



# *SUePDF*: a program to obtain quantitative pair distribution functions from electron diffraction data

Dung Trung Tran,\*‡ Gunnar Svensson and Cheuk-Wai Tai\*

Department of Materials and Environmental Chemistry, Arrhenius Laboratory, Stockholm University, Svante Arrhenius väg 16C, Stockholm, S-10691, Sweden. \*Correspondence e-mail: [trungdungtran@gmail.com](mailto:trungdungtran@gmail.com), [cheuk-wai.tai@mmk.su.se](mailto:cheuk-wai.tai@mmk.su.se)

Received 26 May 2016

Accepted 21 November 2016

Edited by D. Pandey, Indian Institute of Technology (Banaras Hindu University), Varanasi, India

‡ Present address: Johnson Matthey Technology Centre, Blounts Court Road, Sonning Common, Reading RG4 9NH, UK.

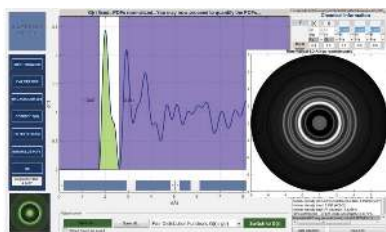
**Keywords:** pair distribution functions; electron diffraction; transmission electron microscopy; disorder.

*SUePDF* is a graphical user interface program written in MATLAB to achieve quantitative pair distribution functions (PDFs) from electron diffraction data. The program facilitates structural studies of amorphous materials and small nanoparticles using electron diffraction data from transmission electron microscopes. It is based on the physics of electron scattering as well as the total scattering methodology. A method of background modeling is introduced to treat the intensity tail of the direct beam, inelastic scattering and incoherent multiple scattering. Kinematical electron scattering intensity is scaled using the electron scattering factors. The PDFs obtained after Fourier transforms are normalized with respect to number density, nanoparticle form factor and the non-negativity of probability density. *SUePDF* is distributed as free software for academic users.

## 1. Introduction

The pair distribution function (PDF) method is widely employed for studying structurally disordered materials (Warren, 1990; Egami & Billinge, 2002; Proffen *et al.*, 2003). This is because it gives more structural information, beyond the information given by the traditional Bragg-peak-based analysis, from scattering data. Standard computer programs for PDFs for neutron and X-ray powder diffraction data are well established (Peterson *et al.*, 2000; Juhás *et al.*, 2013). A number of efforts have been made to obtain PDFs from electron diffraction (ED) data (Cockayne & McKenzie, 1988; Tewes *et al.*, 1994; Takagi *et al.*, 2001; McBride, 2003; Ankele *et al.*, 2005; Ishimaru *et al.*, 2008; Abeykoon *et al.*, 2012; Mu *et al.*, 2013). However, these have not been sufficient to establish electron powder diffraction as one of the major data sources for PDF analysis. The main obstacle has been the multiple scattering of electrons, which may alter the scattering intensities in the ED patterns, making it difficult to extract an undistorted structure function (Uyeda, 1968; Cowley, 1969; Anstis *et al.*, 1988; Cockayne & McKenzie, 1988). Beside this, the electron–matter interaction differs from the cases of X-rays and neutrons, implying that a dedicated procedure of data treatment is needed for ED data.

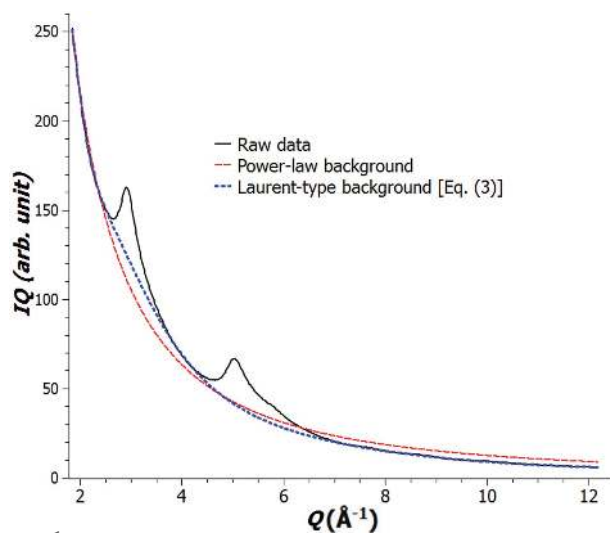
In order to obtain reasonable PDFs, ED data  $I(Q)$  have to be scaled using appropriate electron scattering factors  $f_e(Q)$  to yield a proper reduced structure function  $F(Q)$ , where  $Q$  is the magnitude of the scattering vector. This is not an easy task because the ED data are generally contaminated by inelastic scattering and distorted by multiple scattering, instrument errors and noise. The widely reported solution for inelastic scattering is using an energy filter, but this may also cause additional distortion in the data when only electrons of certain energy-loss range can be blocked. Data scaling was previously



OPEN ACCESS

carried out by multiplying the compositional averages  $\langle f_e^2(Q) \rangle$  and  $\langle f_e(Q) \rangle^2$  with a fitting parameter  $\eta$ , which may be fixed by matching  $\eta \langle f_e^2(Q) \rangle$  with  $I(Q)$  at large  $Q$  (Cockayne & McKenzie, 1988; Cockayne, 2007; Cockayne *et al.*, 2010). An early software attempt to fit  $I(Q)$  with  $\langle f_e^2(Q) \rangle$  for PDF extraction was made by drawing and then subtracting a background of  $I(Q)$  (Hauschild, 2009). However, because the mathematical curve presented for this background is generally not suitable, the background is often constructed manually. The next software attempt was made by Mitchell & Petersen (2012), in which an additional fitting parameter  $\alpha$  was introduced to counter the discrepancy between  $I(Q)$  and  $\eta \langle f_e^2(Q) \rangle$ . This meant that  $F(Q)$  was obtained by matching  $\eta \langle f_e^2(Q) \rangle - \alpha$  with  $I(Q)$  at large  $Q$ . In other work, instead of introducing an additional parameter,  $F(Q)$  was directly adjusted by fitting the distortion feature to a fourth-order polynomial (Mu *et al.*, 2013).

