

## Local structure studies using the pair distribution function

Pierre Bordet

Univ. Grenoble Alpes, Inst NEEL, 38042 Grenoble, France

CNRS, Inst NEEL, 38042 Grenoble, France

**Abstract.** The pair distribution analysis method is a fast spreading structural analysis method allowing to go beyond classical crystallographic analysis by providing quantitative information about local as well as meso-structure. It based on powder diffraction data fourier transformed to direct space. We will present here the main characteristics of the method, and its domain of application.

### 1. Introduction

The knowledge of the local atomic arrangement of materials is of primary importance for the understanding of their physical/chemical properties. Indeed, these are in general directly related to short- or middle-range interactions between atoms, the resultant of which will depend on the structural arrangement. In case of crystals, this arrangement can be described by giving the unit cell dimension, the space group symmetry and the atomic positions and atomic displacement parameters (a.d.p.) for the (generally) small set of atoms constituting the asymmetric unit. In a diffraction experiment using X-rays, neutrons or electrons, due to the lattice periodicity, all the diffracted intensity will be concentrated into Bragg reflections. The positions of these reflections informs about the crystal structure symmetry and unit cell dimensions, their intensities can be used to retrieve the atomic arrangement of the asymmetric unit, and their widths and shapes carry information about the sample microstructure, especially in the case of powders. The nowadays crystallographic analysis methods are able to provide solution and refinement of crystal structures with incomparable accuracy, building the basis of our understanding of materials in various fields of solid state physics, chemistry and biology. However, all compounds are not found as single crystals or well crystallized powders, and crystallographers had to develop original methods in order to gain information on the structural arrangements of materials in the liquid or amorphous state, in the form of nano-crystalline powders and in general for compounds where the local structure is substantially different from the average structure that can be determined by the measurement of Bragg reflections.

One of these methods is the analysis of the pair distribution function (PDF). Experimentally, this function is obtained in a way similar to a powder diffraction experiment. The waves scattered by each pair of atoms in the sample will interfere and the sum of these signals over all the pairs will yield the experimental diffraction pattern which is then directly related to the distribution of atoms in the direct space. In the general case, scattering can occur at any diffraction angle. For a crystalline material, the

This is an Open Access article distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

periodicity of the atomic arrangement will lead to the presence of Bragg peaks where the scattered intensity is concentrated. In other cases, the analysis of the full scattered intensity distribution allows to retrieve the maximum amount of information that can be obtained about the atomic arrangement, that is, the pair distribution function itself. This function is defined in direct space as describing the probability of finding a pair of atoms at a given interatomic distance  $r$ . It can thus be simply calculated for any given model of the sample atomic arrangement. In order to be compared to the PDF of the model, the experimental diffraction pattern must be transformed into direct space, using essentially a Fourier transform. To yield an experimental PDF of sufficient quality to tackle complex structural problems, the diffraction pattern must fulfill requirements such as a large enough value of  $Q_{\max}$ , (i.e. the maximum  $Q = 4\pi \cdot \sin(\theta)/\lambda$  value of the measurement) and high counting statistics. For a long time, these requirements have limited the use of the PDF analysis to the study of liquids or amorphous materials. The availability of hot reactor sources and spallation installations now provide neutron beams allowing to reach high  $Q_{\max}$  values. At the same time, the development of 3G-synchrotrons yields X-ray beams of very high energies and very strong intensities which allow to carry out PDF measurements in a few seconds. This was accompanied by the development of data collection and analysis tools and software which make the technique usable by a vast community of scientists. The PDF analysis is now applied to a wide variety of problems ranging from mineralogy [1] to solid state physics [2]. In the following, we will give a brief introduction to PDF analysis method and how it can be used to provide quantitative structural and microstructural information for compounds regardless of their crystalline/amorphous state.

## 2. What is the PDF?

Historically, as described in the books by Guinier [3] or Warren [4], the pair distribution function analysis was used to characterize the local order in amorphous compounds using data from wide angle X-ray scattering. Indeed, this function is somehow a histogram of interatomic distances in a compound, regardless of its crystalline or amorphous state. Depending on the authors and the communities, there are different definitions of the pair distribution function. We will not go into the details here, the interested reader can report to the article by D. Keen [5] where relationships between different definitions are explained. We will here follow the formalism developed by Egami and Billinge [6].

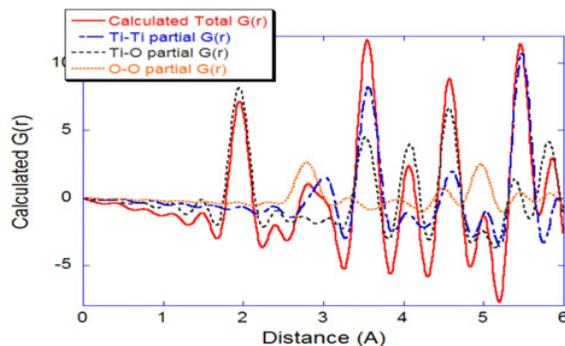
The PDF  $G(r)$  yields the probability of finding a pair of atoms separated by a distance  $r$ . It can be calculated from a structural model describing the distribution of atoms in a sample according to the formula:

$$G(r) = \frac{1}{r} \sum_i \sum_j \left[ \frac{b_i b_j}{\langle b \rangle^2} \delta(r - r_{ij}) \right] - 4\pi r \rho_0 \quad (1)$$

where the sum is propagated throughout the atoms contained in the structural model, separated by a distance  $r_{ij}$ .  $b_i$  is the scattering power of atom  $i$  (its Fermi length for neutrons, its form factor for X-rays or electrons);  $\langle b \rangle$  is the average scattering power of the sample;  $\rho_0$  is the numerical density of the compound, i.e. the number of atoms per unit volume.

This function will present peaks for values of  $r$  corresponding to interatomic distance in the model, the intensity of these peaks is proportional to the product of the scattering factors of the atoms forming the pair. All the contributions of all pairs of atoms add up in the PDF. However, it is possible to calculate partial PDFs by taking into account selected atoms only (Fig. 1).

