JOURNAL OF APPLIED CRYSTALLOGRAPHY
ISSN 1600-5767

Illustrated formalisms for total scattering data: a guide for new practitioners. Corrigendum and addendum

Peter F. Peterson[a,b]* and David A. Keen[c]*

[a]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [b]Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, and [c]ISIS Facility, Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire, United Kingdom. *Correspondence e-mail: petersonpf@ornl.gov, david.keen@stfc.ac.uk

Received 17 June 2021
Accepted 26 July 2021

Edited by J. M. García-Ruiz, Instituto Andaluz de Ciencias de la Tierra, Granada, Spain

Keywords: total scattering; pair distribution function.

Errors and ambiguities in the article by Peterson, Olds, McDonnell & Page [*J. Appl. Cryst.* (2021), **54**, 317–332] are corrected and clarified, respectively.

In the article *Illustrated formalisms for total scattering data: a guide for new practitioners* (Peterson *et al.*, 2021) the authors provide a detailed comparison between the various notations and formalisms used by the total scattering/pair distribution function (PDF) community. The paper repeats the relationships already established by Keen (2001), presents look-up tables for easy conversion between functions, and provides graphical examples based on calculated neutron scattering functions of liquid argon and powdered MnO from molecular dynamics simulations and the cubic crystal structure, respectively. However, there is a confusing choice of units when some of the functions are presented graphically, leading to a mis-labelling the *y* axes of several of the figures. Furthermore, the low-Q limits have been defined incorrectly, such that they are only applicable for monoatomic materials. Given the pedagogical nature of the article (Peterson *et al.*, 2021), we felt it necessary to provide this corrigendum to clarify any unintended confusion. For greater clarity, we provide a worked example based on experimental GEM data from $BaTiO_3$ (Senn *et al.*, 2016) in support of these revisions.

As already pointed out by Keen (2001) and Peterson *et al.* (2021), in this field different communities lay claim to the same function names for subtly different definitions. In order to be completely explicit in this corrigendum we will use the subscript 'K' and superscript 'PDF' to clarify multiply defined functions consistent with the sub/superscripting used by Peterson *et al.* (2021) and Keen (2001), respectively. We have not assigned these additional labels to functions [such as $S(Q)$ and $\rho(r)$] that are defined identically by Keen (2001) and Peterson *et al.* (2021). This somewhat tedious notation will be helpful when presenting clarifying points and is detailed in Table 1.

Neutron scattering lengths were typically given in units of 10^{-12} cm, *i.e.* 10 fm units, a natural working unit since neutron cross sections are expressed in barns (10^{-24} cm^2) [for example *International Tables for X-ray Crystallography*, Vol. III (1962)]; the powers of ten, although implicitly present, could be ignored in practice. They are now more commonly tabulated in fm (Sears, 1992) and these are the values used by Peterson *et al.* (2021). This has introduced an inadvertent

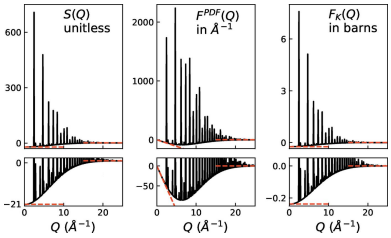


Table 1
Connection of formalisms to those found in previous work.

It is fully expected that future publications will not employ the subscript/superscript used here, but might nonetheless refer to these specific equations when establishing notation or define equations in terms of $S(Q)$ or $\rho(r)$ which have common meaning.

Function	Peterson *et al.* (2021) equation	Keen (2001) equation
$F^{\rm PDF}(Q)$	10	45 (implicit)
$S(Q)$	4	19/20
$F_{\rm K}(Q)$	11/13	9/19
$G^{\rm PDF}(r)$	19	43
$g^{\rm PDF}(r)$	22	41
$\rho(r)$	16	46
$G_{\rm K}(Q)$	26/27	10/44

scaling error in Figs. 1(*e*), 1(*f*), 2(*e*), 2(*f*), 3 [$G_{\rm K}(r)$ and $T(r)$] and 4 [$G_{\rm K}(r)$ and $T(r)$] of Peterson *et al.* (2021); the functions in the figures with *y* axes labelled 'barn' or 'barn/Å^2' should be 100× smaller than presented therein. Those functions that are further scaled by $\langle b_{\rm coh}\rangle^{-2}$ [*e.g.* $F^{\rm PDF}(Q)$, $G^{\rm PDF}(r)$ or $\rho(r)$] are unaffected as the scattering length units cancel out.

The low-Q limit of total scattering structure factors is related in a complex way to thermodynamic functions and fluctuations [see for example Table 30.2 of Cusack (1987)] most of which, for 'well behaved' non-magnetic systems, are assumed to give a small correction to the expected level. It can be further complicated by other sources of low-Q 'small-angle' scattering, such as from longer-ranged scattering density variations (*e.g.* from nanoparticles) and magnetism, all of which are ignored in the following derivations. The low-Q limit of the partial total scattering structure factors, $A_{ij}(Q)$, is

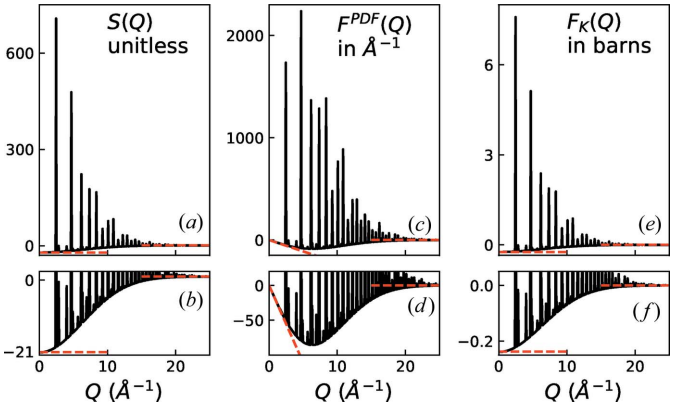


Figure 1
Redrawn version of Fig. 1 of Peterson *et al.* (2021) with corrected limiting values and scaling. The figure shows the comparison between various reciprocal-space neutron total scattering data from MnO: (*a*), (*b*) $S(Q)$; (*c*), (*d*) $F^{\rm PDF}(Q)$; and (*e*), (*f*) $F_{\rm K}(Q)$. The upper plots show an overview of the various functions, although the height of the Bragg peaks is mostly dictated by the width of the resolution function used (the same width was used in all plots). The asymptotes at high and low Q are highlighted with red dashed lines and pertinent values are given in Table 4. The low-Q limit of $F^{\rm PDF}(Q)$ is a sloped line which only remains linear for an extended range of Q because these are synthetic data with infinite sample size (and no magnetic contribution). However, the same behaviour holds true for highly ordered $BaTiO_3$ in Fig. 2. Note that contributions from the terms involving η for MnO are effectively zero on the scale of these figures.

Table 2
Limits of reciprocal-space functions.

This is a corrected version of Table 3 of Peterson *et al.* (2021) with all low-Q limits changed and monatomic low-Q limits added. The 'low-Q limit' column is correct for all materials (noting that these expressions are only approximately followed for measurements of real materials and assume there is no additional 'small-angle' scattering) and the 'monatomic low-Q limit' largely replicates the values in the original article. For solids and many liquids, η is usually considered sufficiently close to zero to be ignored, especially given that these low-Q limits should only be treated as indicative for real materials.

Function	Low Q	Monatomic low Q	High Q	Units
$S(Q)$	$\eta+1-\langle b_{\rm coh}^2\rangle/\langle b_{\rm coh}\rangle^2$	η	1	Unitless
$F^{\rm PDF}(Q)$	0	0	0	Å^{-1}
$F_{\rm K}(Q)$	$\langle b_{\rm coh}\rangle^2\eta-\langle b_{\rm coh}^2\rangle$	$b_{\rm coh}^2(\eta-1)$	0	Barn

given in equation (13) of Keen (2001) and by McGreevy & Mitchell (1982) and is frequently shown in experimental data (Fischer *et al.*, 2006; Bowron *et al.*, 2006):

$$\lim_{Q\to 0} A_{ij}(Q) - 1 \simeq \rho_0 k_{\rm B} T \kappa_{\rm T} - \delta_{ij}/c_i, \qquad (1)$$

where the symbols have their usual meanings [*e.g.* in the text following equation (8) of Peterson *et al.* (2021)]. The first term on the right-hand side of equation (1) is the isothermal compressibility term, η, of Bhatia & Thornton (1970). It is usually small relative to the other terms in the low-Q limit of the total scattering structure factors (given below) and frequently treated as zero for data normalization purposes (although note that this may not always be the case and it may provide useful physical insight; Cusack, 1987). $A_{ij}(Q)$ is defined identically by Keen (2001) and Peterson *et al.* (2021). Since

$$F_{\rm K}(Q) = \sum_{i,j=1}^{n} c_i c_j \bar{b}_i \bar{b}_j \big[A_{ij}(Q) - 1\big], \qquad (2)$$

this gives rise to the following low-Q limits for the total structure factors,

$$\lim_{Q\to 0} F_{\rm K}(Q) \simeq \left(\sum_{i=1}^{n} c_i \bar{b}_i\right)^2 \eta - \sum_{i=1}^{n} c_i \bar{b}_i^2 = \langle b_{\rm coh}\rangle^2 \eta - \langle b_{\rm coh}^2\rangle, \qquad (3)$$

and with $S(Q) = F_{\rm K}(Q)/\langle b_{\rm coh}\rangle^2 + 1$ yields

$$\lim_{Q\to 0} S(Q) \simeq \eta - \frac{\langle b_{\rm coh}^2\rangle}{\langle b_{\rm coh}\rangle^2} + 1. \qquad (4)$$

Here we have explicitly propagated η through the equations above, rather than using it more flexibly [as was done by Keen (2001)], whilst still bearing in mind that the definition in equation (1) is not valid in all circumstances (Cusack, 1987). These limits are different from those given in equations (7) and (64) and Table 3 of Peterson *et al.* (2021), which are only valid for monatomic systems. The corrected behaviour for various total scattering structure factors is summarized in Table 2, and the limits are recalculated using the average neutron scattering length constants in Table 3 to give the results in Table 4. Taking these points together, and as an example, we show in Fig. 1 a corrected version of Fig. 1 of Peterson *et al.* (2021) with very different low-Q limiting values and a much-reduced *y*-axis scale for $F_{\rm K}(Q)$. Fig. 2 and the plots

Table 3
Average neutron scattering length constants for selected materials calculated using neutron scattering lengths and cross sections provided by Sears (1992).

