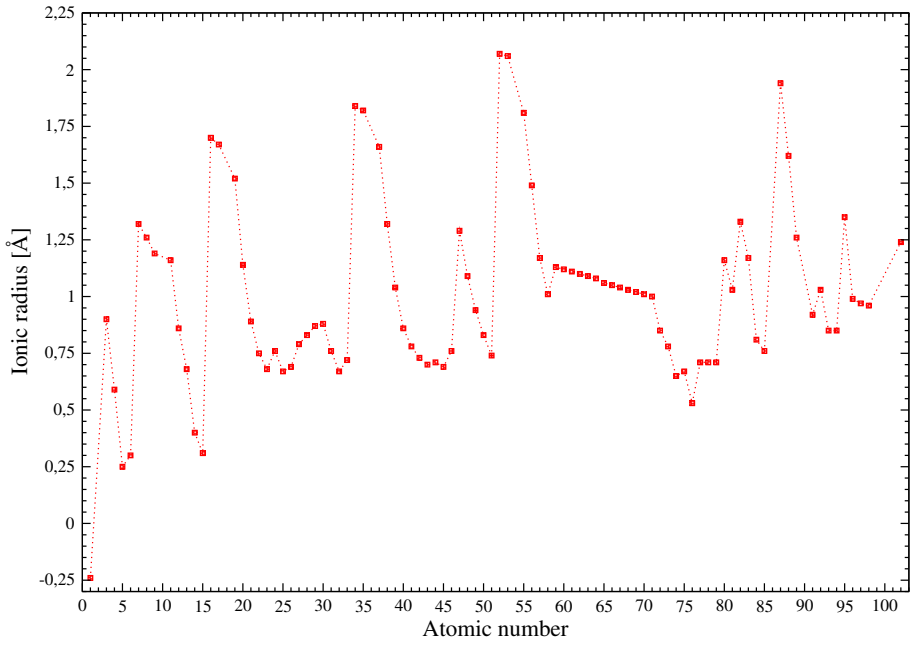


Figure A.2 *Ionic radii used in the I.S.A.A.C.S. program.*

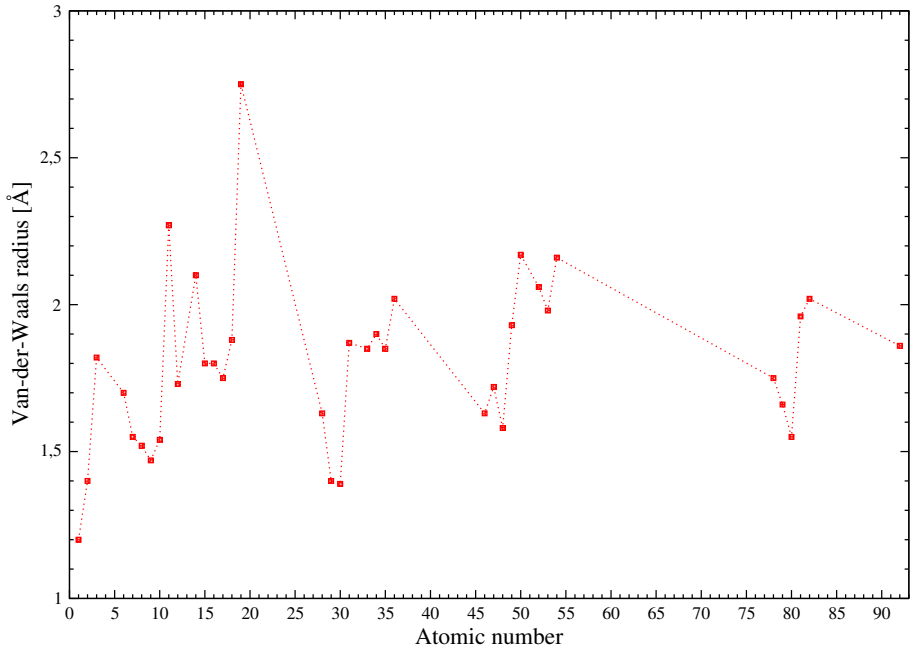


Figure A.3 *Van Der Waals radii used in the I.S.A.A.C.S. program.*

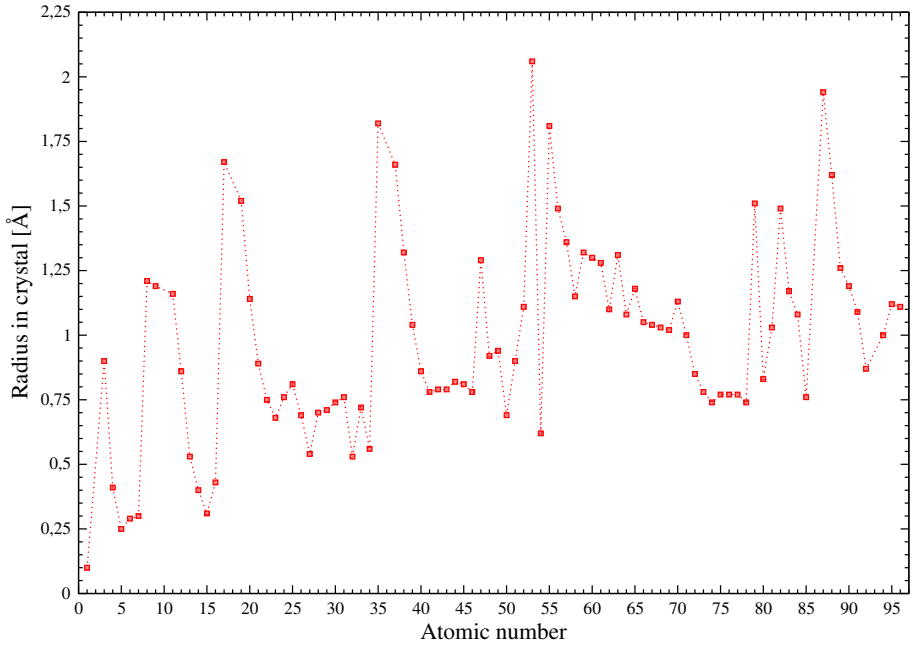


Figure A.4 *Shannon radii in crystals used in the I.S.A.A.C.S. program.*

A.2 Neutron scattering lenghts

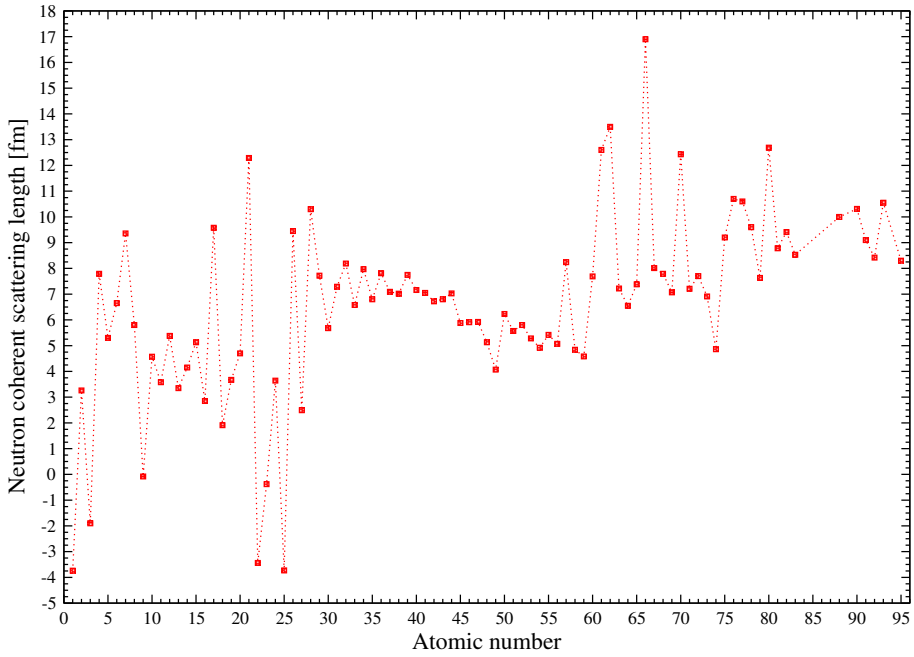


Figure A.5 *Neutron scattering lengths used in the I.S.A.A.C.S. program.*

Figure [Fig. A.5] illustrates the neutron scattering lengths used in I.S.A.A.C.S. see [44, 45] for details.
The atomic numbers are used for the x-ray scattering lengths.

Bibliography

[1] The GTK+ Team (2009).
http://www.gtk.org/.

[2] World Wibe Web Consortium - W3C. Extensible markup language (xml) (2009).
http://www.w3.org/XML/.

[3] Minnesota Supercomputer Center. Xmol (1993).
http://www.hpc.unm.edu/ chem/xmol/xmol.html.

[4] CambridgeSoft. Chem3d (2009).
Cambridge Scientific Computing, Cambridge Massachusetts, USA.

[5] http://www.cpmd.org.

[6] http://cms.mpi.univie.ac.at/vasp/.

[7] M.P. Allen and D.J. Tildesley. *Computer simulation of liquids*. Oxford science publications (1987).

[8] P.S. Salmon. *J. Non-Cryst. Solids.*, **353**:2959–2974 (2007).