It is worth noting that equation (1) is strictly valid only for neutron scattering for which the Fermi length  $b$  is independent of  $Q$  in reciprocal space so that  $\langle b \rangle$  can be calculated unambiguously. For X-rays and electrons, the form factors are  $Q$ -dependent and the shape of  $f(Q)$  depends on the element considered so that Eq. (1) is only an approximation [7]. However, in practice the error introduced is small enough for quantitative analysis of X-rays PDFs to remain valid.



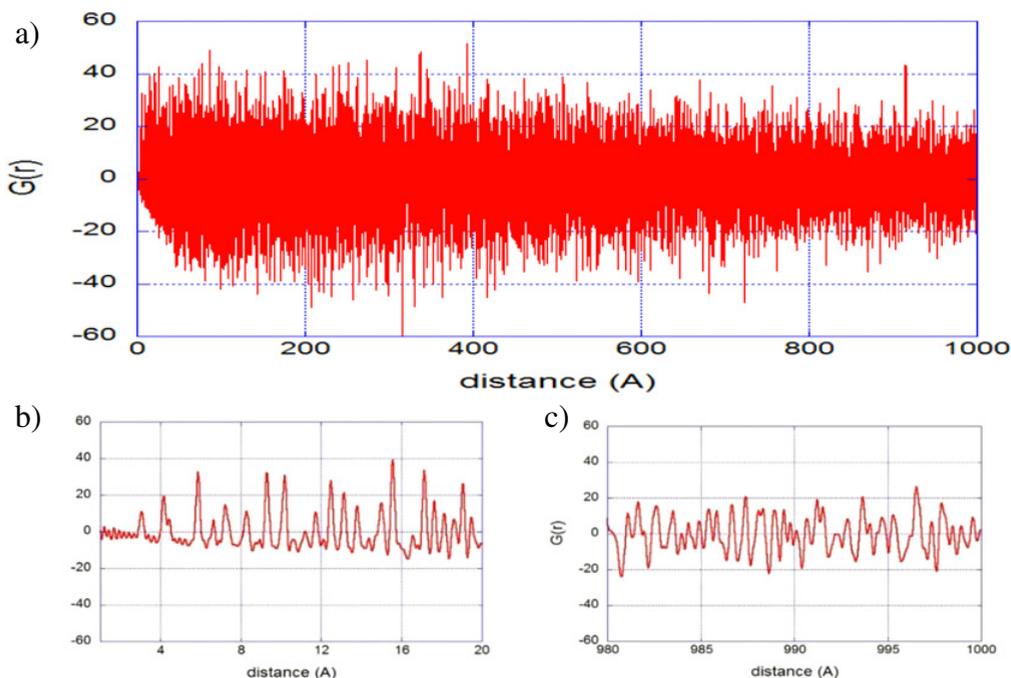
**Figure 1.** Total and partial PDFs calculated for TiO<sub>2</sub> Rutile.

The position of the peaks of the PDF reflects the distribution of distances in the material. In principle, the contribution of each pair consists in a  $\delta$ -function. For a completely random distribution of atoms, the PDF would then be flat and equal to 0 for any  $r$ . The second term  $-4\pi r \rho_0$  in equation (1) is a normalization term introduced to compensate for the increase of the number of atomic pairs with  $r$ , given the numerical density  $\rho_0$ . Therefore, the frequent (i.e. more than average) presence or absence of interatomic distances appears as fluctuations of the PDF about zero. For short distances corresponding to close atomic neighbors, isolated peaks can be observed (as for example at  $\sim 1.9$  or  $3.5$  Å in Fig. 1). The width of the peaks depends on the distribution of the distances around their average value, either due to thermal vibrations or a statistical distribution in the case of structural disorder.

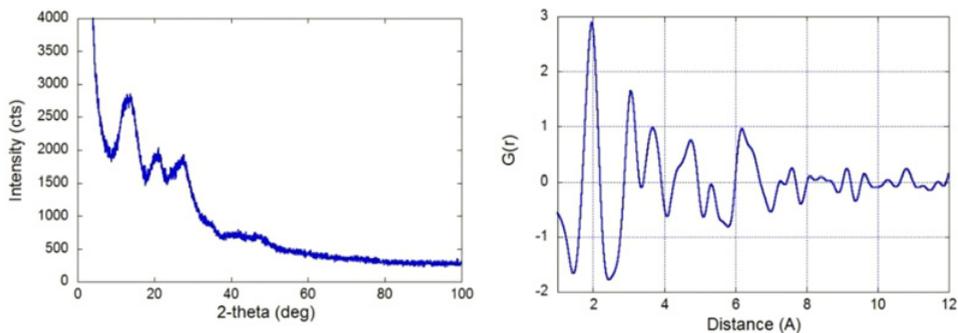
For a crystalline solid, the distribution of interatomic distance is defined by the coordinates of atoms, the atomic displacement parameters and the periodicity of the unit cell. The PDF will yield a peak at each of these distances, the width of the peak being related to the relative atomic displacement parameters of the atoms forming the pair, and its surface to the product of their scattering powers. In principle the PDF will show such peaks up to the largest interatomic distances in the crystal. For example, Fig. 2 displays the experimental PDF of LaB<sub>6</sub> obtained from high resolution synchrotron powder diffraction on the ESRF ID31 beam line up to 1000 Å.

Obviously, structural information can be retrieved from the PDF. For a well ordered and crystallized solid it is in fact equivalent to a 1D projection of the Patterson function. In that case, standard crystallographic techniques can be applied and are generally more convenient than the PDF description to study the structure.

In the case of nanoparticles, the sets of coherently scattering atoms are limited to the volume of the individual particles and therefore no peak in the PDF will be observed above the longest interatomic distance within the particles. Above this distance, only randomly distributed interatomic distances between atoms belonging to different particles will exist, and the PDF will vanish due to the normalization term of Eq. (1). The PDF will thus provide a direct estimate of the particle size, plus the distribution of interatomic distances inside the particles. The result will be similar for amorphous compounds: the PDF will vanish with the loss of structural coherence induced by disorder. The case of titanium oxide nano-particles of a diameter less than 1 nm is exemplified in Fig. 3 for PDF data obtained using a laboratory diffractometer in Bragg-Brentano geometry and MoK $\alpha$  radiation. A particle size of  $\approx 8-9$  Å can be readily deduced by simple observation of the PDF; a model of atomic arrangement inside the particle can be checked and fitted to the data to provide a quantitative estimate of the interatomic distances and ultimately of the structural arrangement.