Note that 1 barn = 10^{-24} cm^2 = 100 fm^2.

Peterson *et al.* (2021) notation	Keen (2001) notation	MnO value (barn)	$BaTiO_3$ value (barn)	SiO_2 value (barn)
$\langle b_{\rm tot}^2\rangle$	$\sum_{i=1}^{n} c_i\overline{b_i^2}$	0.254	0.325	0.282
$\langle b_{\rm coh}^2\rangle$	$\sum_{i=1}^{n} c_i\bar{b}_i^2$	0.238	0.277	0.282
$\langle b_{\rm coh}\rangle^2$	$(\sum_{i=1}^{n} c_i\bar{b}_i)^2$	0.011	0.145	0.276

of $G_{\rm K}(r)$ and $T(r)$ in Figs. 3 and 4 in the original article need to be similarly modified but are not included here.

The above discussion also highlights another important point. Although the scaled functions [$S(Q)$, $F^{\rm PDF}(Q)$ and $G^{\rm PDF}(r)$ *etc*., which are divided through by $\langle b_{\rm coh}\rangle^2$], are useful when comparing with models and calculations, the functions $F_{\rm K}(Q)$, $G_{\rm K}(r)$ *etc*., which are not scaled by $\langle b_{\rm coh}\rangle^{-2}$, permit a much more direct and unambiguous assessment of absolute data normalization when correcting experimental data. $F^{\rm PDF}(Q)$ and $G^{\rm PDF}(r)$ are more challenging in this regard as their respective Q- and r-dependent asymptotes to the origin make determination of the low-Q and low-r trends less obvious 'by eye'. $S(Q)$ should also be used cautiously; although it is unitless this hides the fact that scattering factors are incorporated within the function, and even though many $S(Q)$ have a low-Q limit that is close to zero this does not mean that zero is the limiting value by definition.

As a worked example of this, we show in Fig. 2 data from $BaTiO_3$ measured on GEM (Hannon, 2005) at 15 K (Senn *et al.*, 2016), which have been corrected using the *Gudrun* program (Hannon *et al.*, 1990; Soper, 2017). The Ti atoms in $BaTiO_3$ have a negative neutron scattering length and the relevant average scattering constants are listed in Table 3. The high- and low-Q levels of the corrected differential scattering cross section (they should equal $\langle b_{\rm tot}^2\rangle$ and tend to $\sim\langle b_{\rm tot}^2\rangle - \langle b_{\rm coh}^2\rangle = 0.048$, respectively) are immediate indicators of the quality of the data correction [as are the limits of $F_{\rm K}(Q)$, *i.e.* after subtraction of $\langle b_{\rm tot}^2\rangle$; see Fig. 2(*a*)]. Typically when correcting data, a 'good' low-Q limit is often much harder to achieve than a 'good' high-Q limit. This is especially the case for time-of-flight neutron diffractometers where data corrections are usually more challenging in the lower-Q regime. The low-Q limit of $S(Q)$ should approximately equal $-\langle b_{\rm coh}^2\rangle/\langle b_{\rm coh}\rangle^2 + 1 = -0.911$ [Fig. 2(*c*)]. The low-r levels of $G_{\rm K}(r) = -\langle b_{\rm coh}\rangle^2$ and $g^{\rm PDF}(r) = 0$ [Figs. 2(*b*) and 2(*d*), respectively]. Here a back-transform correction has been applied to direct these PDFs to their theoretical values for $r <$ 1 Å; encouragingly these values are maintained to much higher r, including in the gaps between the first few low-r peaks. For completeness, plots of $F^{\rm PDF}(Q)$ and $G^{\rm PDF}(r)$ are shown in Figs. 2(*e*) and 2(*f*), respectively.

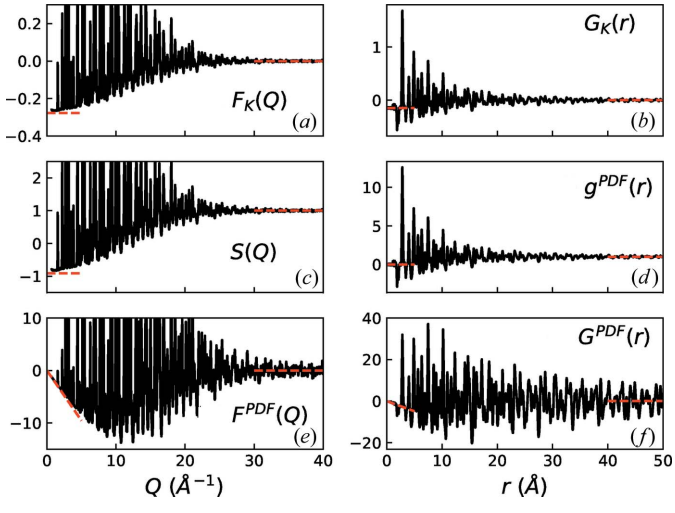


Figure 2
Experimental neutron total scattering functions from $BaTiO_3$ at 15 K (Senn *et al.*, 2016), using ρ_0 = 0.0779 atoms Å^{-3}, determined from Rietveld refinement of the data. The various functions are defined using the equations given in Table 1. The asymptotes at high and low Q (or r, as appropriate) are highlighted using red dashed lines with values for reciprocal-space functions from Table 4; all real-space function limits are clearly seen to be 0 or 1 except for the low-r limit of $G_{\rm K}(r)$, which is $-\langle b_{\rm coh}\rangle^2 = -0.145$ barn.

Table 4
List of reciprocal-space limits for selected materials with the assumption that $\eta = 0$.

Although many materials have an $S(Q)$ with a limiting value at low Q close to zero (*e.g.* SiO_2), for materials containing elements with negative neutron coherent scattering lengths (*e.g.* MnO and $BaTiO_3$) $\lim_{Q\to 0} S(Q)$ is often far from zero.

Material	Function	Low Q	High Q
MnO	$S(Q)$	−21.1	1
	$F^{\rm PDF}(Q)$	0	0
	$F_{\rm K}(Q)$	−0.237	0
$BaTiO_3$	$S(Q)$	−0.911	1
	$F^{\rm PDF}(Q)$	0	0
	$F_{\rm K}(Q)$	−0.277	0
SiO_2	$S(Q)$	−0.022	1
	$F^{\rm PDF}(Q)$	0	0
	$F_{\rm K}(Q)$	−0.282	0

Finally, we note an inconsistency in the discussion of symmetric PDF peaks by Peterson *et al.* (2021). The different r dependencies mean that it is not possible for all definitions of the PDF function to show symmetric peaks centred at the average pairwise distances. A symmetric and centred peak in $G^{\rm PDF}(r)$ will not be symmetric and centred in $g^{\rm PDF}(r)$ [equivalent to $G'_{\rm K}(r)$ of Keen (2001)]. Symmetric peak fitting should only be carried out using PDF functions such as $G^{\rm PDF}(r)$ or $D(r)$ (Olds *et al.*, 2018).

We have worked together on this corrigendum to try to ensure that these corrections to the Peterson *et al.* (2021) paper are clear and that the explicitly labelled functions herein mean that the relational expressions first established by Keen (2001) are not confused by the subtly redefined functions presented by Peterson *et al.* (2021), thus undermining the purpose of both papers. Total scattering notation has evolved over time since the equations were first conceived of by Zernike and Prins in 1927 (see a recent review and references

therein; Keen, 2020), but it has stabilized over the past 20 years within the now mature total scattering community, in part through the cross-referencing of equations following Keen (2001). We as a community have a responsibility to ensure that we do not further compound any perceived notational confusion we might be trying to mitigate. Hopefully Keen (2001) and Peterson *et al.* (2021) with this corrigendum article will continue to provide the necessary clarity in total scattering function definitions.

Acknowledgements

The authors of this contribution acknowledge Daniel Olds, Marshall McDonnell and Katherine Page for feedback and comments during the production of this corrigendum. They are in complete agreement with the information provided herein.

Funding information

Funding for this research was provided by the US Department of Energy, Office of Science (contract No. DE-AC05-00OR22725).

References

Bhatia, A. & Thornton, D. (1970). *Phys. Rev. B*, **2**, 3004–3012.

Bowron, D., Finney, J., Hallbrucker, A., Kohl, I., Loerting, T., Mayer, E. & Soper, A. (2006). *J. Chem. Phys.* **125**, 194502.

Cusack, N. (1987). *The Physics of Structurally Disordered Matter: an Introduction.* Bristol: Adam Hilger/IOP Publishing.

Fischer, H. E., Barnes, A. C. & Salmon, P. S. (2006). *Rep. Prog. Phys.* **69**, 233–299.

Hannon, A. (2005). *Nucl. Instrum. Methods Phys. Res. A*, **551**, 88–107.

Hannon, A., Howells, W. & Soper, A. (1990). *Inst. Phys. Conf. Ser.* **107**, 193–211.

Keen, D. A. (2001). *J. Appl. Cryst.* **34**, 172–177.

Keen, D. A. (2020). *Crystallogr. Rev.* **26**, 143–201.

McGreevy, R. & Mitchell, E. (1982). *J. Phys. C. Solid State Phys.* **15**, 5537–5550.

Olds, D., Saunders, C. N., Peters, M., Proffen, T., Neuefeind, J. & Page, K. (2018). *Acta Cryst.* A**74**, 293–307.

Peterson, P. F., Olds, D., McDonnell, M. T. & Page, K. (2021). *J. Appl. Cryst.* **54**, 317–332.

Sears, V. F. (1992). *Neutron News*, **3**(3), 26–37.

Senn, M., Keen, D., Lucas, T., Hriljac, J. & Goodwin, A. (2016). *Phys. Rev. Lett.* **116**, 207602.

Soper, A. (2017). *Gudrun – Routines for Reducing Total Scattering Data*, https://www.isis.stfc.ac.uk/Pages/Gudrun.aspx.

JOURNAL OF APPLIED CRYSTALLOGRAPHY

ISSN 1600-5767

Received 16 June 2020
Accepted 27 November 2020

Edited by J. M. García-Ruiz, Instituto Andaluz de Ciencias de la Tierra, Granada, Spain

Keywords: total scattering; pair distribution function.