The program *SUePDF* aims to effectively correct ED data by taking into account the multiple scattering features and uniqueness of electron–matter interactions. For amorphous materials, the coherent multiple scattering can be considered as insignificant and incoherent multiple scattering should merely contribute to the background (Cowley, 1992), which may be modeled and subtracted. For crystalline-like nanoparticles, coherent multiple scattering should not affect the peak positions of the resulting PDFs but only the peak intensities. This has been demonstrated to be the case for crystals having a thickness of less than five times the electron mean free path (Anstis *et al.*, 1988). The problem with modulated peak intensities caused by coherent multiple scattering is tackled in *SUePDF* by a renormalization procedure based on number densities, and the probability of having an atom at a certain distance is non-negative. In general, the major distortion of  $I(Q)$  can be considered to form a smooth background, which is built up from the direct-beam tail,



**Figure 1**  
Background modeling for electron powder diffraction data (solid black line) of nanoporous carbon: power-law model (dashed red line) compared with Laurent-type model (dotted blue line) [equation (3)] with  $N = 7$ .

inelastic continuum and incoherent multiple scattering. *SUePDF* employs the combination of an optimizing parameterization and a reasonable mathematical model for the background. Inside *SUePDF*, background subtraction is coupled with data scaling in a loop-based routine to optimize the data treatment. After background subtraction and data scaling, the PDF can be obtained by a Fourier transform of the normalized data. On the basis of the physical meaning of PDFs, noise filtering and normalization procedures have to be routinely carried out. This improves the physical reliability of the outcome PDFs and allows the uncertainties to be evaluated. *SUePDF* also offers the possibility to correct finite size effects present for nanoparticles, by using a nanoparticle form factor computed for a given size and shape.

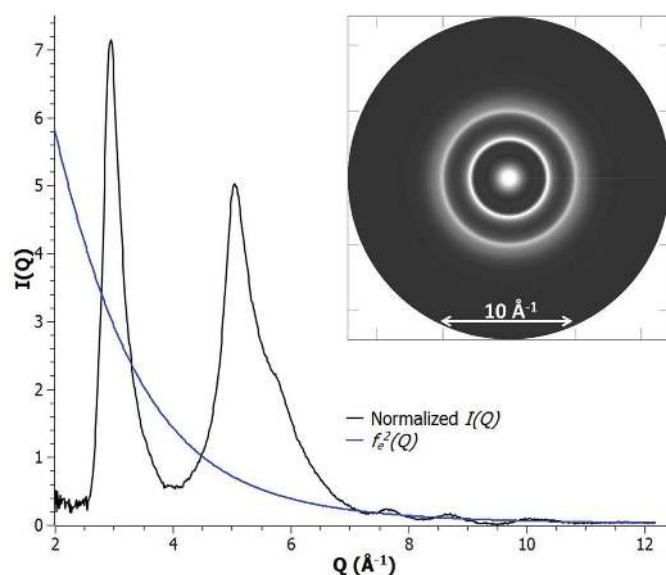
## 2. Methods

### 2.1. Background modeling

The smooth background observed for ED data is considered to be the contributions from the direct-beam tail, inelastic scattering continuum and incoherent multiple scattering. For an elastic scattering of a  $\lambda$ -wavelength electron with a semi-angle  $\theta$ ,  $Q = 4\pi \sin \theta / \lambda$ , but in the case of small-angle inelastic scattering (Egerton, 2011)

$$Q = \left[ \left( \frac{4\pi \sin \theta}{\lambda} \right)^2 + \left( \frac{2\pi \bar{\theta}}{\lambda} \right)^2 \right]^{1/2}, \quad (1)$$

where  $\bar{\theta} = \gamma_E \Delta E / [K_E(1 + \gamma_E)]$  is the characteristic angle corresponding to an energy loss  $\Delta E$ ;  $\gamma_E = (1 + K_E/E_0)$  is the relativistic factor with  $K_E$  the kinetic energy and  $E_0 \simeq 511$  keV the stationary mass-converted energy of the electron. Equation (1) can be rewritten for the inelastic component when  $\theta = 0$  as



**Figure 2**  
Normalized  $I(Q)$  of a nanoporous carbon sample (back solid) and the corresponding  $\langle f_e^2(Q) \rangle = f_e^2(Q)$ ; the inset shows a ring pattern reconstructed from this normalized  $I(Q)$ .

$$Q_{\text{inel}}(\Delta E) = \frac{2\pi\gamma_E \Delta E}{\lambda K_E(1 + \gamma_E)}. \quad (2)$$

For example, if the electron energy loss is counted up to 2000 eV for an incident beam of 200 keV, then the inelastic beam tail is limited to below  $Q_{\text{inel}}(2000) \simeq 1.46 \text{ \AA}^{-1}$ . For low-loss electrons ( $\Delta E < 50 \text{ eV}$ ), the inelastic component  $Q_{\text{inel}}(<50)$  is less than  $\sim 0.036 \text{ \AA}^{-1}$ . Equation (2) means that the inelastic error is more significant for electrons of high energy loss. It is well known that high-loss electrons can be excluded using an energy filter. However, according to our experience, energy filtering also modifies the background, making an accurate quantitative background modeling difficult. The following mathematical model is introduced to fit the background  $B(Q)$  of the electron powder diffraction pattern:

$$B_N(Q) = \sum_{k=1}^N c_k/Q^k. \quad (3)$$

Equation (3) is actually a positive-degree portion of a Laurent-type series where  $c_k$  are the fitting parameters and  $N$  is the fitting order, which may vary for different sample compositions. A comparison between the power-law model and the Laurent-type model [described in equation (3), with  $N = 7$ ] for background fitting of electron powder diffraction data acquired from a nanoporous carbon sample is shown in Fig. 1. In addition, this background model works in both thin and thick samples. This is demonstrated in a study of amorphous silica discussed in §4.2. The measured bond lengths and the values reported in the literature are listed there and are in good agreement.

### 2.2. Optimization procedure for data scaling and background modeling

Electron scattering factors, denoted as  $f_c(Q)$ , are used for data scaling. These factors may be obtained using the following Mott–Bethe formula (Mott & Massey, 1965):

$$f_c(Q) = \frac{2\gamma}{a_0} \left[ \frac{Z - V - f_x(Q)}{Q^2} \right], \quad (4)$$

where  $Z$  is the atomic number,  $V$  is the valence number,  $a_0 \simeq 0.53 \text{ \AA}$  is the Bohr radius and  $f_x(Q)$  is the X-ray scattering factor (Brown *et al.*, 2006). For neutral atoms, it is recommended to use the available DFT-computation-based parameterizations (Kirkland, 2010), which are considered to be more accurate than those from the Mott–Bethe formula, particularly at low  $Q$  values. For samples composed of more than one element, the chemical compositions are required as

input, together with the corresponding valence for each element.