**Figure 2.** Experimental PDF of  $\text{LaB}_6$  measured on the ID31 beam line of ESRF up to 1000 Å.



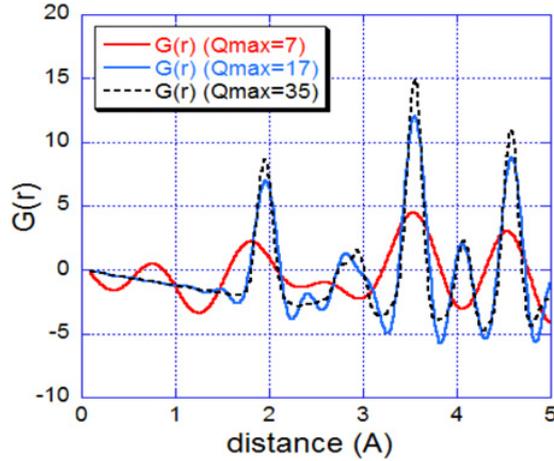
**Figure 3.** Diffraction pattern (left) and PDF (right) for  $\text{TiO}_2$  based nanoparticles.

### 3. How to measure the PDF

Experimentally, the PDF is obtained from a powder diffraction pattern by realizing the Fourier transform as:

$$G(r) = 4\pi r[\rho(r) - \rho_0] = \frac{2}{\pi} \int_0^{\infty} Q[S(Q) - 1] \sin(Qr) dQ \quad (2)$$

where  $\rho(r)$  is the microscopic pair density,  $S(Q)$  is the total structure function, i.e. the normalized coherent scattered intensity. It is worth noting that the whole diffraction pattern is used to obtain  $G(r)$ , not just the Bragg peak intensities as for a classical crystallographic analysis. Therefore the contributions of disorder, defects, etc. which contribute to the pattern through diffuse scattering outside of the Bragg peaks, will also be included in  $G(r)$ . Obtaining  $S(Q)$  from diffraction data requires to perform a set



**Figure 4.** Comparison of theoretical PDFs of rutile calculated for increasing values of  $Q_{\max}$ .

of corrections to eliminate all incoherent or inelastic contributions (Compton, fluorescence in the case of X-rays, Placzek and incoherent scattering in the case of neutrons), the signal coming from the environment (sample holder, cryostat, air scattering, etc.), as well as the effects of the absorption, multiple scattering, etc.. These corrections require additional measurements of these contributions (empty sample holder, etc.). In the case of neutron diffraction, the corrected intensity  $I$  is obtained with an additional measurement of a vanadium sample used for calibration by calculating:

$$I = \frac{(S - S_B) - \alpha(C - C_B)}{V - V_B} \quad (3)$$

where  $S$  stands for sample,  $C$  for container,  $V$  for vanadium and the suffix  $B$  refers to background measurements. Different background data may be necessary in the case where different measurement conditions (slits, etc.) are used for the three sets of data  $S$ ,  $C$  and  $V$ .

The self-scattering contribution is then subtracted and the pattern is normalized by  $\langle b \rangle^2$ , resulting in  $S(Q)$ . Then the function  $F(Q) = Q(S(Q) - I)$  is Fourier transformed according to Eq. (2). User-friendly programs to calculate the PDF from experimental patterns are now available for neutrons as well as X-rays and the procedure has become almost routine.

The calculation of the Fourier transform on a bounded domain will have consequences on the experimental PDF, it will be necessary to take into account when modeling by the function  $G(r)$ . These effects include the presence of termination wiggles, broadening of the peaks of the PDF and a decrease of their intensity at long distances. In principle, these effects can be taken into account in the calculation of the theoretical PDF. However, if one is interested in fine structural detail on a wide range of distances, it is necessary to minimize them. For that we must measure the powder diffraction pattern up to high values of  $Q$ , ideally well above the disappearance of a distinct scattering signal due to thermal and/or static disorder. Figure 4 shows a comparison between the PDFs simulated for values of  $Q_{\max}$  corresponding to the  $\text{CuK}\alpha$  ( $\approx 7 \text{ \AA}^{-1}$ ) and  $\text{MoK}\alpha$  ( $\approx 17 \text{ \AA}^{-1}$ ) radiations and for a value of  $Q_{\max}$  significantly higher ( $35 \text{ \AA}^{-1}$ ) corresponding to a wavelength of  $0.35 \text{ \AA}$  reachable at a synchrotron instrument. It is clear that the increase in the value of  $Q_{\max}$  provides access to much more detailed structural information. In principle, PDFs from patterns collected only up to only low  $Q_{\max}$  (e.g. using  $\text{CuK}\alpha$  radiations) may be used [8], however the PDF obtained are relatively featureless and dominated by spurious effects from the Fourier transform or poor normalization, which can lead to misinterpretations of the data. It is thus more favorable to use shorter wavelength radiations for PDF analysis [9]. Note that the angular

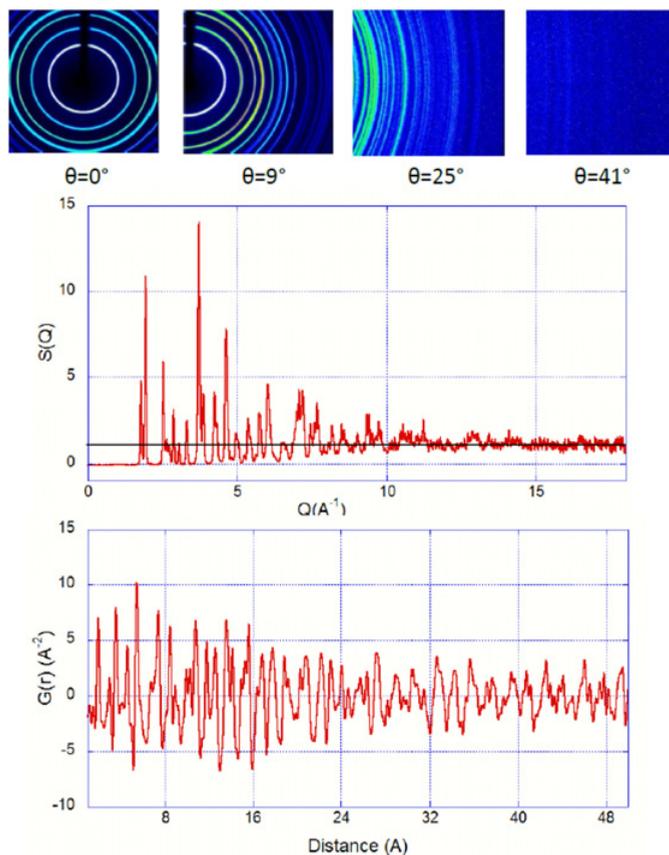
resolution of the diffractometer is less restrictive, though too low resolution in reciprocal space leads to a rapid decrease of the PDF intensity with distance in direct space.