Illustrated formalisms for total scattering data: a guide for new practitioners

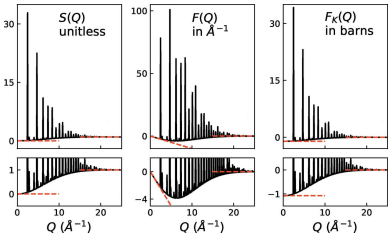
Peter F. Peterson,[a,c]* Daniel Olds,[b]* Marshall T. McDonnell[a] and Katharine Page[c,d]

[a]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, [b]National Synchrotron Light Source-II, Brookhaven National Laboratory, Upton, NY, USA, [c]Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA, and [d]Materials Science and Engineering Department, University of Tennessee, Knoxville, TN, USA. *Correspondence e-mail: petersonpf@ornl.gov, dolds@bnl.gov

The total scattering method is the simultaneous study of both the real- and reciprocal-space representations of diffraction data. While conventional Bragg-scattering analysis (employing methods such as Rietveld refinement) provides insight into the average structure of the material, pair distribution function (PDF) analysis allows for a more focused study of the local atomic arrangement of a material. Generically speaking, a PDF is generated by Fourier transforming the total measured reciprocal-space diffraction data (Bragg and diffuse) into a real-space representation. However, the details of the transformation employed and, by consequence, the resultant appearance and weighting of the real-space representation of the system can vary between different research communities. As the worldwide total scattering community continues to grow, these subtle differences in nomenclature and data representation have led to conflicting and confusing descriptions of how the PDF is defined and calculated. This paper provides a consistent derivation of many of these different forms of the PDF and the transformations required to bridge between them. Some general considerations and advice for total scattering practitioners in selecting and defining the appropriate choice of PDF in their own research are presented. This contribution aims to benefit people starting in the field or trying to compare their results with those of other researchers.

1. Introduction

Pair distribution function (PDF) analysis is a broad term encompassing the use of experimentally generated atomic pair density functions in real space to study short-range order in materials. The approach has its roots in the work of Warren and co-workers, who developed it to study the short- and intermediate-range structure of glasses (Warren, 1934, 1937). Notable extensions have included work on molecular liquids (Narten, 1972), the study of glasses and crystalline materials (Srolovitz *et al.*, 1981), the study of liquids and glasses with neutron diffraction (Wright *et al.*, 1989; Wright, 1990, 1994), the use of high-energy X-rays (Egelstaff, 1967; Root *et al.*, 1986; Neuefeind & Poulsen, 1995; Neuefeind *et al.*, 1996), the adoption of the technique for reverse Monte Carlo (RMC) methods (McGreevy & Pusztai, 1988; McGreevy, 1995, 2001), and the application to nanostructured and disordered crystalline materials (Billinge, 1992; Egami & Billinge, 2012). These advancements and their applications to materials are distinguished in a number of review articles and text books appearing in the past two decades, for liquids and glasses (Neuefeind, 2002; Egelstaff, 2003; Kohara & Suzuya, 2003; Fischer *et al.*, 2005; Kohara *et al.*, 2007; Benmore, 2012), and for disordered crystalline and nanostructured materials



(Egami, 2007; Billinge & Kanatzidis, 2004; Billinge, 2004; Billinge & Levin, 2007; Young & Goodwin, 2011; Egami & Billinge, 2012; Playford *et al.*, 2014; Hou *et al.*, 2018).

The PDF is a probability distribution function that measures the probability of finding pairs of atoms separated by a given distance. There are many different forms of the PDF with subtly different functional forms, units, normalizations and use in research communities, but they all contain the same information: the probability of finding atoms separated by a distance (Dinnebier & Billinge, 2008). For example, from a statistical mechanics definition of the radial distribution function, this probability can be given relative to the ideal gas state, where no correlations exist. Therefore, deviations in such a distribution function will give a factor to multiply the bulk density by to get a local density (McQuarrie, 2000). All the different forms of the PDF give information about the changes in local density with distance and thus insight into the local structure.

The use of PDF analysis has grown exponentially from a specialized technique employed for the study of liquids, glasses and other amorphous materials [where Rietveld (1969) analysis is not possible] to one that encompasses disordered materials more broadly through the study of local atomic structure and disorder in nanocrystalline and crystalline materials. As various material communities have adopted PDF analysis, they have refined the methodology and developed corresponding analysis software to address their specific scientific needs. The weighting and normalization of different features, either from the measured reciprocal-space data or modeled real-space atomic coordinates, has led to no less than eight different published forms of real-space distribution function (Keen, 2001), which can all claim in one way or another to be analogous to the PDF. Despite being functionally similar, the differences between these varied forms has led to some confusion and redundancy across the different communities.

This contribution aims to untangle many of these semantic and terminological confusions through a consistent derivation of the relationships between many different forms of the PDF, demonstrated through simple examples. We reintroduce many fundamental concepts and formalisms, and directly relate them to the physical and distinguishing features they represent.

The equations are presented here with neutron scattering formalism, where there is no Q dependence in the atomic scattering lengths. The added complexity brought by Q-dependent X-ray form factors (Narten, 1972) has been, broadly, addressed in three approaches: explicit corrections during data normalization (Qiu, Thompson & Billinge, 2004), *ad hoc* or approximative corrections (Juhás *et al.*, 2013; Billinge & Farrow, 2013) during data normalization, or leaving the data uncorrected and instead forward calculating the effects in a refined model (Gereben *et al.*, 2007; Tucker *et al.*, 2007). The variability of these different approaches and their implications are outside the scope of this paper. Herein, the presented derivations assume that the reduction of X-ray scattering data successfully mitigates the effects of these form factors.

For the remainder of this text, the authors make the following assumptions which are being made explicit with the aim of aiding the reader even further in seeing the bridge between various conventions. First, the term 'total scattering' was adopted for the PDF method within the past few decades to bring attention to the fact that it provides an examination of both Bragg and diffuse scattering. The modern adoption of this colloquial name should be distinguished from the foundational terminology 'total scattering' in use by the general time-of-flight neutron scattering community (Squires, 2012). Note also that most neutron and X-ray PDF measurements are energy-integrated scattering functions. Throughout this manuscript 'total scattering' will be used to refer to the PDF technique in general. Second, as will be further explained in Section 3, the $I(Q)$ presented here is proportional to the differential cross section. Note that $I(Q)$ is different from the measured intensity often employed in Rietveld refinements, sometimes utilizing the same nomenclature (Rietveld, 1969).

Previous work went into great detail on the derivation and conversion of different functional forms of the PDF, and is widely cited in the community (Keen, 2001; Fischer *et al.*, 2005). Subtle differences in nomenclature of various approaches have led to some confusion in the PDF communities, particularly when converting between the different formalisms. This manuscript expands on this work, beginning with the conventions of the disordered crystalline material communities (Egami & Billinge, 2012) and bridging to other derivations. Visual comparisons of various functional forms guide the reader and frame discussion of each use case. Effort has been made to reference both initial derivations and examples of utilization in scientific literature.

The derivations and examples in this paper are presented first in reciprocal space then in real space. A collection of appendices provide both reference and further detail supporting the derivations. Appendix *A* provides a list of simple conversions between the various real-space functions. Appendix *B* provides details of the molecular dynamics simulations used for the liquid example data. Appendix *C* is an overview of the process of converting from measured intensities to differential cross sections to provide a frame of reference for understanding experimental concerns. Appendix *D* details calculation of the normalized Laue term employed in some formalisms, while Appendix *E* details calculation of the number density. Appendix *F* provides a brief overview of partial structure functions.

2. Methods

To illustrate the different functions contained herein, we have simulated neutron nuclear scattering data for two systems of representative material types: bulk binary oxide manganese oxide (MnO) (Sasaki *et al.*, 1979) for crystalline and disordered materials; and liquid argon (Ar) (Yarnell *et al.*, 1973) for liquid, amorphous and glass materials. Mn has a negative neutron scattering length and O has a positive one, which emphasizes certain differentiating characteristics of various forms of the PDF. MnO has a magnetic structure that will be

Table 1
Summary of structure of MnO used for examples (Sasaki *et al.*, 1979).

The structure is $Fm\bar{3}m$ with a lattice constant of 4.446 (1) Å.

Atom	(x, y, z) (fractional units)	B_{eq} (Å^2)
Mn^{2+}	(0.0, 0.0, 0.0)	0.617 (5)
O^{2-}	(0.5, 0.5, 0.5)	0.72 (1)

ignored to allow for focusing on the atomic structure. Simulated data from liquid argon (Ar) are also included as a monatomic example. Note that the term 'monatomic system' here assumes a single element with a single isotope for neutron scattering and a single element with a single charge state for X-ray scattering.

To generate the presented MnO data, real-space patterns were simulated with the *PDFgui* software (Farrow *et al.*, 2007) using the crystal structure in Table 1. Patterns were calculated for $0 \leq r \leq 160$ Å with a bin width of $\delta r = 0.01$ Å. This pattern was then inverse transformed to generate the presented reciprocal-space data.

To generate the Ar data, molecular dynamics simulations were performed using the *Large-Scale Atomistic/Molecular Massively Parallel Simulator* (*LAMMPS*) open-source code (Plimpton, 1995, 2018). Details of the simulations are provided in Appendix *B*.