The data scaling is constrained by a mathematical feature of the structure function  $S(Q)$ . That is,  $S(Q) \rightarrow 1$  when  $Q \rightarrow \infty$ . The background-subtracted and normalized scattering intensity  $I(Q)$  is converted into  $S(Q)$  by using the composition-averaged  $\langle f_c^2(Q) \rangle$  and  $\langle f_c(Q) \rangle^2$  (Warren, 1990):

$$S(Q) = 1 + \frac{I(Q) - \langle f_c^2(Q) \rangle}{\langle f_c(Q) \rangle^2}. \quad (5)$$

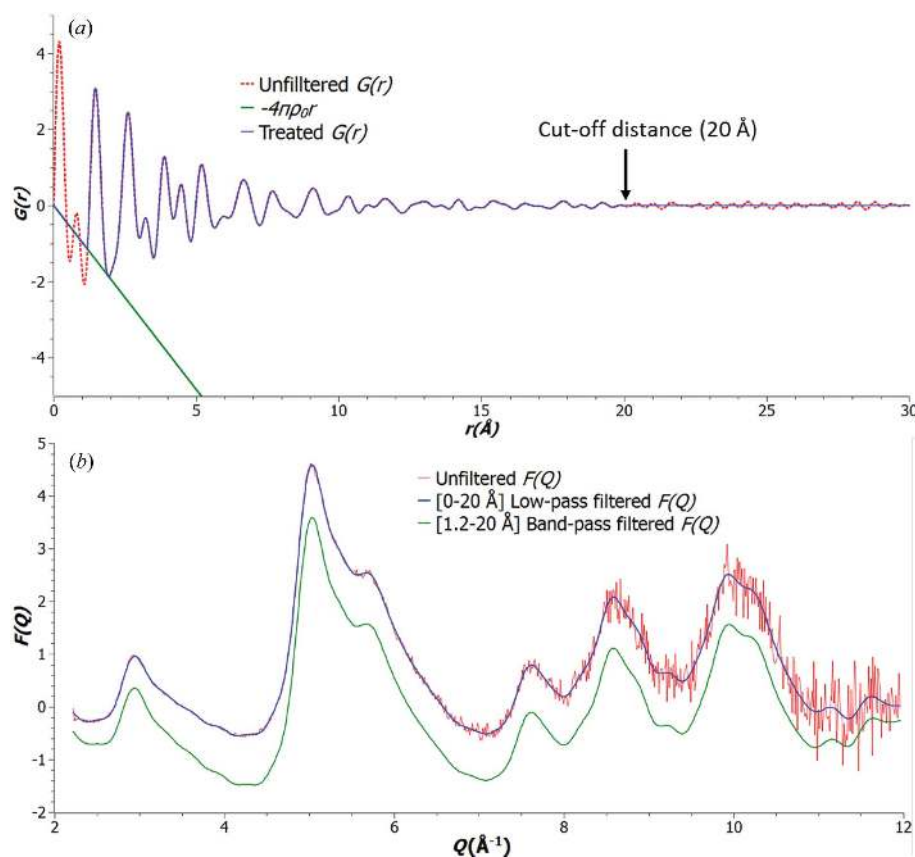
The data scaling is done according to

$$I(Q) = \frac{\int_{Q_{\text{min}}}^{Q_{\text{max}}} \langle f_c^2(Q') \rangle dQ'}{\int_{Q_{\text{min}}}^{Q_{\text{max}}} [I_{\text{raw}}(Q') - B_N(Q')] dQ'} [I_{\text{raw}}(Q) - B_N(Q)], \quad (6)$$

where  $I_{\text{raw}}(Q)$  is the raw data of scattering intensity. The feature  $S(Q) \rightarrow 1$  when  $Q \rightarrow \infty$  restricts  $I(Q)$  to attenuate around  $\langle f_c^2(Q) \rangle$  at high  $Q$  values. This suggests a necessary minimization of the following quantity, defined at the tail of  $I(Q)$ :

$$\chi_{\text{tail}}^2 = \int_{\text{tail}} [I_{\text{tail}}(Q) - \langle f_c^2(Q) \rangle_{\text{tail}}]^2 dQ \rightarrow \min. \quad (7)$$

In *SUEPDF*, the minimization of equation (7) is carried out in a loop procedure by varying the background reference and the



**Figure 3** (a) Normalization and noise treatment for  $G(r)$  of nanoporous carbon, with  $r_{\text{min}} \simeq 1.2 \text{ \AA}$  and  $r_{\text{max}} \simeq 20 \text{ \AA}$ ; (b) [0–20 Å] low-pass-filtered  $F(Q)$  (blue solid line) and [1.2–20 Å] band-pass-filtered  $F(Q)$  (green solid line) compared with the unfiltered  $F(Q)$  (red).

fitting order  $N$ , which are initially input by users, to optimize the normalized  $I(Q)$ . A normalized  $I(Q)$  of a nanoporous carbon sample, scaled by the corresponding  $\langle f_c^2(Q) \rangle = f_c^2(Q)$ , is shown in Fig. 2.

### 2.3. Fourier transform of $F(Q)$ to yield reduced PDF

The reduced PDF  $G(r)$  is obtained by a Fourier transform of the reduced structure function  $F(Q) = Q[S(Q) - 1]$ :

$$G(r) = \frac{2}{\pi} \int_{Q_{\min}}^{Q_{\max}} F(Q) \sin(Qr) dQ. \quad (8)$$

The red dotted curve shown in Fig. 3(a) is the unfiltered  $G(r)$  for nanoporous carbon after the Fourier transform [equation (8)]. Besides being convolved with the termination function  $[\sin(Q_{\max}r) - \sin(Q_{\min}r)]/(\pi r)$  (Peterson *et al.*, 2003), this unfiltered  $G(r)$  exhibits some artifact peaks at low ( $<1 \text{ \AA}$ ) and high ( $>20 \text{ \AA}$ ) values of  $r$ , which are equivalent to the low- and high-frequency noise, respectively, of the experimental data.

### 2.4. PDF normalization and noise filtering

The unfiltered  $G(r)$  obtained from equation (8) has to be adjusted by physical and mathematical constraints. These constraints refer back to the definition of PDF  $g(r)$ , which represents the probability density of finding a pair of two atoms separated by distance  $r$  (Egami & Billinge, 2002). Therefore  $g(r)$ , the probability density, has to be non-negative:

$$g(r) = \gamma(r) + \frac{G(r)}{4\pi\rho_0r} \geq 0, \quad (9)$$

where  $\gamma(r)$  is the nanoparticle form factor (Kodama *et al.*, 2006; Gilbert, 2008; Tran *et al.*, 2016) [for bulk,  $\gamma(r) \equiv 1$ ] and  $\rho_0$  is the average number density of the sample. The normalization of the PDFs is based on equation (9). Prior knowledge about the shortest interatomic distance existing in the sample is used as a physical constraint, such that the PDF  $g(r)$  at distances smaller than a value  $r_{\min}$  is set to zero:

$$\begin{aligned} g(r) &= 0 \\ G(r) &= -4\pi\rho_0r\gamma(r) \end{aligned} \quad \text{when } r < r_{\min}. \quad (10)$$