Due to the decrease vs.  $Q$  of the scattering power in the case of X-rays, the normalization of  $S(Q)$  by dividing with  $\langle b \rangle^2$  will significantly amplify the data at large  $Q$ , which may be very noisy. It is therefore important to ensure sufficiently high counting statistics in the high  $Q$  range, otherwise a high noise level at high  $Q$  for  $F(Q) = Q(S(Q) - I)$  will introduce a spurious signal in the Fourier transform and therefore the PDF.

To obtain a powder diffraction pattern usable for PDF analysis, one must combine a high  $Q_{max}$  (meaning small wavelength/high energy radiation) and high counting statistics on a wide  $Q$  range. Both conditions can be fulfilled at synchrotron sources. Two main kind of set ups can be used at synchrotron sources: high resolution powder diffraction using a parallel beam and analyzer crystals (e.g. ID31 at ESRF, CRISTAL at SOLEIL, etc.). This has the advantage of eliminating the inelastic scattering and thus simplifies the corrections, but leads to yet long measurement times (several hours). The high  $Q_{max}$  and ultimate angular resolution of such an instrument allows obtaining PDFs up to several hundreds of Å (Fig. 2). The alternate set up relies on the access to very high energies ( $\sim 100$  keV) at which a flat plate 2D detector allows to collect data to high  $Q$  in a short time (for example at ESRF beam lines ID11, ID15, etc.) [10]. In this case, although the spatial resolution may be poorer, *in situ* studies, investigations of phase transitions, solidification, etc. become feasible.

Nevertheless, good quality PDF measurements are also possible with modern laboratory X-ray diffraction instruments. In order to reach high  $Q_{max}$ , it is necessary to use the MoK $\alpha$  ( $\lambda = 0.7107$  Å,  $Q_{max} = 17$  Å $^{-1}$ ) or better AgK $\alpha$  ( $\lambda = 0.5608$  Å,  $Q_{max} = 21.7$  Å $^{-1}$ ) radiation and measure up to the highest accessible angles ( $> 150^\circ$ ). The diffractometer must be set to optimize intensity, with relaxed monochromatization and a 1D detector which must remain efficient for the wavelength used. Either Bragg-Brentano or Debye-Scherrer geometry can be used. In the first case, a set of variable slits will help increasing the high angle intensity. For the latter case, a 2D detector can also be used. This can be done for instance by using a single crystal diffractometer equipped with a CCD camera. Diffraction images taken at increasing values of  $2\theta$  are converted to a  $I(2\theta)$  diffraction pattern which is used to compute the PDF (Fig. 5). In any case, the collection time for measuring a PDF with a modern laboratory X-ray diffractometer is of the order of one day, depending on the instrument performances and the sample scattering power.

Although in principle the PDF can be obtained with X-rays and neutrons, the latter have seen a strong development in the recent years and present some advantages. High  $Q$  data can readily be obtained at neutron spallation sources, (for example at ISIS, in England) or at hot sources in neutrons reactors and dedicated instruments such as NPDF at LANL, USA, are now available. The characteristic of nuclear neutron scattering play an important role in calculating and measuring the PDF. We have seen that the PDF peak height is proportional to the product of the scattering powers of the atoms forming the pair. It is well-known that for neutrons the scattering power due to the nuclear interaction between the neutron and the atom nucleus, varies in a very different way than that of X-rays, coming from the electromagnetic interaction between X-rays and the electron cloud. Two consequences are that for X-rays, the scattering power – the so-called atomic form factor – is mainly proportional to  $Z$  and decays rapidly with increasing  $Q$ , whereas for neutrons the scattering power – the Fermi length – varies in a seemingly erratic manner with  $Z$  and is constant as function of  $Q$ . This is used in diffraction studies to increase scattering contrast between atoms of neighboring  $Z$ 's and two detect low  $Z$  atoms (typically H or O) which have a large fermi length. The same is true for the PDF. Figure 6 shows the comparison of the PDFs calculated for LaMnO $_3$  up to 10 Å for X-rays (black non continuous line) and neutrons (blue line). The first peak just below 2 Å corresponds to the first nearest neighbor Mn-O distance. It is a quite small peak for X-rays, due to the relatively small scattering power of oxygen. It is however much stronger in the case of neutrons, and negative, because the Fermi lengths of Mn and O are both strong in absolute value, and the Fermi length of Mn is negative. The Mn-O peak height being proportional the product of these

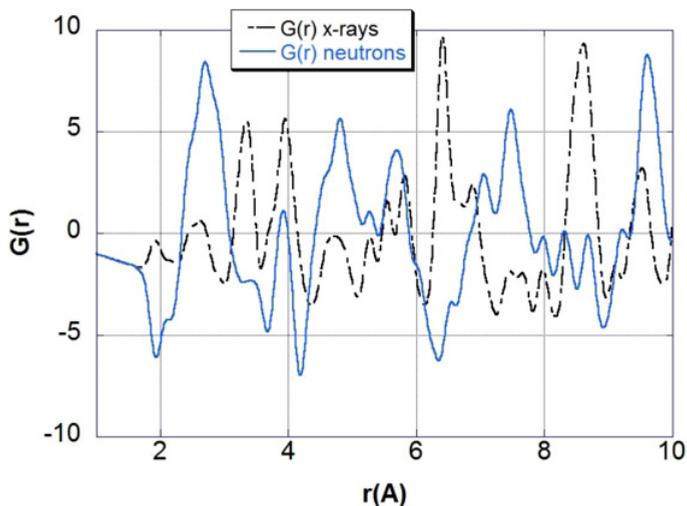


**Figure 5.** Measurement of the PDF of a mixture of  $\text{TiO}_2$  oxides using a laboratory single crystal diffractometer with a CCD camera. Top: selection of diffraction images taken at increasing  $\theta$  angles. Middle: the  $S(Q)$  scattering function obtained after integration and summation of images. Bottom: the corresponding PDF.

two Fermi length will then be negative, something that does never happen for X-rays. A consequence of the scattering power for neutrons being constant in  $Q$ , is that substantial diffracted intensity can still be connected at quite high  $Q$  values, the decay of the diffraction pattern being then essentially due to the effect of thermal and static disorder. On one hand this will provide more information after Fourier transformation to the PDF, but care must be taken to collect data at even higher  $Q$ s where most of the structural contributions have vanished.