3. Reciprocal-space functions

We begin the derivation assuming data in the form of the fully corrected and normalized scattering intensity, $I(Q)$, obtained from experimentally measured intensities. Interestingly, Debye himself, along with Menke, performed the first PDF experiments using X-ray scattering in 1930 to obtain such data (Debye, 1930). The details for a protocol used to reduce such measured data to fully corrected patterns based on literature and reduction software manuals can be found in Appendix *C*. $I(Q)$ can be directly related to a set of atomic coordinates through the Debye (1915) scattering equation:

$$I(Q) = \sum_{\nu\mu} b_{\mathrm{coh},\nu}\, b_{\mathrm{coh},\mu} \frac{\sin(Q r_{\nu\mu})}{Q r_{\nu\mu}}, \tag{1}$$

where $b_{\mathrm{coh},\nu}$ is the coherent scattering length of atom ν and $r_{\nu\mu} = |\mathbf{r}_\nu - \mathbf{r}_\mu|$ is the interatomic pairwise vector of atoms ν and μ (Lovesey, 1986; Farrow & Billinge, 2009; Page *et al.*, 2011). Debye's formalism, slightly modified to include the effects of thermal atomic displacements through a Debye–Waller term $\sigma_{\nu\mu}$ (Debye, 1913; Waller, 1923), is written as

$$I(Q) = \sum_{\nu\mu} b_{\mathrm{coh},\nu}\, b_{\mathrm{coh},\mu} \frac{\sin(Q r_{\nu\mu})}{Q r_{\nu\mu}} \exp\left(\frac{-\sigma_{\nu\mu}^2 Q^2}{2}\right). \tag{2}$$

In this formalism $I(Q)$ is the scattering from the sample as a whole. The total scattering community most commonly employs $I(Q)/N$ (the differential cross section), where N is the number of atoms illuminated in the sample (described further in Appendix *E*).

Peak profile refinement methods (*e.g.* Rietveld analysis) most commonly define the scattering per sample, rather than scattering per atom. This convention can be traced back to Rietveld's initial aims of fitting models against the peak profiles of the relative intensities as directly generated from instrument measurements. Use of an arbitrary scale factor during modeling was convenient and sufficient for this purpose. The total scattering formalism, on the other hand, allows for a fully atomistic model comparison with data. Accurate corrections to remove experiment artifacts are required to compare the data with atomistic models, the importance of which has been shown elsewhere (Egelstaff, 1992; Wright, 1994; Fischer *et al.*, 2005). However, in practice, many practitioners studying crystalline materials also apply an arbitrary scale factor to data during modeling (Farrow *et al.*, 2016). In fact, the departure of a scale factor from unity for standard (known) samples is sometimes applied as a quality criterion for assessing the success of data reduction procedures (Peterson *et al.*, 2003).

When defined only through isotropic atomic displacement parameters (commonly referred to as U_{iso}), the Debye–Waller term $\sigma_{\nu\mu}$ can be written as (Jeong *et al.*, 1999, 2003; Proffen & Billinge, 1999)

$$\sigma_{\nu\mu} = \left(\sigma_\nu^2 + \sigma_\mu^2\right)^{1/2}, \tag{3}$$

where σ_ν and σ_μ are the amplitudes of the uncorrelated thermal motion of atoms ν and μ. This relationship is more complicated in the case of anisotropic atomic displacement (Dunitz *et al.*, 1988; Jeong *et al.*, 2003), but its effect on the normalized intensity is similar: exponential dampening of the Bragg intensities at high Q.

Another form of the normalized and corrected scattering data is the 'structure function', $S(Q)$. This form of the scattering data is employed in the generation of many atomic pair–pair representations of data, which accounts for its widespread description in past work (Yarnell *et al.*, 1973; Billinge & Egami, 1993; Keen, 2001; Peterson *et al.*, 2003; Farrow & Billinge, 2009; Page *et al.*, 2011; Olds *et al.*, 2015) where details of the derivation can be found. The structure function is related to the normalized $I(Q)$ function through the relationship

$$S(Q) = \frac{I(Q)}{N\langle b_{\mathrm{coh}}\rangle^2} - \frac{\langle b_{\mathrm{tot}}^2\rangle - \langle b_{\mathrm{coh}}\rangle^2}{\langle b_{\mathrm{coh}}\rangle^2}, \tag{4}$$

where $\langle b_{\mathrm{tot}}^2\rangle = \langle\sigma_{\mathrm{tot}}\rangle/4\pi$ is the average total scattering power of the system, σ_{tot} is the total cross section and $\langle b_{\mathrm{coh}}\rangle^2$ is the average coherent scattering power of all atoms in the sample. Appendix *D* provides a more complete discussion of the total scattering length term, $\langle b_{\mathrm{tot}}^2\rangle$. The second term in equation (4) contributes a constant factor called the normalized Laue monatomic diffuse scattering term. The normalized Laue term, often written simply as L, is

$$L = \frac{\langle b_{\mathrm{tot}}^2\rangle - \langle b_{\mathrm{coh}}\rangle^2}{\langle b_{\mathrm{coh}}\rangle^2}. \tag{5}$$

Table 2
η for the materials chosen at room temperature (300 K).

The bulk modulus for MnO is taken from the article by Zhang (1999).

Material	$K_0 = 1/\kappa$ (GPa)	ρ_0 (atoms Å^{-3})	η
Ar	0.552	0.02138	0.046
MnO	148	0.0455	0.0013

L is zero in the case of single-element scattering (though not strictly so for neutron scattering, since naturally abundant elemental samples often comprise a mix of isotopes). L is calculated for our simulated examples in Appendix D.

The limits of $S(Q)$ are

$$\lim_{Q\to\infty} S(Q) = 1 \tag{6}$$

and

$$\lim_{Q\to 0} S(Q) = \eta, \tag{7}$$

where η is unitless and proportional to the isothermal compressibility of the sample (Lovesey, 1986; Egelstaff, 1992; Wang *et al.*, 2014). Thermodynamically, η is defined as

$$\eta = k_{\rm B}T\left(\frac{\partial\rho}{\partial P}\right)_T = \rho_0\kappa k_{\rm B}T, \tag{8}$$

where κ is the isothermal compressibility (equal to the inverse of the bulk modulus, K_0), ρ_0 is the number density, $k_{\rm B}$ is Boltzman's constant, T is temperature and P is pressure. This is only strictly correct for monatomic, homogeneous, isotropic systems and is incorrect for a fluid close to its critical point. Additional details and references for more complex cases such as mixtures of molecular liquids and ions in aqueous solution are given by Fischer *et al.* (2005). The isothermal compressibility can be calculated as

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = \frac{\langle\sigma_v^2\rangle}{Vk_{\rm B}T}, \tag{9}$$

where V is the volume and σ_v^2 is the variance of the volume. η is often negligible (Bhatia & Thornton, 1970; Wagner, 1985; Egelstaff, 1992; Keen, 2001), as shown in Table 2 for our simulated examples. Note that in cases where nanostructured features exist (such as materials where small-angle scattering is present) the measured low-Q behavior will deviate (Mildner & Carpenter, 1984; Farrow & Billinge, 2009; Olds *et al.*, 2015).

The 'reduced total scattering structure function' is defined as

$$F(Q) = Q[S(Q) - 1]. \tag{10}$$

This representation of reciprocal-space data has a limit of 0 at high Q and is linearly weighted by Q, such that noise and resolution are highlighted features that have dramatic effects on the resultant real-space PDF). Another advantage to this formalism is that associated uncertainties increase linearly with Q (Egami & Billinge, 2012; Olds *et al.*, 2018).

In the derivations of Keen (2001), a similar function also referred to as $F(Q)$ is presented. This alternative function, here referred to as $F_{\rm K}(Q)$, is scaled by $\langle b_{\rm coh}\rangle^2$ and not by Q. Thus, these three reciprocal-space function are related as

$$F_{\rm K}(Q) = \langle b_{\rm coh}\rangle^2[S(Q) - 1] = \frac{\langle b_{\rm coh}\rangle^2}{Q}F(Q). \tag{11}$$

The normalized and corrected intensity, $I(Q)$, is related to $F(Q)$ and $F_{\rm K}(Q)$ as

$$F(Q) = \frac{Q}{\langle b_{\rm coh}\rangle^2}\left[\frac{I(Q)}{N} - \langle b_{\rm tot}^2\rangle\right] \tag{12}$$

and

$$F_{\rm K}(Q) = \frac{I(Q)}{N} - \langle b_{\rm tot}^2\rangle. \tag{13}$$

A visual comparison of $S(Q)$, $F(Q)$ and $F_{\rm K}(Q)$ is shown for the case of MnO in Fig. 1 and for the case of Ar in Fig. 2. A

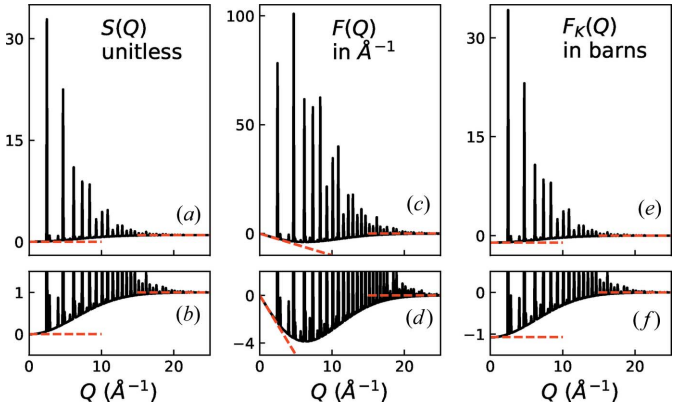


Figure 1
Comparison of reciprocal-space representations of MnO total scattering data: (*a*), (*b*) $S(Q)$, (*c*), (*d*) $F(Q) = Q[S(Q) - 1]$, and (*e*), (*f*) $F_{\rm K}(Q) = \langle b_{\rm coh}\rangle^2[S(Q) - 1]$. The upper plots show an overview of the various functions. The asymptotes are highlighted with dashed lines. In this specific case, $\langle b_{\rm coh}\rangle^2 = 1.074$ fm^2 such that the difference between $S(Q)$ and $F_{\rm K}(Q)$ coincidentally appears to be a vertical shift of one (1). Also note that for MnO $\eta = 0.0013$, which appears to be zero on the scale of this figure.

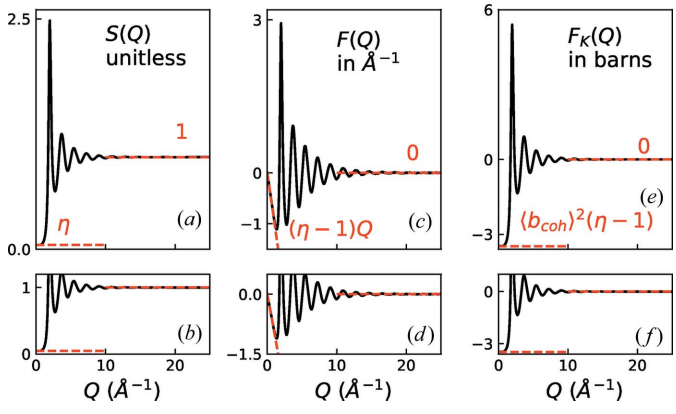


Figure 2
Comparison of reciprocal-space representations of Ar total scattering data: (*a*), (*b*) $S(Q)$, (*c*), (*d*) $F(Q) = Q[S(Q) - 1]$, and (*e*), (*f*) $F_{\rm K}(Q) = \langle b_{\rm coh}\rangle^2[S(Q) - 1]$. The upper plots show an overview of the various functions. The asymptotes are highlighted with dashed lines.