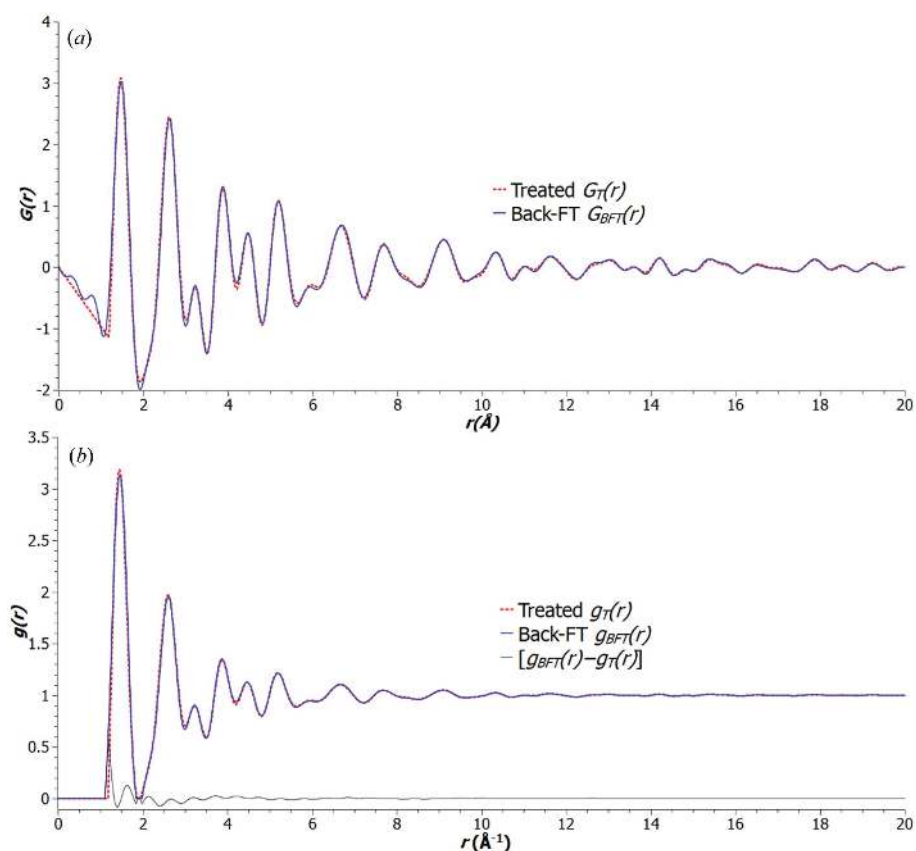
Equation (10) is used to filter off low-frequency noise. The high-frequency noise can be filtered off by setting an upper cut-off distance  $r_{\max}$ , where  $G(r) = 0$  at  $r > r_{\max}$ . This noise treatment of  $G(r)$  is demonstrated in Fig. 3(a). A Fourier back-transform of the treated  $G(r)$  then yields the noise-

filtered  $F(Q)$  data, shown in Fig. 3(b), for the here given example from a nanoporous carbon sample.

### 2.5. Evaluation of uncertainty

The treated reduced PDF  $G_T(r)$  is compared with  $G_{\text{BFT}}(r)$ , which is the Fourier transform of the  $[1.2\text{--}20 \text{ \AA}]$  band-pass-filtered  $F(Q)$ , in Fig. 4(a). The observed difference between  $G_T(r)$  and  $G_{\text{BFT}}(r)$  is the consequence of the cutting of low and high frequencies which is propagated through the Fourier transform. Therefore, the uncertainty considered here is mainly associated with the low- and high-frequency noise, which correspond to the artifact short- and long-distance peaks found in the unfiltered  $G(r)$ , respectively.

From Fig. 3(b), it is clear that the low-frequency noise can be attributed to some low-frequency distortions of the  $F(Q)$  data. These distortions may be caused by the imperfection of the instrumental setup, the errors encountered during the background subtraction and data scaling using the approximated electron scattering factors. One of the typical errors is the beam convergence, which generally does not affect the positions of PDF peaks but does affect their intensities (McBride, 2003). Beside these errors, in the case of coherent structures, the low-frequency distortions may be related to the coherent multiple scattering which might not be treated



**Figure 4** Uncertainty in the noise filtering of PDFs: (a) treated  $G_T(r)$  (red dotted line) compared with  $G_{\text{BFT}}(r)$  (blue solid line) which has been transformed from  $[1.2\text{--}20 \text{ \AA}]$  band-pass-filtered  $F(Q)$ ; (b) the corresponding normalized PDFs  $g_T(r)$  (red dotted line) and  $g_{\text{BFT}}(r)$  (blue solid line) and the difference between these (black line). From these, the evaluated uncertainty is  $\sim 3.7\%$ .

properly by background subtraction. The high-frequency noise, which is believed to be more random, may come from various sources. Some probable sources are the electromagnetic environment, mechanical instability of the instrument *etc.* Note that the properties of the recording media are of importance and can influence the results in both images and diffraction significantly, in particular resolution and sensitivity (Ruskin *et al.*, 2013). In general, the higher the number of pixels in the detector, the better the resolution that can be achieved. The current generation of detectors, which are of CCD and CMOS type and have either  $1000 \times 1000$  or  $2000 \times 2000$  pixels installed in a typical electron microscope configuration, can provide significantly high  $Q_{\max}$  for most applications. A detector with better dynamical range and sensitivity can have the advantage of handling strong diffraction in the low- $Q$  range and acquiring a weaker signal in the higher- $Q$  range. As a result, the extended  $Q$  range can improve the overall quality of the ED-based PDF. The optimization of the acquisition condition for a particular detector can help to obtain quality electron diffraction patterns and therefore ED-based PDFs, especially to minimize artifacts given by the detector, such as streaking, the trace of the beam path given by the shutter, blooming *etc.*

The uncertainty of the normalized PDFs can be evaluated from the relative r.m.s. difference between the corresponding  $g_T(r)$  and  $g_{BFT}(r)$ :

$$U_{g(r_{\min}, r_{\max})} = \frac{\text{r.m.s.}[g_T(r) - g_{BFT}(r)]}{\text{r.m.s.}[g_T(r)]} \quad (11)$$

$U_{g(1.2 \text{ \AA}, 20 \text{ \AA})} \simeq 3.7\%$  for the case of the nanoporous carbon sample shown in Fig. 4(b).

## 3. The graphical user interface of SUEPDF

### 3.1. Electron total scattering profile input

The input data for SUEPDF v1.0 is a one-dimensional electron total scattering intensity profile. The input file must have a simple  $x$ - $y$  two-column format. The  $x$  column is the  $s$  values (in  $\text{\AA}^{-1}$  when calibrated; note that  $Q = 2\pi s$ ) and the  $y$  column is the corresponding intensity  $I(Q)$ . Multiple input files may be selected for integration of different diffraction data of the same sample and under the same experimental conditions. Multiple input files must be synchronized with the same format and the same data size. Fig. 5 shows an

overview of the graphical user interface (GUI) after inputting a data file using 'INPUT BROWSER'.

### 3.2. Loading electron scattering factors

Databases of parameterization for both electron (Kirkland, 2010) and X-ray (Brown *et al.*, 2006) scattering factors (or atomic form factors) are implemented in SUEPDF. In cases of neutral atoms the electron database will be used; otherwise the X-ray database will be loaded in order to calculate the electron scattering factors for ions *via* the Mott-Bethe formula [equation (4)]. Besides the electron energy (in keV), chemical information, including the elemental composition, molar ratios and valences, is required as input to load appropriate electron scattering factors for data scaling. Fig. 6 shows the GUI panel for inputting chemical information and electron kinetic energy and calculation of electron scattering factors.