#### 4. Parameters influencing the PDF

The appearance of the PDF may depend strongly on a number of parameters related to the structure of microstructure of the sample, but also on the experimental conditions. We have already seen the strong influence of the  $Q_{\text{max}}$  value and the nature of the scattered radiation (neutrons or X-rays). Here we will discuss a few of the most important remaining parameters, starting from experimental ones, which may be helpful to design a well-adapted experiment when dealing with a particular problem.



**Figure 6.** Comparison of the PDFs calculated for  $\text{LaMnO}_3$  for X-rays and neutrons, everything else unchanged.

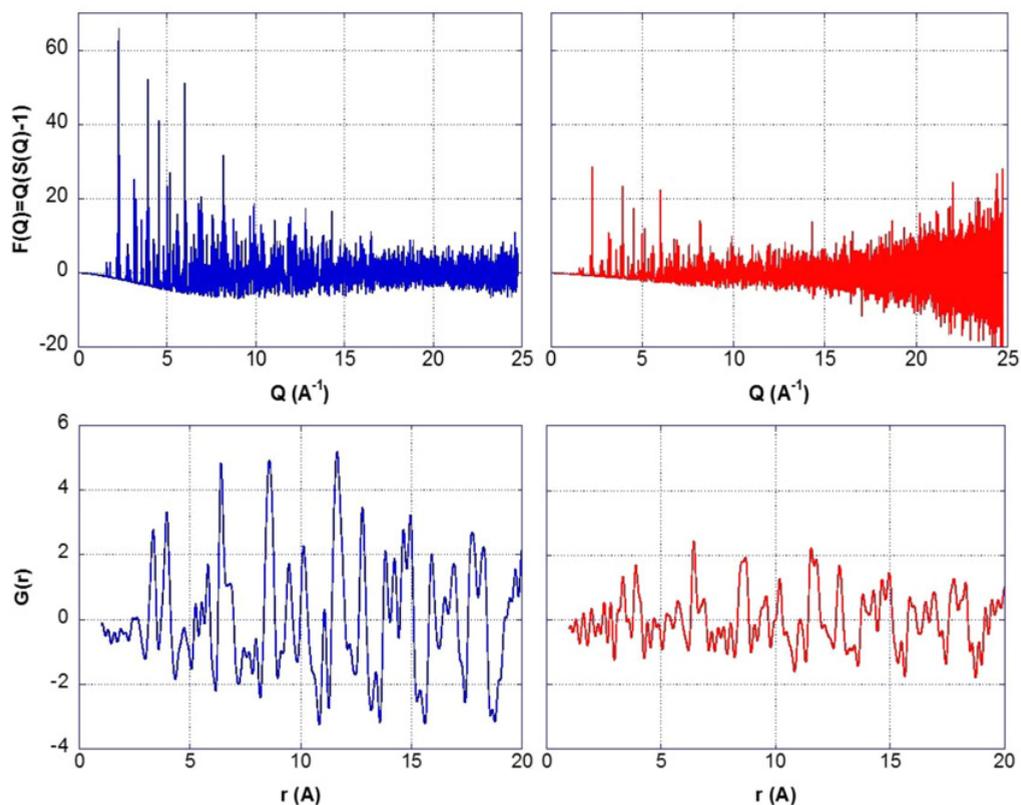
#### 4.1 Counting statistics

We have seen how important it is to measure up to high  $Q_{\text{max}}$  values (i.e.  $20\text{--}30 \text{ \AA}^{-1}$ ), however at such values, visible features in the diffraction signal are usually extremely weak if not completely absent. It must be reminded here, that the normalization procedure will considerably enhance the high- $Q$  part of the diagram with respect to the low- $Q$  one. It is most obvious for X-rays where normalization implies dividing the pattern by  $f^2$ , where  $f$  is the form factor rapidly decaying with  $Q$ . This is also true for neutron since the signal is damped by thermal and disorder effects. In Fig. 7 below, the  $F(Q)$  and  $G(r)$  function are simulated in the same conditions except that on the right part the peak/background ratio was divided by a factor of 10. It is clear that  $F(Q)$  is largely dominated by noise at high  $Q$ , resulting in a poor quality of the obtained PDF. It must be emphasized that these diffraction data would be quite satisfactory for a classical Rietveld analysis generally using data up to  $7\text{--}8 \text{ \AA}^{-1}$ .

It is necessary to adapt a data collection strategy to avoid such detrimental effect. A way of doing is to increase the counting time with increasing  $Q$ , ideally in a way inversely proportional to the global decrease of intensity. This may lead to exceeding long experimental time for high  $Q$  data collections. Thus data collections requiring high- $Q$  data, i.e. dealing with rather well-crystallized samples, should be preferentially carried out on instrument with high incident flux, and using 1D or 2D detectors efficient at high energies. When not available, there remains the possibility to make use of damping functions during the extraction of the PDF. Such functions, such as a Gaussian or the Lorch function, strongly decrease with  $Q$  and when multiplied to  $F(Q)$  prior to Fourier transform, will damp the effects of high  $Q$  noise. Their main drawback is to increase the PDF peak.