Table 3
Limits of reciprocal-space functions.

Although atoms are normally not included in the unit definition, they are included here for clarity.

Function	Low-Q behavior	High-Q behavior	Units
$I(Q)$	$N[\langle b_{\rm coh}\rangle^2(\eta-1)+\langle b_{\rm tot}^2\rangle]$	$N\langle b_{\rm tot}^2\rangle$	Barn atom
$S(Q)$	η	1	Unitless
$F(Q)$	$(\eta-1)Q$	0	Å^{-1}
$F_{\rm K}(Q)$	$\langle b_{\rm coh}\rangle^2(\eta-1)$	0	Barn

summary of the limiting behaviors of these functions and $I(Q)$ can be found in Table 3.

4. Real-space distribution functions

The pair distribution function is a general concept describing the distribution of distances between pairs of objects contained in a volume. Zernike & Prins (1927) were the first to report the theoretical expression for the atomic density at a given separation in real space via their Fourier transform relationship, leading to the origin of the PDF formalisms. Yet, throughout the years of literature on the PDF, many different functional forms have spawned from this origin.

When defined independently of the atomic origin, μ, this is termed a radial distribution function (RDF), an entity that finds prevalent use as a descriptor for the atomic structure of amorphous, liquid, disordered and nanocrystalline materials. The same name can be associated with different functional forms, ever increasing confusion. For example, the name 'radial distribution function' is associated with both equation (14) and equation (22) in previous literature (Thorpe *et al.*, 1998; McQuarrie, 2000). In this section, we derive and relate a number of functions used in various research communities for representing real-space PDFs (see Figs. 3 and 4 below). These are generally related by multiplicative or additive constants and thus contain the same underlying information. We will explain some of the relative merits and related preferences for these formalisms at the close of the section.

A conceptually different quantity is the RDF which, containing no relationship to scattering weights and thus not directly measurable, is presented here for comparison. We begin by defining a configuration of N atoms arranged such that each atom has a position defined through the vector $\mathbf{r}_\nu$. The interatomic distance between any pair of atoms, μ and ν, is thus $r_{\mu\nu}=|\mathbf{r}_\nu-\mathbf{r}_\mu|$. An unweighted radial distribution function, labeled here RDF(r), can be constructed through the sum of Dirac delta functions, δ, which describe the full set of these pair–pair distances [of which there will be $N(N-1)/2$ total pairs]. RDF(r) can be written as

$$\mathrm{RDF}(r)=\frac{1}{N}\sum_{\mu\nu}\delta(r-r_{\mu\nu}). \tag{14}$$

A radial PDF can be generated from the measured scattering intensities of various physical measurements, including light scattering, electron diffraction, X-ray diffraction and neutron diffraction, with the last three all used to produce atom–atom PDFs. RDF(r) is straightforward to calculate but is only straightforward to measure with monatomic systems. More easily measured, radiation-specific PDFs can be calculated from atomic models by accounting for the scattering power of each atom. This results in the weighted radial distribution function, $R(r)$, defined as

$$R(r)=\frac{1}{N}\sum_{\mu\nu}\frac{b_{{\rm coh},\mu}b_{{\rm coh},\nu}}{\langle b_{\rm coh}\rangle^2}\delta(r-r_{\mu\nu}). \tag{15}$$

For the case of monatomic systems, the weighting prefactor becomes unity and $R(r)$ simplifies to the equation for RDF(r).

A similar formalism often encountered is the density function, $\rho(r)$, which is the radial distribution function normalized by the surface area of a sphere of radius r, such that (Warren, 1990)

$$\rho(r)=\frac{R(r)}{4\pi r^2}. \tag{16}$$

For isotropic and three-dimensional systems, the density function can be directly related to $S(Q)$ through the following pair of transforms (Warren, 1990; Billinge, 1992):

$$S(Q)-1=\frac{4\pi}{Q}\int_0^\infty[\rho(r)-\rho_0]\,r\sin(Qr)\,{\rm d}r \tag{17}$$

and

$$\rho(r)=\rho_0+\frac{1}{2\pi^2 r}\int_0^\infty Q[S(Q)-1]\sin(Qr)\,{\rm d}Q. \tag{18}$$

The heavily used form of the PDF encountered in disordered crystalline material literature is the reduced pair distribution function, $G(r)$ (Egami & Billinge, 2012), which is defined in relation to the density function as

$$G(r)=4\pi r\big[\rho(r)-\rho_0\gamma_0(r)\big]. \tag{19}$$

Here, ρ_0 is the average number density of N atoms in the volume V such that $\rho_0=N/V$, and $\gamma_0(r)$ is the characteristic shape function or nanoparticle form factor (Guinier & Fournet, 1955; Azaroff, 1968; Farrow & Billinge, 2009; Olds *et al.*, 2015). In the case of bulk materials, $\gamma(r)=1.0$, and thus the term is often neglected in the literature.

This reduced pair distribution function can be generated from reciprocal-space data via the sine transform of $F(Q)$, such that

$$G(r)=\frac{2}{\pi}\int_0^\infty F(Q)\sin(Qr)\,{\rm d}Q. \tag{20}$$

Therefore, the Fourier inversion theorem holds that

$$F(Q)=\int_0^\infty G(r)\sin(Qr)\,{\rm d}r. \tag{21}$$

An alternative formalism of the PDF often encountered in studies of amorphous and liquid materials is $g(r)$. It is frequently called the pair distribution function by the liquids/

amorphous community and the pair density function by the disordered crystalline materials community (Benmore, 2012). $g(r)$ is functionally identical to the density function $\rho(r)$; however, it has been scaled by the average number density, resulting in the relationship

$$g(r) = \rho(r)/\rho_0. \tag{22}$$

Note that the $g(r)$ function is related to isothermal compressibility, defined in equation (9), via

$$\kappa = \frac{1}{\rho_0 k_B T}\left[1 + \rho_0 \int_0^\infty 4\pi r^2 [g(r) - 1]\,\mathrm{d}r\right]. \tag{23}$$

Up to this point, all described atomic PDFs have assumed a sum over all atom–atom pairs in a defined volume. However, one can define a 'partial PDF', $g_{\mu\nu}(r)$, which includes contributions from only those atoms in a given pair type. By definition, the sum of all possible partial PDFs will reconstruct the corresponding all-atom PDF. The most common convention requires that 'Faber–Ziman partial structure factors' be calculated for each atom pair. Yet other formalisms exist, each with their respective advantages. The Bhatia–Thornton formalism is an alternative representation of the system as the mean square fluctuations in the particle number, fluctuations in concentration and the correlation between these two correlations (Bhatia & Thornton, 1970). These can be, for a two-component system, directly mapped to Faber–Ziman using the equations in Bhatia and Thornton's seminal 1970 work. The Ashcroft–Langreth formalism is another that is commonly used in theoretical and computational work owing to its connection to direct correlation functions (Ashcroft & Langreth, 1967). The mapping of the Ashcroft–Langreth to the Faber–Ziman equations is most readily accessible in equation 2.35 of the review paper by Fischer *et al.* (2005) The connection to the Faber–Ziman partial structure factors and both the total scattering structure factor and partial pair distribution functions is presented in Appendix *F*. The weighted sum of the partial PDFs will result in a $g(r)$ such that

$$g(r) = \sum_\mu \sum_\nu W_{\mu\nu}\, g_{\mu\nu}(r), \tag{24}$$

where $W_{\mu\nu}$ is the associated weighting factor for the pair of atoms μ and ν. Note that $g_{\mu\nu}(r)$ is a 'true' distribution function as it does not include weighting by scattering lengths. However, this is not a distribution function in the statistical sense as the normalization is

$$\int \sum_\mu \sum_\nu g_{\mu\nu}(r) = N - 1 \simeq N \tag{25}$$

rather than one (McQuarrie, 2000). Different communities employ different normalization schemes for these weighting factors. Some communities will add an 'x' or 'n' superscript to $g(r)$ to denote the weighting. Herein, we normalize them such that $\sum_\mu \sum_\nu W_{\mu\nu} = 1$ unless explicitly noted otherwise. For monatomic systems, the weighting factor is always one.

A form of confusion within the greater PDF community is the differences between the reduced pair distribution function, $G(r)$, and the total radial distribution function, which is often also labeled $G(r)$ (Keen, 2001). For clarity, we here refer to the total radial distribution function as $G_K(r)$. This form of the PDF is constructed from the sum of all partials, $g_{\mu\nu}(r)$, weighted according to concentration of atomic species, c, and associated coherent scattering power, b_{coh}, such that

$$G_K(r) = \sum_\mu \sum_\nu c_\mu c_\nu b_{coh,\mu} b_{coh,\nu}\big[g_{\mu\nu}(r) - 1\big]. \tag{26}$$

The relationship between $G_K(r)$ and $G(r)$ is therefore

$$G_K(r) = \frac{\langle b_{coh}\rangle^2}{4\pi\rho_0}\frac{G(r)}{r}. \tag{27}$$

An important note, and an example of where confusion can occur for new practitioners, is that $G_K(r)$ in this work is equivalent to $G(R)$ in the review paper of Fischer *et al.* (2005) This can be seen by comparing equation (27) in this work with equation 2.40 of Fisher *et al.*

A third variation commonly found in the crystalline PDF community, referred to as the differential correlation function, $D(r)$ (Tucker *et al.*, 2007, 2017), is identical to $G(r)$ apart from a constant scaling factor such that

$$D(r) = \langle b_{coh}\rangle^2 G(r). \tag{28}$$

Related to the previous note about $G_K(R)$, $D(r)$ here is different from the $D(r)$ found in the review paper of Fischer *et al.* (2005) The $D(r)$ in this work is equivalent to $G(R)$ in the review paper of Fischer *et al.* This can be seen by comparing $G(r)$ in equation (34) in this work with equation 2.26 of Fisher *et al.*

Another version of the PDF, primarily used in the liquids and glass community and referred to as the total correlation function, is $T(r)$ (Soper, 1989; Hannon *et al.*, 1990). $T(r)$ is related to $G(r)$ as

$$T(r) = \langle b_{coh}\rangle^2\big[G(r) + 4\pi r^2\rho_0\big]. \tag{29}$$

Additional minor variations of pair distribution function relationships can be found, but while some do surface occasionally in the modern literature, many are no longer actively utilized.