### 3.3. Background optimization

The background optimization is based on user input of the following:

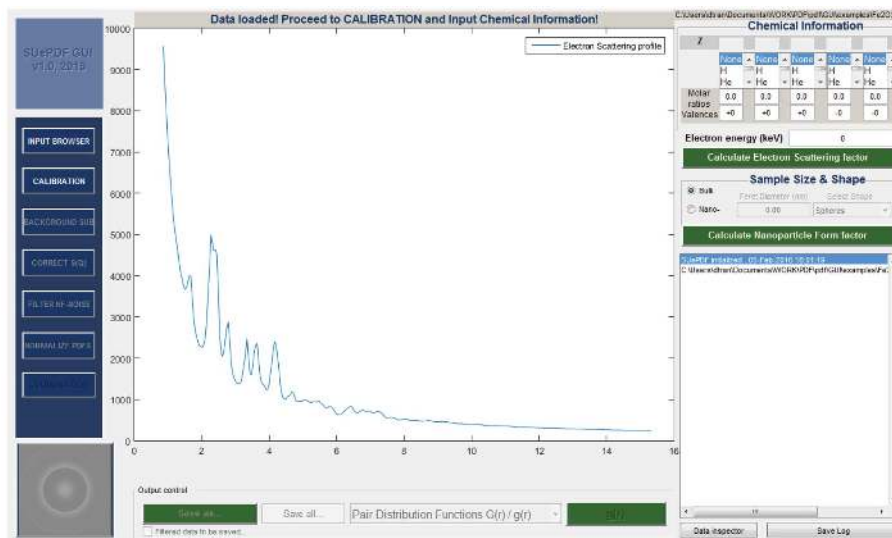


Figure 5  
An overview of the SUEPDF GUI after inputting the data file of an electron scattering profile.

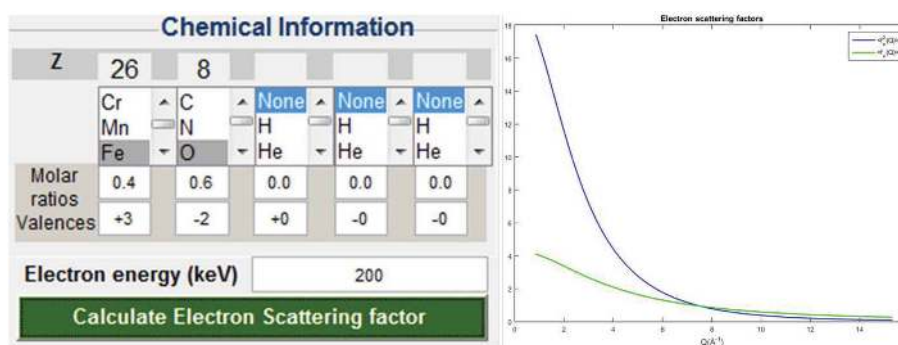


Figure 6  
The GUI for calculating the electron scattering factor by inputting chemical information and electron kinetic energy.

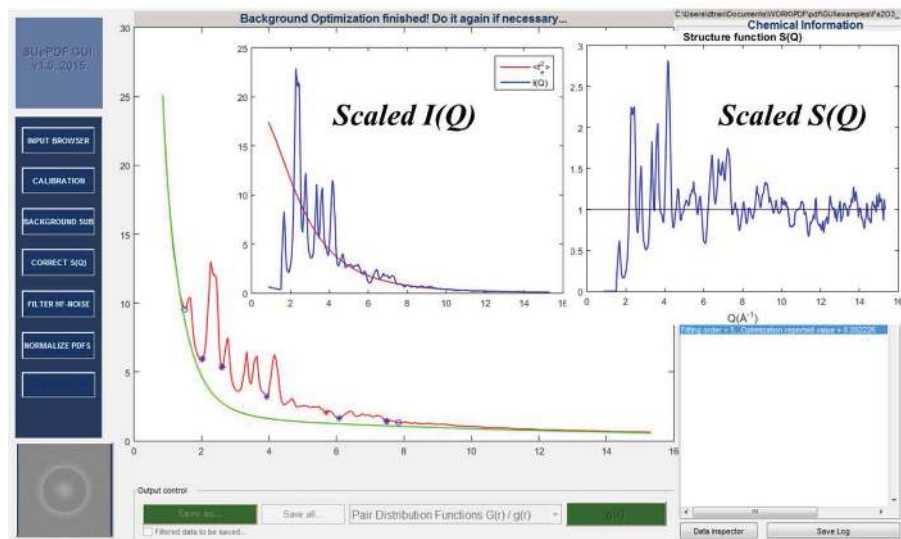


Figure 7  
An optimized background and the corresponding scaled  $I(Q)$  and  $S(Q)$ .

(i) Two points specifying the pre-peak background (corresponding to the lowest momentum transfer) and the tail.

(ii) The number of middle background reference points: these points will be automatically positioned as initial conditions between the previous two points of the pre-peak and the tail. Their positions will vary along the curve of the raw scattering profile to optimize the background. It is noted that these reference points generally do not lie on the background; their distances to the background are refined while their positions vary. The typically recommended number of these reference points is 3–8. Use of larger numbers of reference points consumes more computation time.

(iii) The maximum fitting order: typically recommended orders are 5–15. Larger fitting orders consume more computation time.

Fig. 7 shows a background optimized in the *SUEPDF* GUI, together with the scaled  $I(Q)$  and  $S(Q)$ .

### 3.4. $S(Q)$ correction and high-frequency noise filtering

*SUEPDF* offers a routine for correction of the  $S(Q)$  tail. This routine is optional and only recommended when a good enough solution for the  $S(Q)$  scaling problem cannot be found by background optimization. The correction procedure is based on calculation of the median curve of the  $S(Q)$  tail. There are two steps:

- (1) Specifying the to-be-corrected tail of  $S(Q)$ .
- (2) Tuning the order of the median fitting to achieve a corrected tail of  $S(Q)$ .

The tuning may also be done automatically by pressing the ‘Auto optimization’ button.

High-frequency noise filtering is recommended because ED data usually contain high-frequency noise, which is more visible at high  $Q$  values as spiky oscillations. The filter is based on forward-and-back Fourier transforms. A cut-off distance is required as a user input. The cut-off distance serves as the ‘highest frequency’ allowed in the ED data and is supposed to relate to the atomic structure.

### 3.5. Nanoparticle form factor

The nanoparticle form factor takes into account the size and shape of the sample and quantifies how much they affect the normalized PDF. For a bulk sample this factor is unity. *SUEPDF* v.1.0 offers the calculations for four basic shapes: sphere, cuboctahedron, cube and truncated cube (Tran *et al.*, 2016). When working with nanoparticle samples, the user must input their size and shape to load the appropriate form factor. If this information is not loaded, the default form factor will be unity (for bulk). Fig. 8 shows the GUI panel for calculation of the nanoparticle form factor.