#### 4.2 Instrumental resolution

Suppose the instrumental resolution in reciprocal space (e.g. the instrumental contribution to the Bragg peak widths) could be described by a Gaussian function of constant width  $H$ . The experimental diffraction pattern can then be regarded as the convolution of the “pure” pattern and this Gaussian resolution function. After Fourier transform, the PDF will then be obtained in direct space as the product of the “pure” PDF and the Fourier transform of the Gaussian resolution function, i.e. a Gaussian function of inverse width  $2\pi/H$ . The latter will tend to damp the amplitude of the PDF peaks at higher  $r$  values.

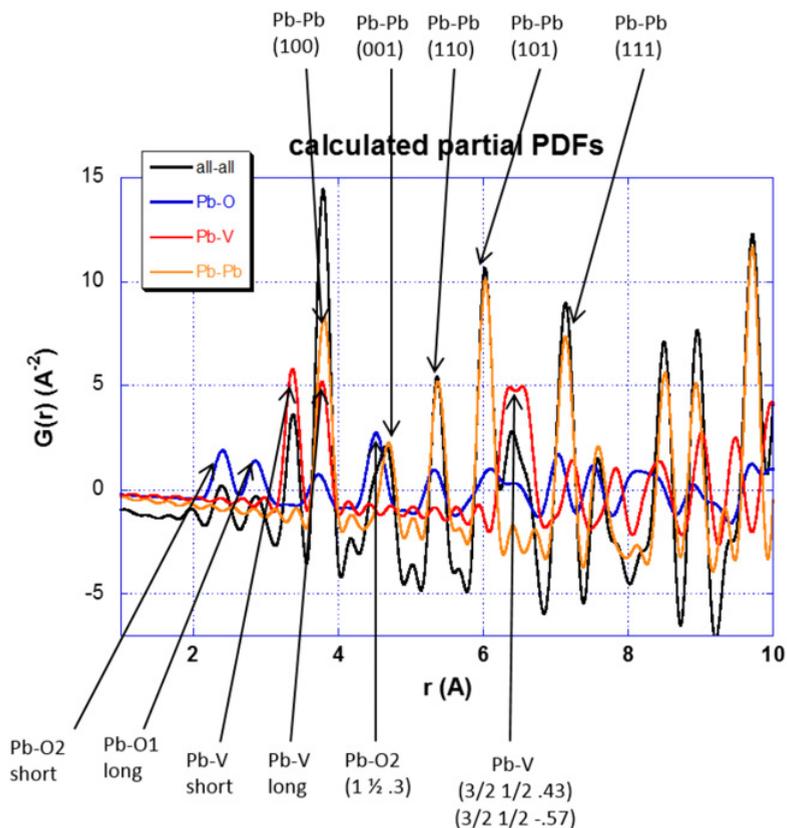


**Figure 7.** Comparison of simulated  $F(Q)$  and  $G(r)$ . In the right side, the peak/background ratio was divided by a factor of 10 with respect to the left part.

The worse the reciprocal space resolution (i.e. the high  $H$  value), the more will this damping occur, which may prevent to follow the PDF  $p$  to large interatomic distances. It is clear that the PDF of Fig. 2 above could only be obtained on a synchrotron beamline optimized for ultimate resolution. Neutron diffractometer cannot compete in this instance and generally show obvious effects of instrument resolution. This can be approximately modelled using a gaussian damping function, but it is only a gross approximation since it is well-known that the instrumental resolution strongly varies with  $Q$ . Therefore it is advisable to choose an instrumental resolution adapted to the type of problem the user wants to deal with. For example, the D4c instrument of the ILL is optimized for liquid and amorphous material and the PDF signal will hardly exceed a few tens of Å. Thus, it can be preferentially used for the studies of local structural effects or small nanoparticles which concentrate on short range distances.

### 4.3 The PDF peak width

Since the density of interatomic distance rapidly increases with  $r$ , it becomes difficult to identify a PDF peak with a single atomic pair past the first few Å even for simple structures. This may be possible for the first few peaks, as can be seen below in Fig. 8, where the total and partial x-ray PDFs are represented. Here, several peaks corresponding to well identified pairs of atoms can be observed, either because they are isolated (the first 3 peaks) or because they strongly dominate the PDF (the Pb-Pb peaks). Although they correspond to interatomic distances in a crystalline phase, these peaks clearly



**Figure 8.** Calculated total and partial X-ray PDFs for the  $\text{PbVO}_3$  tetragonal perovskite structure. The atomic pairs and the direction of the bonds are indicated.

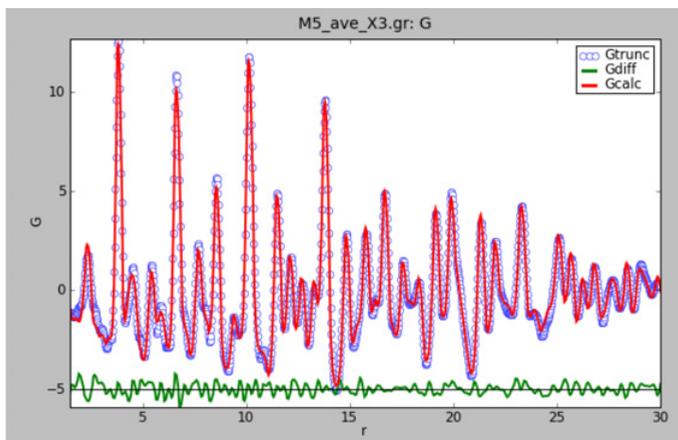
have a finite width. This width is due to the distribution of distances for the given atomic pair, due to thermal displacements and static disorder. It can be modeled with a Gaussian distribution the width of which is defined by an pair displacement parameter, in a similar fashion as the Debye-Waller factor for diffraction. However, here we have to deal with the relative displacements of the atoms forming the pair, in contrast to individual displacement as seen in diffraction. The Debye-Waller factor for PDF can be calculated as:

$$\underbrace{\sigma_{ab}^2}_{PDF} = \underbrace{\sigma_a^2}_{Diffraction} + \sigma_b^2 - 2\sigma_a^2\sigma_b^2\rho$$

Here,  $\rho$  is a factor related to the correlation of displacements between the atoms forming the pair. It can generally be considered as nil, unless for near neighbors where  $\rho > 0$  and correlations become sizeable, leading to a decrease of the peak width. This effect can be used to evidence such correlations, and it must be taken into account to model correctly the PDF at short distances.