[9] M.T. Dove, M.G. Tucker, and D.A. Keen. *Eur. J. Mineral.*, **14**:331–348 (2002).

[10] B. Thijsse. *J. Appl. Cryst.*, **17**:61–76 (1984).

[11] T. E. Faber and Ziman J. M. *Phil. Mag.*, **11**(109):153–173 (1965).

[12] N. W. Ashcroft and D. C. Langreth. *Phys. Rev.*, **156**(3):685–692 (1967).

[13] N. W. Ashcroft and D. C. Langreth. *Phys. Rev.*, **159**(3):500–510 (1967).

[14] N. W. Ashcroft and D. C. Langreth. *Phys. Rev.*, **166**(3):934 (1968).

[15] A. B. Bhatia and D. E. Thornton. *Phys. Rev. B*, **2**(8):3004–3012 (1970).

[16] S. V. King. *Nature*, **213**:1112 (1967).

[17] D. S. Franzblau. *Phys. Rev. B*, **44**(10):4925–4930 (1991).

[18] L. Guttman. *J. Non-Cryst. Solids.*, **116**:145–147 (1990).

[19] K. Goetzke and H. J. Klein. *J. Non-Cryst. Solids.*, **127**:215–220 (1991).

[20] X. Yuan and A. N. Cormack. *Comp. Mat. Sci.*, **24**:343–360 (2002).

[21] F. Wooten. *Acta Cryst. A*, **58**(4):346–351 (2002).

[22] S. Le Roux and P. Jund. *Comp. Mat. Sci.*, **49**:70–83 (2010).

[23] P. Steinhardt, D.R. Nelson, and M. Ronchetti. *Phys. Rev. B*, **28**(2):784–805 (1983).

[24] I.D. Brown (2006).
http://www.ccp14.ac.uk/ccp/webmirrors/i_d_brown/bond_valence_param/bvparm2006.cif.

[25] D. Altermatt and I.D. Brown. *Acta Cryst. B*, pages 244–247 (1985).

[26] I.D. Brown and R.D. Shannon. *Acta Cryst. A*, pages 266–282 (1973).

[27] V. Petkov, S. J. L. Billinge, S. D. Shastri, and B. Himmel. *Phys. Rev. Lett.*, **85**(16):3436–3439 (2000).

[28] Grace Development Team. Graphing, advanced computation and exploration of data (1996-2011).
http://plasma-gate.weizmann.ac.il/Grace/.

[29] J. P. Rino, I. Ebbsjö, R. K. Kalia, A. Nakano, and P. Vashishta. *Phys. Rev. B*, **47**(6):3053–3062 (1993).

[30] R. M. Van Ginhoven, H. Jónsson, and L. R. Corrales. *Phys. Rev. B*, **71**(2):024208 (2005).

[31] M. Cobb, D. A. Drabold, and R. L. Cappelletti. *Phys. Rev. B*, **54**(17):12162–12171 (1996).

[32] X. Zhang and D. A. Drabold. *Phys. Rev. B*, **62**(23):15695–15701 (2000).

[33] D. N. Tafen and D. A. Drabold. *Phys. Rev. B*, **71**(5):054206 (2005).

[34] S. Le Roux and P. Jund. *RINGS*, pages http://rings–code.sourceforge.net (2009).

[35] W. Jin, P. Vashishta, R. K. Kalia, and J. P. Rino. *Phys. Rev. B*, **48**(13):9359–9368 (1993).

[36] S. Blaineau and P. Jund. *Phys. Rev. B*, **69**(6):064201 (2004).

[37] S. Le Roux and P. Jund. *J. Phys.: Cond. Mat.*, **19**(19):169102 (2007).

[38] A. Baranyai, A. Geiger, P.R. Gartrell-Mills, K. Heinzinger, R. McGreevy, G. Pálinkás, and I. Ruff. *Chem. Soc. Faraday Trans. 2*, **83**(8):1335–1365 (1987).

[39] B. Cordero. *Dalton Trans.*, pages 2832–2838 (2008).

[40] J. C. Slater. *J. Chem. Phys.*, **41**(10):3199 (1964).

[41] A. Bondi. *J. Phys. Chem.*, **68**(3):441 (1964).

[42] R.D. Shannon and C.T. Prewitt. *Acta Cryst.*, **B25**:925 (1969).

[43] R.D. Shannon. *Acta Cryst.*, **A23**:751–76A (1976).

[44] http://www.ncnr.nist.gov/resources/n-lenghs/.

[45] V.F. Sears. *Neutrons News*, **3**(3):26–37 (1992).

This document has been prepared using the Linux operating system and free softwares:

The text editor	"gVim"
The GNU image manipulation program	"The Gimp"
The WYSIWYG plotting tool	"Grace"
And the document preparation system	"LaTeX 2ε".