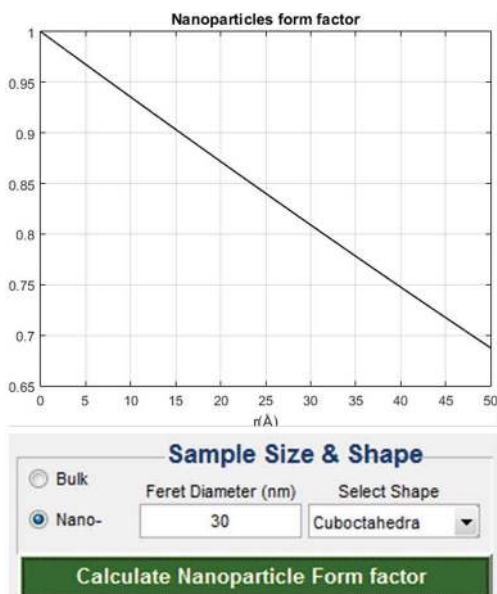


Figure 8  
GUI panel for calculation of nanoparticle form factors.

### 3.6. PDF renormalization

This renormalization is to achieve the quantitative PDF  $g(r)$ . The renormalization procedure is based on the following:

- (a) The non-negativity of  $g(r)$  as probability density.
- (b) The cutting off of short unphysical distances producing low-frequency distortion in the data (low-frequency filtering). This requires a user input of a lower cut-off distance (based on general prior knowledge of the shortest interatomic distance existing in the sample).
- (c) The revision of number density. *SUEPDF* is able to deduce a value of number density from the normalized ED

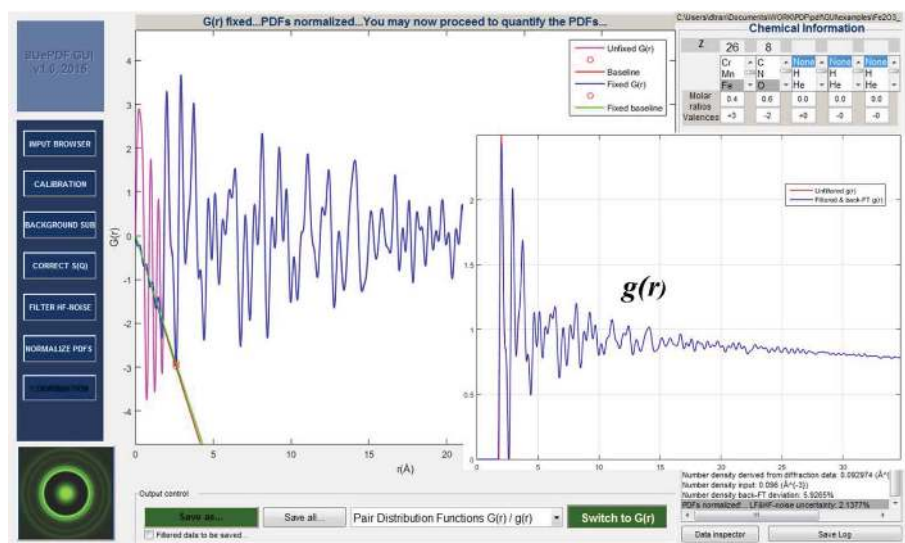


Figure 9 PDF renormalization: the cutting off of low-frequency distortion and number density revision;  $g(r)$  is shown in the inset.

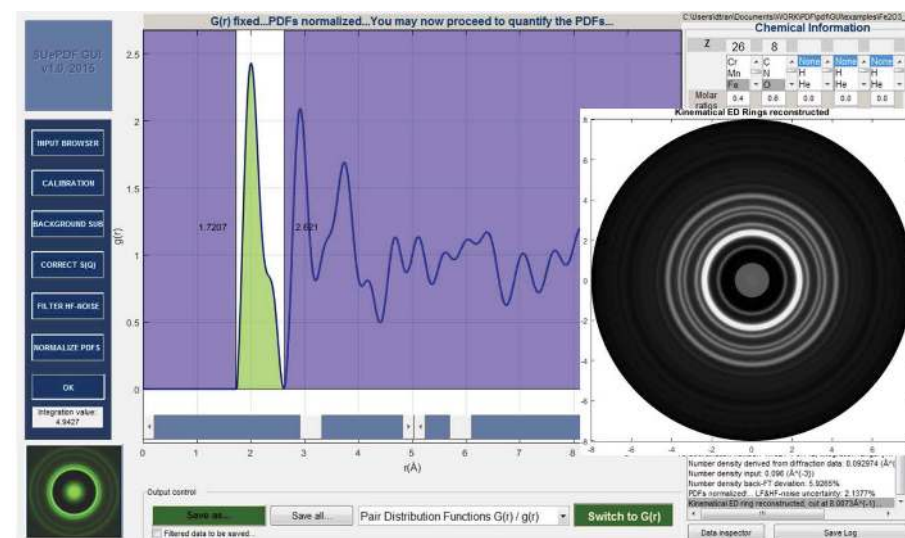


Figure 10 Measuring coordination number and reconstruction of the kinematical ED pattern.

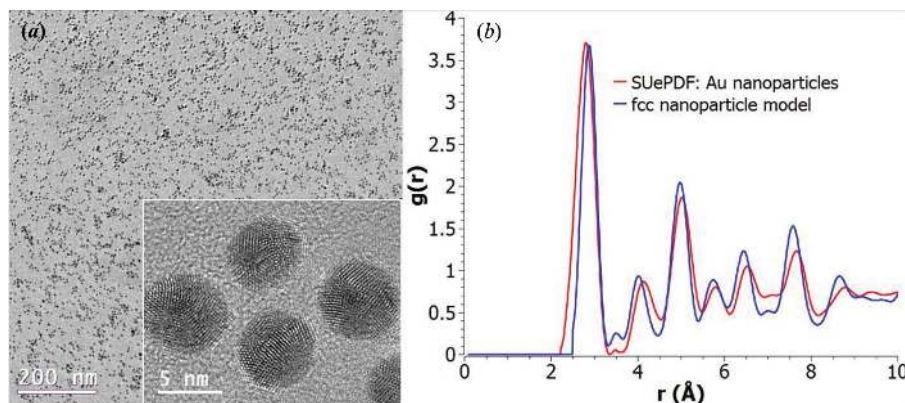


Figure 11 (a) TEM images of  $\sim 5$  nm sized Au nanoparticles including a high-resolution image (inset); (b) ED-based PDF obtained using *SUEPDF* (red) compared with the theoretical PDF of a 5 nm spherical Au nanoparticle model having perfect fcc structure (blue).

data. If a better value of number density is known already, it should be used as the correction for the deduced value. In a good case of data processing of nanoscale samples, the deduced value can be very close to the generally accepted one.

Fig. 9 shows the PDF renormalization by cutting off low-frequency distortions and number density revision.