Figs. 3 and 4 graphically display examples of $R(r)$, $\rho(r)$, $g(r)$, $G_K(r)$, $G(r)$ and $T(r)$ for the cases of crystalline MnO and liquid Ar, respectively. The inherent information content of all forms is the same. All PDFs show atomic pair–pair correlations as peaks centered at average pairwise distances in real space, with the height of these peaks informing on the frequency of these pairwise distances (often with the scattering power of atoms involved) and the widths related to the distribution of the pairwise distances. The functions feature different limiting behaviors at low and high r, summarized in Table 4. These limits, and the accompanying scaling of peak intensities as r increases, emphasize different features of interatomic order. Preferred usage has developed in various research communities according to some of these distinguishing behaviors.

The weighted radial distribution function, $R(r)$, shown at the top of Figs. 3 and 4, and the radial distribution function, RDF(r), find limited use because they rapidly increase towards infinity with increasing r. Thus it is a challenge to

visually inspect the local correlations on the same scale as the mid-to-long-range correlations.

In materials which lack long-range order, few important structural details exist at high r, and several functions are commonly used. We introduced definitions of pair distribution functions based on the density function, $\rho(r)$, which is a straightforward quantity to calculate from atomistic simulations and models. It is shown in the panels second from the top in Figs. 3 and 4. $g(r)$ is simply the number density divided by the average density, and it has found wide adoption in the amorphous and liquids community (Benmore, 2012). The limits of $g(r)$ are absolutely defined to be zero prior to the first pair correlation and 1 at high r, which in practice can aid with data reduction and normalization procedures. An example of $g(r)$ can be seen in the third panel from the top of Figs. 3 and 4. $G_K(r)$, shown fourth from the top in Figs. 3 and 4, shares many qualitative features with $g(r)$, and while they appear nearly identical at first glance, they feature different units, limits and scaling behavior (refer to Appendix A, Figs. 3 and 4, and Table 4 for details).

To resolve certain details of local structure, researchers occasionally find it useful to preferentially weight a structural refinement towards features at low r, ignoring or downplaying longer-range features. While there is nothing implicitly wrong with this approach, this is a decision best applied at the time of modeling and stated clearly in analysis discussions. The issue with functions that inherently carry their own r-dependent weighting [such as $G_K(r)$ damping as $1/r$] is that they require additional r-dependent normalization of residuals to uniformly treat misfit at different length scales. It can be argued that representations of the measured data should not themselves contain such an r-dependent feature bias.

Table 4
Limits of real-space functions.

For materials with long-range order (*i.e.* crystalline), the high-r behavior is often obscured by the peaks in the distribution function. Like radians, atoms is normally not listed as a unit, but this table explicitly mentions it for added clarity.

Function	Low-r behavior	High-r behavior	Units
$R(r)$	0	$4\pi\rho_0 r^2$	Atom Å^{-1}
$\rho(r)$	0	ρ_0	Atom Å^{-3}
$g(r)$	0	1	Unitless
$G_K(r)$	$-\langle b_{coh}\rangle^2$	0	Barn
$G(r)$	$-4\pi\rho_0 r$	0	Atom Å^{-2}
$D(r)$	$-4\pi\rho_0\langle b_{coh}\rangle^2 r$	0	Barn atom Å^{-2}
$T(r)$	0	$4\pi\rho_0\langle b_{coh}\rangle^2 r$	Barn atom Å^{-2}
$N(r)$	0	$\propto r^3$	Atom

The reduced pair distribution function, $G(r)$ (second from the bottom in Figs. 3 and 4), is the most prevalent formalism used in the study of disordered crystalline materials and nanocrystalline materials and is the version of data compatible with the popular real-space PDF modeling program *PDFgui* (Farrow *et al.*, 2016). $G(r)$ is also sometimes used in glass or molecular liquid studies, particularly when longer-range ordering is present. Arguably, the most advantageous feature of the $G(r)$ formalism is that the amplitude of the oscillations is independent of R value. This means that the nature of a material's structural coherence can be readily interpreted via visual inspection of $G(r)$. It also means that residual differ-

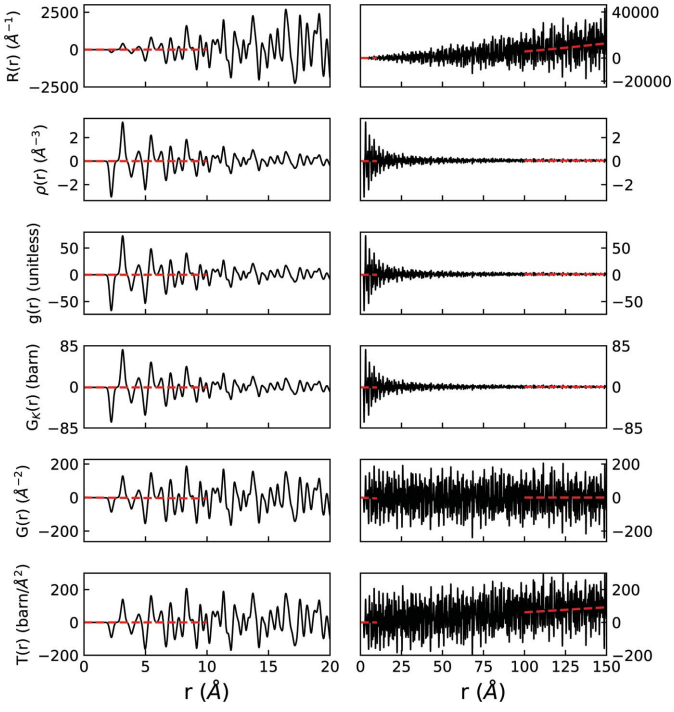


Figure 3
Comparison of long-range behavior of different real-space functions for MnO. The dashed lines highlight the asymptotic behavior. Since $D(r) = \langle b_{coh}\rangle^2 G(r)$, it is not shown.

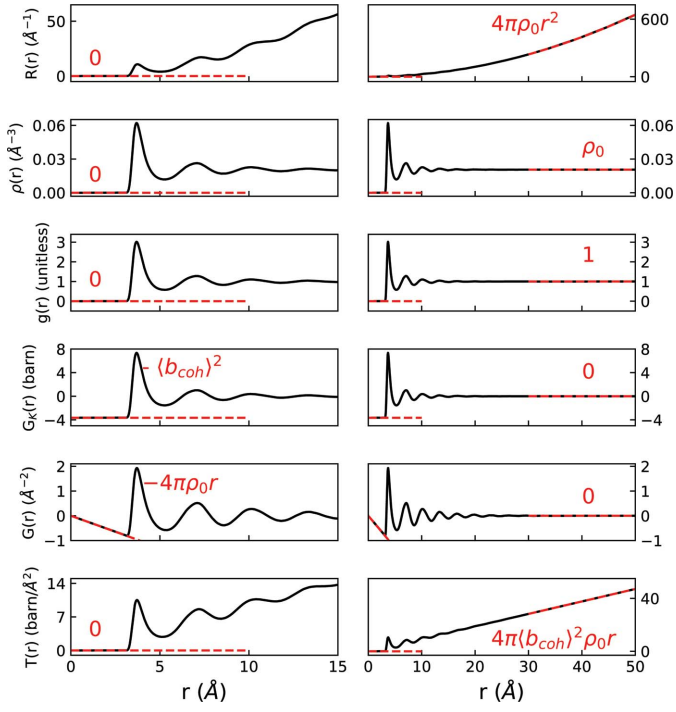


Figure 4
Comparison of long-range behavior of different real-space functions for Ar. The dashed lines highlight the asymptotic behavior. Since $D(r) = \langle b_{coh}\rangle^2 G(r)$, it is not shown.

ences between models and data are equally weighted at all R values.

It is sometimes asserted that $G(r)$ is the most directly calculable function from experimental data (Egami & Billinge, 2012), as it is the direct Fourier transform of $S(Q)$ and does not require any assumptions of number density or average scattering power. In practice, the data reduction procedures employed to generate $G(r)$ typically involve a number of optimization steps, which effectively estimate various sample-dependent corrections either analytically (Peterson *et al.*, 2000; Jeong *et al.*, 2001; Qiu, Božin *et al.*, 2004) or via *ad hoc* methods (Neuefeind *et al.*, 2012; Juhás *et al.*, 2013).

$D(r)$ shares many of the same features as $G(r)$, as it is equivalent to $\langle b_{\rm coh}\rangle^2 G(r)$. Because of this subtle difference, there has been some confusion in the community about when to use $D(r)$ compared with $G(r)$ in different analysis methods. In practice, employing either $D(r)$ or $G(r)$ when using small-box modeling approaches (where a scale parameter for the data set can be freely refined) will produce identical model results. However, they cannot be used interchangeably in those methods that rely on absolutely normalized data, such as RMC-based modeling. This can be particularly tricky in those cases where $\langle b_{\rm coh}\rangle^2$ is near to one, as the results of a large-box modeling approach may appear to be converging, but the results will be incorrect. Researchers are advised to carefully verify what form of the PDF they are employing, especially when using data from a new beamline (where data reduction protocols may differ) or employing a new form of analysis.

The neutron glass community tends to favor the $T(r)$ formalism shown at the bottom of Figs. 3 and 4 (Hannon *et al.*, 1990; Ellison *et al.*, 1993). $T(r)$ [and $G(r)$] scale relative to the number density as a function of r, as opposed to functions such as $g(r)$ or $R(r)$ (see Appendix *A*4). In the harmonic approximation of atomic motion, peaks are broadened symmetrically in $T(r)$ [and $G(r)$] by thermal motions (Wright *et al.*, 1989), which is cited as a considerable advantage in differentiating between static and thermal disorder (Hannon *et al.*, 1990).