#### 4.4 Nanoparticle size effects

In the case of nanoparticles, the PDF peaks vanish for distances longer than the longest interatomic distances in the particle, e.g. for spherical nanoparticles, above the particle diameter. This can be seen on



**Figure 9.** PDFGui fit of the X-ray PDF of ceria nanoparticles, yielding a spherical particle diameter of about 63 Å.

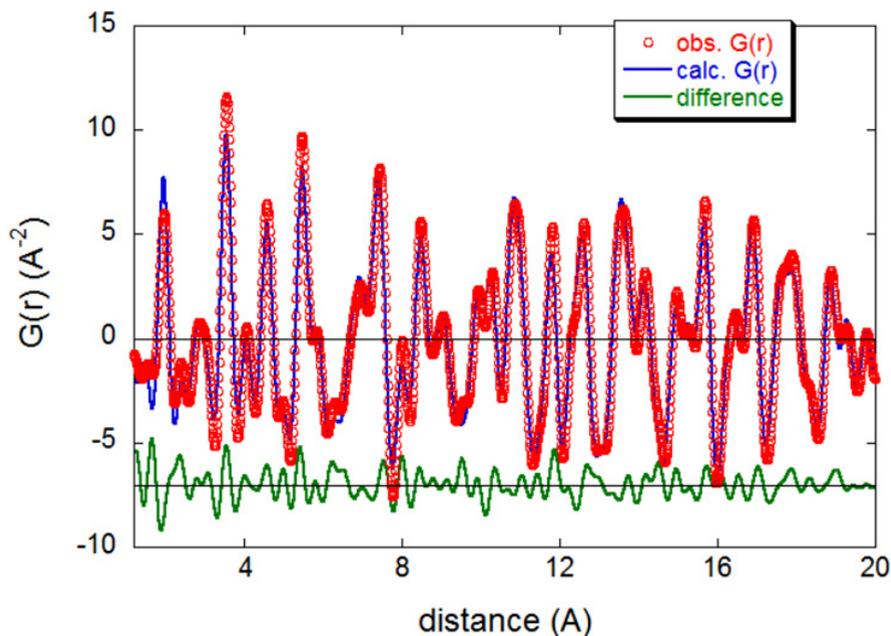
the PDF of Fig. 3, where no PDF peaks can be distinguished above 8–9 Å, yielding directly an estimate of the particle diameter. Above this value, all the interatomic pairs in the sample correspond to atoms of different, randomly oriented nanoparticles, thus yielding a continuous distribution corresponding to the average pair density. Therefore the PDF becomes nil, except from noise. Below this value, the PDF is the sum of intra- and inter-particle contributions.

In the case of nanoparticles, the number density  $\rho_0$  used to calculate the PDF cannot be considered anymore as a constant: it first increases with  $r$  until it reaches a maximum and then decreases to reach zero at the longest interatomic distance. In principle, the shape of  $\rho_0(r)$  will depend on the nanoparticle shape. Several contributions have been reported to take this effect into account, either by multiplying the PDF by a suitable shape function [11], or taking in to account a correct shape for  $\rho_0(r)$  [12]. In the case of spherical nanoparticles, an analytical expression can be derived [13] which allows to introduce the size effect in the refinement process and have an experimental determination of the particle size. An example is shown below for CeO<sub>2</sub> nanoparticles of about 6 nm using AgK $\alpha$  X-ray laboratory data (Fig. 9).

## 5. Modelling of the PDF

The main feature of the PDF is to provide a representation of the structure of a compound through the distribution of interatomic distances and this in a variety of scales ranging from the local order (some Å) to average structure (a few tens of nm). It then constitutes a true multi-scale structure investigation tool and will be fruitfully applicable in cases where these local and average structures differ, either because of the coherent domain/grain size (nano-crystalline and amorphous compounds), or because of localized defects or structural disorder (static and dynamic) which cannot be taken into account efficiently by conventional crystallographic techniques. We have seen that the PDF can be obtained experimentally from total scattering measurements and can also be calculated provided some kind of distribution of atoms is given. As usual, in order to validate and improve the structural model it will be necessary to search for the configuration yielding the best agreement between the observed and calculated PDF.

Two major types of methods are used to model the PDF. The first is based on numerical minimization techniques of the Reverse Monte Carlo type (e.g. program RMCProfile [14]). The structural model is built by distributing more or less randomly atoms in a (large) box, and is optimized by using a Reverse Monte Carlo algorithm minimizing the difference between the observed and calculated PDF's [15]. It therefore requires large computing resources when the problem size increases and its chances of



**Figure 10.** Plot of the X-ray PDF refinement of a mixture of TiO<sub>2</sub> anatase/rutile. Red circles: experimental points; blue line: calculated PDF; green line: difference.

success are heavily dependent on the choice of the physico-chemical constraints applied to the model (distances/angles limits, coordination imposed, bond valence sum...). It requires dedicated tools for statistical analysis of the resulting structure. This method is particularly well suited to amorphous or nearly amorphous compounds and liquids.

The second type of techniques (e.g. program PDFGui [16]) consists in a least squares profile refinement of the experimental PDF with a periodic structural model, sort of Rietveld refinement in direct space (Fig. 10).

This method is faster and better suited to compounds which are close to a crystalline state. Symmetry constraints can be imposed to all or part of the structure and refinements can be applied to different ranges of distances, to compare local and average structure. The effects of various instrumental parameters ( $Q_{max}$ , spatial resolution...) and sample dependent effects (atomic displacement correlations, spherical nanoparticle size...) can be applied. The model can also comprise several different phases with individual scale factors, allowing quantitative analysis to be performed regardless of the crystalline state of the different phases.

This method is faster and better suited to compounds which are close to a crystalline state. Symmetry constraints can be imposed to all or part of the structure and refinements can be applied to different ranges of distances, to compare local and average structure. The effects of various instrumental parameters ( $Q_{max}$ , spatial resolution...) and sample dependent effects (atomic displacement correlations, spherical nanoparticle size...) can be applied. The model can also comprise several different phases with individual scale factors, allowing quantitative analysis to be performed regardless of the crystalline state of the different phases.