### 3.7. PDF quantification

As shown in Fig. 10, the coordination number can be measured by specifying an integration window for the relevant  $g(r)$  peak. The background-subtracted and noise-filtered electron scattering profile (which should be the extracted kinematical scattering data) can be reconstructed back into a ring pattern, shown as the inset of Fig. 10.

## 4. Examples

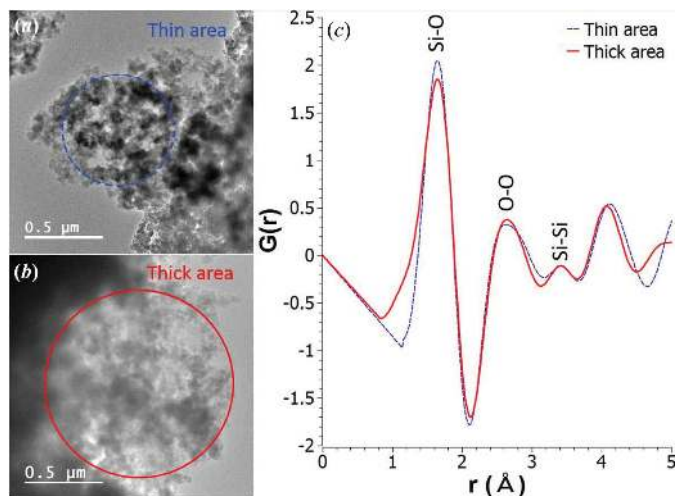
### 4.1. Gold nanoparticles

An example of using *SUEPDF* to study  $\sim 5$  nm sized Au nanoparticles supported on an amorphous carbon film is shown in Fig. 11. The particles are considered to be close to spherical in shape (thus, the nanoparticle form factor of a 5 nm sphere was chosen), although some facets resembling cuboctahedral morphology can be seen in the high-resolution transmission electron microscopy (TEM) image (inset of Fig. 11a). To obtain reliable ED data for the Au nanoparticles, an ED data set of an equivalent blank carbon film was collected as the substrate reference. The PDF based on ED data with  $Q_{\min} = 1.5 \text{ \AA}^{-1}$  and  $Q_{\max} = 12.5 \text{ \AA}^{-1}$  is shown in Fig. 11(b) (red) in a comparison with the theoretical PDF of a 5 nm spherical model of a perfect face-centered cubic (fcc) Au nanoparticle (blue).

### 4.2. Amorphous silica

Fig. 12 shows examples of amorphous silica at thin and thick areas of a sample, in order to demonstrate the validity of the background modeling for samples with different thicknesses. These areas are shown in the TEM images (Figs. 12a and 12b) with the selected-area

apertures (marked with dashed blue and solid red circles for the thin and thick areas, respectively) defining the regions for ED acquisition. The corresponding (reduced) PDFs obtained using *SUePDF* are shown in Fig. 12(c). The PDF of the thin area (dashed blue line) was obtained from ED data with  $Q_{\min} \approx 0.4 \text{ \AA}^{-1}$  and  $Q_{\max} \approx 12 \text{ \AA}^{-1}$ . The PDF of the thick area (solid red line) was obtained with the same  $Q_{\max}$  but a slightly higher  $Q_{\min}$  ( $0.55 \text{ \AA}^{-1}$ ) in order to cut off the possible increase of inelastic scattering. Note that, because of the limited  $Q_{\max}$ , the ED-based PDFs are generally broader than X-ray/neutron-based PDFs and the termination ripples may interfere significantly with some low and broad peaks of amorphous materials (*e.g.* Si–Si peaks). The quantitative measurements of bond lengths and coordination numbers for both the thin and the thick areas are listed in Table 1, together the reference data from neutron scattering (Keen & Dove, 1999) and molecular dynamics simulation of bulk amorphous silica (Hoang, 2007). The bond lengths found by *SUePDF* do not change significantly from the thin to the thick areas. On the other hand, the average coordination numbers (measured with a number density of  $0.065 \text{ \AA}^{-3}$ ) do vary from the thin to the thick region. Apart from the possible sources of error specified in §2.5, this reasonable variance could be attributed to the finite size of the studied regions and local effects when the ED data are obtained from small amounts of sample in a transmission electron microscope, which are not large enough to be fully considered as bulk samples.



**Figure 12** Comparison between thin and thick areas of amorphous silica for obtaining ED-based PDFs using *SUePDF*. (a), (b) TEM images of thin and thick areas, respectively, for ED acquisition; (c) the corresponding ED-based PDFs for the thin (dashed blue line) and thick (solid red line) areas.

**Table 1**

Bond lengths and coordination numbers of amorphous silica found by using *SUePDF* on the ED data obtained from thin and thick areas.

These values are compared with the results from neutron (N) scattering (Keen & Dove, 1999) and molecular dynamics (MD) simulations (Hoang, 2007). Note that the first peak of the total PDF measures the average of the (Si–O) and (O–Si) coordination numbers, which is calculated as  $[N_{(\text{Si-O})}^{\text{Coord}} + 2N_{(\text{O-Si})}^{\text{Coord}}]/3$ .

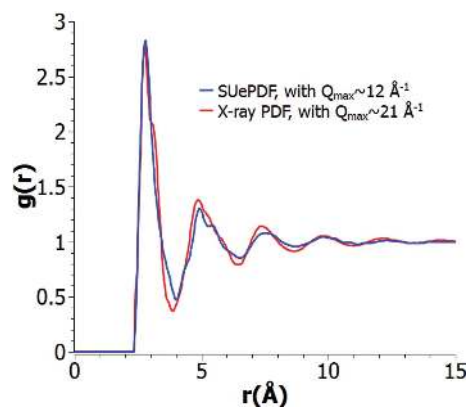
	Bond lengths (Å)			Coordination numbers		
	Si–O	O–O	Si–Si	(Si–O) & (O–Si)	O–O	Si–Si
<i>SUePDF</i> thin area	1.63	2.62	3.40	2.49	5.49	3.89
<i>SUePDF</i> thick area	1.63	2.64	3.41	2.69	5.18	4.44
References	1.62 (N)	2.63 (N)	3.10 (N)	2.66 (MD)	6.07 (MD)	3.78 (MD)

### 4.3. Metallic glass

Fig. 13 shows an ED-based PDF of  $\text{Cu}_{0.475}\text{Zr}_{0.475}\text{Al}_{0.05}$  metallic glass in comparison with the corresponding X-ray data (Kaban *et al.*, 2015). The ED data have  $[Q_{\min}, Q_{\max}]_{\text{ED}} = [0.9, 12.2] \text{ \AA}^{-1}$ , while the X-ray data have  $[Q_{\min}, Q_{\max}]_{\text{X}} = [0.7, 21.1] \text{ \AA}^{-1}$ . Note that the difference in  $Q$  range can cause different termination effects on the ED-based PDF and X-ray PDF. Besides this, the sample amount in a TEM-based ED experiment is much less than the amount in an X-ray experiment, suggesting that the information given by the ED-based PDF is more local than the information given by the X-ray PDF.