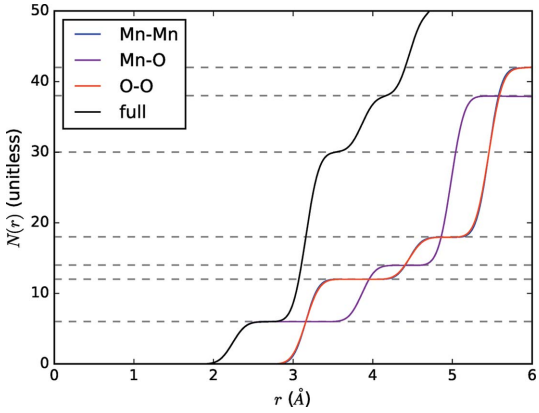


Figure 5
Accumulated correlation number for MnO plotted with its partials. The partial and full coordination numbers are marked with dashed lines. Note that the partial coordination numbers for Mn–Mn and O–O are identical owing to symmetry, but the isotropic displacement parameters, $U_{\rm iso}$, are different, which is visible near the plateaus in the Mn–Mn and O–O partials.

Table 5
Conversions to and from $\rho(r)$.

A complete list of conversion factors is given in Appendix *A*.

Formalism in terms of $\rho(r)$	$\rho(r)$ in terms of formalism
$R(r) = 4\pi r^2\rho(r)$	$\rho(r) = \dfrac{R(r)}{4\pi r^2}$
$g(r) = \rho(r)/\rho_0$	$\rho(r) = \rho_0 g(r)$
$G_{\rm K}(r) = \langle b_{\rm coh}\rangle^2\left[\dfrac{\rho(r)}{\rho_0} - 1\right]$	$\rho(r) = \rho_0\left[\dfrac{G_{\rm K}(r)}{\langle b_{\rm coh}\rangle^2} + 1\right]$
$G(r) = 4\pi r\left[\rho(r) - \rho_0\right]$	$\rho(r) = \dfrac{G(r)}{4\pi r} + \rho_0$
$D(r) = 4\pi r\langle b_{\rm coh}\rangle^2[\rho(r) - \rho_0]$	$\rho(r) = \dfrac{D(r)}{4\pi r\langle b_{\rm coh}\rangle^2} + \rho_0$
$T(r) = \langle b_{\rm coh}\rangle^2 4\pi r\rho(r)$	$\rho(r) = \dfrac{T(r)}{4\pi r\langle b_{\rm coh}\rangle^2}$

This symmetry is noted by the glass community to lead to more accurate fitting for extracting coordination numbers (Benmore, 2012). However, $T(r)$'s overall r-dependent scaling means it is not practical for viewing wide ranges in real space, and thus it is not commonly used in studies of nanocrystalline or disordered crystalline materials.

Arguably, the density function is one of the most straightforward functions to calculate from an atomic model. We here present the conversions from this common quantity, $\rho(r)$, and the other six common all-atom PDFs we have discussed herein. Table 5 provides conversions between some of the real space functions. The full list of conversions originating from each function in turn can be found in Appendix *A*.

Finally, PDF data are often interpreted using the coordination number, $N(r)$. This is the number of atoms between $r_{\rm min}$ and $r_{\rm max}$. The coordination number is described mathematically in terms of the partial $g_{\mu\nu}(r)$ functions (Soper, 2010):

$$N_{\nu\mu}(r_{\rm min}, r_{\rm max}) = 4\pi\rho_0 c_\mu \int_{r_{\rm min}}^{r_{\rm max}} r^2 g_{\nu\mu}(r)\,{\rm d}r, \qquad (30)$$

which explicitly does not include the scattering lengths. A closely related function is the accumulation of this summed over all atoms:

$$N(r) = 4\pi\rho_0 \sum_{\nu\mu} c_\mu \int_0^r (r')^2 g_{\nu\mu}(r')\,{\rm d}r'. \qquad (31)$$

This quantity is shown in Fig. 5. Note that coordination number is not weighted by scattering length, such that it cannot be easily transformed to other real-space functions without employing partial PDF functions.

5. Summary

This paper has provided a resource to understand the relationships of and convert between eight real-space pair distribution functions commonly found in the scientific literature.

This purely instructional work offers direct mathematical conversions, graphical representations and a practical discussion of function characteristics, meant as an updated step-by-step reference for new practitioners and those seeking to compare their results with those of other researchers. While the appearance and weighting of these representations can vary, often to emphasize certain features of interest, the inherent structural information must be the same among these different forms. Although we respect the decisions of individual researchers to use whichever PDF formalism they find most appropriate, we encourage convergence and standardization where possible. On the basis of the comparison and discussion presented herein, we endorse the use of two primary formalisms: $G(r)$ and $g(r)$. The reduced pair distribution function, $G(r)$, is recommended because of its uniform weighting at all R values. It has been broadly employed by the disordered crystalline and nanocrystalline communities. In contrast, the pair distribution function, $g(r)$, is recommended as it is conveniently bounded at 0 and 1 (simplifying normalization procedures), features symmetric peaks and emphasizes low-r features. Thus it has been widely adopted by the liquids, amorphous and glass communities.

We also encourage that, when reporting a PDF, authors overtly define which real-space distribution functions they present. Reciprocal-space functions should be described in terms of $S(Q)$ (which has an agreed upon and consistent definition across fields). For example, 'we show the structure function, $F(Q) = Q[S(Q) - 1]$'. Similarly, real-space functions should be described in terms of $\rho(r)$. For example, 'we fit the pair distribution function, $G(r) = 4\pi[\rho(r) - \rho_0]$'. It is our hope that the derivations, tables and figures presented in this work help serve as a reference tool for researchers to easily navigate the PDF landscape and guide towards a convergence of total scattering data formalisms.

APPENDIX *A*
Conversions between real-space functions

This appendix will provide transformations between the various real-space functions. Several notes are needed before the listing of equations. First, we will assume a shape function of $\gamma_0(r) = 1$. This is true for sufficiently bulk material measurements (*i.e.* material structures exhibit translational periodicity and thus do not have finite size, shape or morphology effects contributing to the experimental PDF). Second, the correlation number, $N(r)$, will not be mentioned since it can only be correctly calculated from the partial distributions. Note that the rise of $N(r)$ at the location of an isolated peak corresponds to the coordinate number of that pairwise peak. Third, the normalization of $g(r)$ provided in the main text is applicable here.

The original purpose of this appendix was to provide simple recipes for fellow researchers to transform their data or model calculations between the real-space forms. An additional benefit of generating this appendix, as well as the limits in Tables 3 and 4, was further validation that the conversions between conventions and the presentation of the conventions themselves are correct.

A1. Conversions from the radial distribution function, $R(r)$

$$\begin{aligned}
\rho(r) &= \frac{R(r)}{4\pi r^2} \\
G(r) &= \frac{R(r)}{r} - 4\pi r\rho_0 \\
G_{\rm K}(r) &= \langle b_{\rm coh}\rangle^2\left[\frac{R(r)}{4\pi r^2\rho_0} - 1\right] \\
g(r) &= \frac{R(r)}{4\pi r^2\rho_0} \\
D(r) &= \langle b_{\rm coh}\rangle^2\left[\frac{R(r)}{r} - 4\pi r\rho_0\right] \\
T(r) &= \frac{\langle b_{\rm coh}\rangle^2}{r}R(r)
\end{aligned} \tag{32}$$

A2. Conversions from the number density, $\rho(r)$

$$\begin{aligned}
R(r) &= 4\pi r^2\rho(r) \\
G(r) &= 4\pi r\big[\rho(r) - \rho_0\big] \\
G_{\rm K}(r) &= \langle b_{\rm coh}\rangle^2\left[\frac{\rho(r)}{\rho_0} - 1\right] \\
g(r) &= \rho(r)/\rho_0 \\
D(r) &= 4\pi r\langle b_{\rm coh}\rangle^2\big[\rho(r) - \rho_0\big] \\
T(r) &= \langle b_{\rm coh}\rangle^2 4\pi r\rho(r)
\end{aligned} \tag{33}$$

A3. Conversions from $g(r)$

$$\begin{aligned}
R(r) &= 4\pi r^2\rho_0 g(r) \\
\rho(r) &= \rho_0 g(r) \\
G(r) &= 4\pi r\rho_0[g(r) - 1] \\
G_{\rm K}(r) &= \langle b_{\rm coh}\rangle^2[g(r) - 1] \\
D(r) &= 4\pi r\rho_0\langle b_{\rm coh}\rangle^2[g(r) - 1] \\
T(r) &= 4\pi r\rho_0\langle b_{\rm coh}\rangle^2 g(r)
\end{aligned} \tag{34}$$

A4. Conversions from the total radial distribution function, $G_{\rm K}(r)$

$$\begin{aligned}
R(r) &= 4\pi r^2\rho_0\left[\frac{G_{\rm K}(r)}{\langle b_{\rm coh}\rangle^2} + 1\right] \\
\rho(r) &= \rho_0\left[\frac{G_{\rm K}(r)}{\langle b_{\rm coh}\rangle^2} + 1\right] \\
G(r) &= \frac{4\pi r\rho_0}{\langle b_{\rm coh}\rangle^2}G_{\rm K}(r) \\
g(r) &= \frac{G_{\rm K}(r)}{\langle b_{\rm coh}\rangle^2} + 1 \\
D(r) &= 4\pi r\rho_0 G_{\rm K}(r) \\
T(r) &= 4\pi r\rho_0\big[G_{\rm K}(r) + \langle b_{\rm coh}\rangle^2\big]
\end{aligned} \tag{35}$$

A5. Conversions from the reduced pair distribution function, $G(r)$

$$\begin{aligned} R(r) &= rG(r) + 4\pi r^2\rho_0 \\ \rho(r) &= \frac{G(r)}{4\pi r} + \rho_0 \\ G_{\rm K}(r) &= \frac{\langle b_{\rm coh}\rangle^2}{4\pi r\rho_0} G(r) \\ g(r) &= \frac{G(r)}{4\pi r\rho_0} + 1 \\ D(r) &= \langle b_{\rm coh}\rangle^2 G(r) \\ T(r) &= \langle b_{\rm coh}\rangle^2\left[G(r) + 4\pi r^2\rho_0\right] \end{aligned} \tag{36}$$