## 6. Applications

Here we will not present examples which can easily be found in literature. The interest of using the PDF lies in studies where standard crystallographic techniques cannot be used. This can be first because

the object under study cannot be considered as a crystal due to lack of long range translational order. This is of course the case of glasses and liquids, the traditional domain of application of the PDF. Now, more and more, the PDF analysis method is applied to the study of nanoparticles. Indeed, it is one of the rare methods which can provide at the same time information about the particle dimensions and internal structure. Here it is important to point out that what we call particle dimensions corresponds to the coherently scattering domains for the radiation considered. For instance, if nanoparticles are formed by subdomains randomly oriented inside the particles, the PDF, in a similar way as for diffraction, will inform about the characteristics of these subdomains and not about the particles as a whole. Thus, comparisons with electron microscopy observations must be made with caution. For such problems, the analysis of the PDF still has to make some progress in taking accurately into account the particle shapes. Recent synchrotron beamline developments allowing a simultaneous measurement of PDF and small angle scattering could help solving this issue.

Another type of studies where PDF analysis is important are those related to local structural effects: here the local structure will differ from the average one as seen by diffraction-based crystallographic analysis. The PDF is multi-scale in nature and will permit fitting the structure at different distance ranges, which allows to evidence and characterize local modifications of the average structure. These are of fundamental importance for the correct understanding of the physical properties of materials, particularly electric and magnetic properties. Numerous examples can be found in literature for High-Tc cuprate superconductors, CMR manganites, piezoelectrics, etc.. It may also be possible to quantitatively refine the local structure in the case of compounds presenting micro-twinning or stacking faults, provided the analysis is carried out up to distances smaller than the domain sizes or distance separating faults.

Most recent developments of PDF analysis have been applied to inorganic and often simple compounds. In principle, any of these applications are also valid for small molecule compounds. However, these materials are generally much more complex, with low symmetry structures and are constituted of weakly diffracting atomic species. Also, the lack of implementation of practical rigid-block molecular description in PDF refinement programs has hindered its use in structural analysis of small molecule materials, where the number of structural parameters becomes rapidly too large. Most applications of the PDF in the case of pharmaceuticals have then relied on direct inspection and comparison between PDF patterns or with simulations. More recently, Billinge et al. [17] demonstrated the interest of PDF to investigate nano-crystalline and amorphous molecular compounds. Using carbamazepine and indomethacin as test materials, they showed the possibility to unambiguously identify a melt quenched molecular compound from its PDF signature, or identify polymorphs of the same molecule in an amorphous state. This opens new perspectives for the application of PDF analysis to molecular compounds.

## 7. Conclusion

With the trend towards elaboration of drugs in the form of nano- or amorphous materials, new structure investigation tools become necessary to determine the crystalline arrangement and microstructure of badly diffracting samples. The pair distribution function analysis is especially well suited to be such a tool. Instruments are now widely available at neutron spallation and reactor sources, and synchrotrons around the world to provide high quality PDF data. Moreover, laboratory diffractometers equipped with short wavelength sources can also be used successfully. Efficient and user friendly software is available to analyze the data. The lack of a software allowing refinement of the PDF using more sophisticated constraints such as rigid block molecular models is still a drawback for small molecule compounds, but this problem is expected to be solved very soon. PDF analysis provides a way of refining (and even maybe solving) the structures of materials regardless of their crystallinity, and provides at the same time information on the size as well as the internal structure of nanoparticles. It makes no doubt to use that

the use of PDF analysis is soon going to increase very rapidly in a broad field going from materials science to small molecule and pharmaceuticals.

The author wants to express its gratitude to the following persons: M. Brunelli, I.E. Grey and I. Madsen for fruitful discussions and experimental facilities, and instrument support at beamlines (among others) D2AM-ESRF, ID11-ESRF, CRISTAL-SOLEIL, and D4c-ILL.

## References

- [1] Grey, I. E., Bordet, P., Wilson, N. C., Townend, R., Bastow, T. J., Brunelli, M., *American Mineralogist* **95-1** 161(2010)
- [2] Božin E. S., Schmidt M., DeConinck A. J., Paglia G., Mitchell J. F., Chatterji T., Radaelli P. G., Proffen Th., and Billinge S. J. L., *Phys. Rev. Lett.* **98**, 13720 (2007)
- [3] Guinier A., *X-Ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*, Freeman (San Francisco) (1963)
- [4] Warren B. E., *X-Ray Diffraction*, Addison-Wesley Publishing Co. (1969)
- [5] Keen D. A., *J. Appl. Cryst.* **34**, 172-177 (2001)
- [6] Egami T. and Billinge S. J. L., *Underneath the Bragg Peaks: Structural Analysis of Complex Materials*, 2nd edn, Pergamon, Oxford (2012)
- [7] Masson O. and Thomas Ph., *J. Appl. Cryst.* **46**, 461 (2012)
- [8] Bates S., Zografi G., Engers D., Morris K., Crowley K., Newman A., *Pharm. Res.* **23**(10), 2333 (2006)
- [9] Dykhne T., Taylor R., Florence A. and Billinge S. J. L., *Pharm. Res.* **28**, 1041 (2011)
- [10] Chupas P. J., Qiu X., Hanson J. C., Lee P. L., Grey C. P., and Billinge S. J. L., *J. Appl. Cryst.* **36**, 1342 (2003)
- [11] Kodama K., Iikubo S., Taguchi T. and Shamoto S, *Acta Cryst.* **A62**, 444 (2006)
- [12] Neder R. B. and Korsunskiy V. I., *J. Phys.: Condens. Matter* **17**, S125–S134 (2005)
- [13] Howell R. C., Proffen T. and Conradson S. D., *Phys. Rev.* **B 73**, 094107 (2006)
- [14] Keen D. A., Tucker M. G. and Dove M. T., *J. Phys.: Condens. Matter* **17**, S15 (2005)
- [15] McGreevy R. L., *J. Phys.: Condens. Matter* **13**, R877 (2001)
- [16] Farrow C. L., Juhas P., Liu J. W., Bryndin D., Božin E. S., Bloch J., Proffen Th. Billinge and S. J. L., *J. Phys.: Condens. Matter* **19**, 335219 (2007)
- [17] Billinge S. J. L., Dykhne T., Juhas P., Bozin E., Taylor R., Florence A. J. and Shankland K., *Cryst. Eng. Comm.*, **12**, 1366 (2010)