## 5. Environment and distribution of *SUePDF*

*SUePDF* is written in the MATLAB language and is compiled as a stand-alone GUI program, which only requires the free MATLAB Runtime R2015a (or newer) environment (<http://www.mathworks.com>) installed on a 64 bit Windows platform (Windows XP or newer is recommended). The installer (*SUePDF\_Installer.exe* file) will automatically download (internet connection required) and install the MATLAB Runtime environment before installing *SUePDF* when executed.



**Figure 13** ED-based PDF (blue) of  $\text{Cu}_{0.475}\text{Zr}_{0.475}\text{Al}_{0.05}$  metallic glass obtained using *SUePDF* compared with the corresponding X-ray data (Kaban *et al.*, 2015) shown in red.



*SUePDF* is distributed as free software for academic users, with an installer file and a manual document available for free download at <https://osf.io/c2jq8/>.

### 6. Summary

We have described the implementation of *SUePDF*, a GUI program dedicated to structural analysis based on electron diffraction data. *SUePDF* facilitates TEM-based structural studies of amorphous materials and nanoparticles by converting the electron diffraction data in the reciprocal space into quantitative PDFs in the direct space. *SUePDF* employs the scattering physics of electrons as well as the physical meaning of PDFs to achieve reliable data normalization. Noise is treated in *SUePDF* by band-pass Fourier filtering that also allows the evaluation of uncertainties caused by experimental conditions and data treatment procedures. Examples of using *SUePDF* to obtain quantitative PDFs of crystalline Au nanoparticles, amorphous silica and amorphous  $\text{Cu}_{0.475}\text{Zr}_{0.475}\text{Al}_{0.05}$  metallic glass have been demonstrated.

### Acknowledgements

The Knut and Alice Wallenberg (KAW) Foundation is acknowledged for providing the electron microscopy facilities and financial support under the project 3DEM-NATUR.

### References

- Abeykoon, M., Malliakas, C. D., Juhás, P., Bozin, E. S., Kanatzidis, M. G. & Billinge, S. J. L. (2012). *Z. Kristallogr.* **227**, 248–256.
- Ankele, J., Mayer, J., Lamparter, P. & Steeb, S. (2005). *Z. Naturforsch. Teil A*, **60**, 459–468.
- Anstis, G. R., Liu, Z. & Lake, M. (1988). *Ultramicroscopy*, **26**, 65–69.
- Brown, P. J., Fox, A. G., Maslen, E. N., O'Keefe, M. A. & Willis, B. T. M. (2006). *International Tables for Crystallography*, Vol. C, *Mathematical, Physical and Chemical Tables*, 1st online ed., edited by E. Prince, ch. 6.1, pp. 554–595. Chester: International Union of Crystallography.
- Cockayne, D. J. H. (2007). *Annu. Rev. Mater. Res.* **37**, 159–187.
- Cockayne, D. J. H., Chen, Y., Li, G. & Borisenko, K. B. (2010). *J. Phys. Conf. Ser.* **241**, 012006.
- Cockayne, D. J. H. & McKenzie, D. R. (1988). *Acta Cryst.* **A44**, 870–878.
- Cowley, J. M. (1969). *Acta Cryst.* **A25**, 129–134.
- Cowley, J. M. (1992). *Electron Diffraction Techniques*, Vol. 1. Oxford University Press.
- Egami, T. & Billinge, S. J. L. (2002). *Underneath the Bragg Peaks: Structural Analysis of Complex Materials*. Amsterdam: Elsevier Science.
- Egerton, R. F. (2011). *Electron Energy-Loss Spectroscopy in the Electron Microscope*, 3rd ed. New York: Springer.
- Gilbert, B. (2008). *J. Appl. Cryst.* **41**, 554–562.
- Hauschild, D. (2009). *EDDA*, <http://dirkhauschild.de/edda.php>.
- Hoang, V. V. (2007). *J. Phys. Chem. B*, **111**, 12649–12656.
- Ishimaru, M., Hirata, A., Naito, M., Bae, I.-T., Zhang, Y. & Weber, W. J. (2008). *J. Appl. Phys.* **104**, 033503.
- Juhás, P., Davis, T., Farrow, C. L. & Billinge, S. J. L. (2013). *J. Appl. Cryst.* **46**, 560–566.
- Kaban, I., Jóvári, P., Escher, B., Tran, D. T., Svensson, G., Webb, M. A., Regier, T. Z., Kokotin, V., Beuneu, B., Gemming, T. & Eckert, J. (2015). *Acta Mater.* **100**, 369–376.
- Keen, D. A. & Dove, M. T. (1999). *J. Phys. Condens. Matter*, **11**, 9263–9273.
- Kirkland, E. J. (2010). *Advanced Computing in Electron Microscopy*, 2nd ed. New York: Springer Science + Business Media.
- Kodama, K., Iikubo, S., Taguchi, T. & Shamoto, S. (2006). *Acta Cryst.* **A62**, 444–453.
- McBride, M. (2003). *J. Non-Cryst. Solids*, **318**, 233–238.
- Mitchell, D. R. & Petersen, T. C. (2012). *Microsc. Res. Tech.* **75**, 153–163.
- Mott, N. F. & Massey, H. S. W. (1965). *The Theory of Atomic Collisions*, 3rd ed. Oxford: Clarendon Press.
- Mu, X., Neelamraju, S., Sigle, W., Koch, C. T., Totò, N., Schön, J. C., Bach, A., Fischer, D., Jansen, M. & van Aken, P. A. (2013). *J. Appl. Cryst.* **46**, 1105–1116.
- Peterson, P. F., Božin, E. S., Proffen, Th. & Billinge, S. J. L. (2003). *J. Appl. Cryst.* **36**, 53–64.
- Peterson, P. F., Gutmann, M., Proffen, Th. & Billinge, S. J. L. (2000). *J. Appl. Cryst.* **33**, 1192.
- Proffen, T., Billinge, S. J. L., Egami, T. & Louca, D. (2003). *Z. Kristallogr.* **218**, 132–143.
- Ruskin, R. S., Yu, Z. & Grigorieff, N. (2013). *J. Struct. Biol.* **184**, 385–393.
- Takagi, T., Ohkubo, T., Hirotsu, Y., Murty, B. S., Hono, K. & Shindo, D. (2001). *Appl. Phys. Lett.* **79**, 485–487.
- Tewes, M., Zweck, J. & Hoffmann, H. (1994). *J. Phys. Condens. Matter*, **6**, 835–848.
- Tran, D. T., Svensson, G. & Tai, C.-W. (2016). *arXiv:1602.08078* [cond-mat.mtrl-sci].
- Uyeda, R. (1968). *Acta Cryst.* **A24**, 175–181.
- Warren, B. E. (1990). *X-ray Diffraction*. New York: Dover Publications.