A6. Conversions from the differential correlation function, $D(r)$

$$\begin{aligned} R(r) &= \frac{r}{\langle b_{\rm coh}\rangle^2} D(r) + 4\pi r^2\rho_0 \\ \rho(r) &= \frac{D(r)}{4\pi r\langle b_{\rm coh}\rangle^2} + \rho_0 \\ G(r) &= \frac{D(r)}{\langle b_{\rm coh}\rangle^2} \\ G_{\rm K}(r) &= \frac{D(r)}{4\pi r\rho_0} \\ g(r) &= \frac{D(r)}{4\pi r\rho_0\langle b_{\rm coh}\rangle^2} + 1 \\ T(r) &= D(r) + \frac{4\pi r^2\rho_0}{\langle b_{\rm coh}\rangle^2} \end{aligned} \tag{37}$$

A7. Conversions from the total correlation function, $T(r)$

$$\begin{aligned} R(r) &= \frac{rT(r)}{\langle b_{\rm coh}\rangle^2} \\ \rho(r) &= \frac{T(r)}{4\pi r\langle b_{\rm coh}\rangle^2} \\ G(r) &= \frac{T(r)}{\langle b_{\rm coh}\rangle^2} - 4\pi r\rho_0 \\ G_{\rm K}(r) &= \frac{T(r)}{4\pi r\rho_0} - \langle b_{\rm coh}\rangle^2 \\ g(r) &= \frac{T(r)}{4\pi r\rho_0\langle b_{\rm coh}\rangle^2} \\ D(r) &= T(r) - \frac{4\pi r^2\rho_0}{\langle b_{\rm coh}\rangle^2} \end{aligned} \tag{38}$$

APPENDIX *B*
Details of liquid argon molecular dynamics simulations

For all Ar simulations, the system consisted of a cubic simulation cell with 50 000 atoms and a velocity-Verlet-like (Verlet, 1967) time integrator was used with a 1 fs timestep. A Lennard–Jones pair potential with a 15 Å cutoff was used with a tail correction applied, with $\epsilon = 0.238067$ kcal mol^{-1} and $\sigma =$ 3.405 Å (Yarnell *et al.*, 1973). Initially, the atoms were randomly placed in the box and then force minimized with a force tolerance of 10^{-8} kcal (mol Å)$^{-1}$. To produce the real- and reciprocal-space patterns, canonical (NVT) ensemble simulations were carried out with T = 86.56 K and ρ_0 = 0.02138 atoms Å^{-3} to reproduce previous results (Yarnell *et al.*, 1973). A Nosé–Hoover-style thermostat (Nosé, 1984; Hoover, 1985) was used to keep constant temperature with a relaxation time of 0.1 ps. For the isothermal compressibility calculations, isobaric isothermal (NPT) simulations were carried out at the same temperature as the NVT simulation but with different pressures. A Nosé–Hoover-style barostat (Nosé, 1984; Hoover, 1985) was used to keep constant pressure with a relaxation time of 1 ps. The PDF was then calculated over a real-space range $0.075 \le r \le 100$ Å with a bin width of δr = 0.05 Å. The reciprocal-space data were generated through inverse Fourier transform. The reciprocal-space data are scaled such that forward Fourier transforms recreate the original PDFs precisely.

APPENDIX *C*
From experimentally observed scattering intensities to differential cross sections

The authors mostly ignored the steps involved in reducing data from measured scattering intensities (Section 3) to reciprocal-space functions in the main text. This appendix will shed light on some of the corrections necessary to transform measured intensities to a single scattering differential cross section (DCS), $\mathrm{d}\sigma^{\rm s}/\mathrm{d}\Omega$, which can then be used in equation (4). For Rietveld analysis, this level of rigor for corrections is not usually needed for most samples as these effects are frequently addressed as part of the measurement's 'background intensity', which is normally fitted with a polynomial to reduce its influence on the profile refinement. However, for real-space total scattering measurements, these 'minor' effects can produce a signal comparable to the reciprocal-space features of local disorder and therefore must be accounted for.

Repeating equation (4),

$$S(Q) = \frac{I(Q)}{N\langle b_{\rm coh}\rangle^2} - \frac{\langle b_{\rm tot}^2\rangle - \langle b_{\rm coh}\rangle^2}{\langle b_{\rm coh}\rangle^2}, \tag{39}$$

where $I(Q)$ is the scattering from the sample (using the same convention as Rietveld). Its relation to the differential cross section (DCS) is simply (Egami & Billinge, 2012)

$$I(Q) = N\frac{\mathrm{d}\sigma^{\rm s}}{\mathrm{d}\Omega}(Q), \tag{40}$$

where $N = \rho_{0,\rm eff}V$ is the number of atoms illuminated, V is the illuminated volume and $\rho_{0,\rm eff}$ is the effective number density as described in Appendix *E*. Because corrections are easier to explain in terms of the DCS, we rewrite equation (4) as

$$S(Q) = \frac{1}{\langle b_{\rm coh}\rangle^2}\left[\frac{\mathrm{d}\sigma^{\rm s}}{\mathrm{d}\Omega}(Q) - \langle b_{\rm tot}^2\rangle - \langle b_{\rm coh}\rangle^2\right]. \tag{41}$$

The DCS can only be found by fully correcting the measured intensity. Proper treatment of total scattering data requires

correcting for effects such as attenuation through the sample material and its environment, subtraction of multiple scattering events, and recoil or inelastic effects. That being said, the notations for the various corrections do not have a standard set of definitions. The associated measurements for these corrections (*e.g.* empty container, empty sample environment, normalization) are no longer as unusual to perform as they used to be. Table 6 describes the standard set of measurements required and associates subscripts to refer to them. Additionally, Paalman & Pings (1962) developed a detailed notation for the attenuation correction terms which will be used here. Table 7 describes the notation for the correction terms applied to the experimental measurements.

For the measurement of the sample and container combined there are four terms of scattering and attenuation (also called absorption) in the experiment. For example, the scattering that occurs in only the sample but is attenuated by both the sample and container component is denoted as $A_{\rm s,sc}$. Similarly, the multiple scattering is denoted as M_μ for the μth measurement, and thus the sample and container are denoted $M_{\rm sc}$. Finally, the inelastic recoil correction is denoted as $P_\nu^{\rm ie}$ for the νth species; thus the container inelastic correction would be $P_{\rm c}^{\rm ie}$.

Having the various functions and subscript definitions in hand, the description of what is actually experimentally measured, $I_i^{\rm E}$, can be introduced. The experimentally measured intensities are summarized in Table 8.

These experimentally measured intensities are represented by the functional forms below with the necessary correction terms included from Table 7:

$$\begin{aligned}
I_{\rm sca}^{\rm E}(Q,\omega) &= \Phi\left[A_{\rm s,sca}N_{\rm s}\frac{{\rm d}^2\sigma^{\rm s}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega) + A_{\rm c,sca}N_{\rm c}\frac{{\rm d}^2\sigma^{\rm c}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\right.\\
&\quad \left. + A_{\rm a,sca}N_{\rm a}\frac{{\rm d}^2\sigma^{\rm a}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega) + M_{\rm sca} + \frac{{\rm d}^2\sigma^{\rm e}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\right],\\
I_{\rm ca}^{\rm E}(Q,\omega) &= \Phi\left[A_{\rm c,ca}N_{\rm c}\frac{{\rm d}^2\sigma^{\rm c}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega) + A_{\rm a,ca}N_{\rm a}\frac{{\rm d}^2\sigma^{\rm a}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\right.\\
&\quad \left. + M_{\rm ca} + \frac{{\rm d}^2\sigma^{\rm e}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\right],\\
I_{\rm a}^{\rm E}(Q,\omega) &= \Phi\left[A_{\rm a,a}N_{\rm a}\frac{{\rm d}^2\sigma^{\rm a}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega) + M_{\rm a} + \frac{{\rm d}^2\sigma^{\rm e}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\right],\\
I_{\rm e}^{\rm E}(Q,\omega) &= \Phi\frac{{\rm d}^2\sigma^{\rm e}}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega).
\end{aligned} \tag{42}$$

Note that the double-differential cross-section term appears in these equations, with $\hbar\omega$ being the energy loss. For the ultimate goal of producing a suitable reciprocal-space function, one must integrate over ω to obtain the energy-integrated DCS. If all scattering were purely elastic from the sample, we would simply have

$$\frac{{\rm d}\sigma^i}{{\rm d}\Omega}(Q) = \int \frac{{\rm d}^2\sigma^i}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\,{\rm d}\omega. \tag{43}$$

Yet, there are undoubtedly inelastic scattering events that occur and necessary terms must be included in the above equation for the specific experimental technique. The equation for an experimental integration over ω is given as

$$\frac{{\rm d}\sigma^i}{{\rm d}\Omega}(Q) + P_i^{\rm ie} = \int \beta(k_{\rm i},k_{\rm f})\frac{{\rm d}^2\sigma^i}{{\rm d}\Omega\,{\rm d}\omega}(Q,\omega)\,{\rm d}\omega, \tag{44}$$

where $\beta(k_{\rm i},k_{\rm f})$ accounts generally for terms such as detector efficiency, incident spectrum and other terms needed based on the nature of the experimental method (*i.e.* constant wavelength or time of flight). $P_i^{\rm ie}$ is the inelastic correction term that occurs due to both the recoil of the atomic nuclei from the neutron collision and, for time of flight, the fact that the measurement is carried out at a fixed angle and not fixed momentum transfer, Q (Soper, 2009). For the functional forms of $\beta(k_{\rm i},k_{\rm f})$ for constant wavelength or time of flight, we refer the reader to reports in the series by Powles (1973, 1978*a*,*b*) or the *GUDRUN* manual (Soper, 2010).

Real-space techniques are often colloquially referred to as 'total scattering' because they use all of the reciprocal-space information rather than just the Bragg peaks. As mentioned in the main text, this term has a second meaning when it is applied to describe integration over energy transfer. This is not the same as the elastic scattering, which can only be

Table 6
Description of subscripts, i, for measured intensities, $I_i(Q)$, and differential cross sections, ${\rm d}\sigma^i/{\rm d}\Omega$.

Subscript	Term
s	Sample
c	Container
a	Apparatus
sc	Sample + container
sca	Sample + container + apparatus
e	Background / empty diffractometer
n	Normalization

Table 7
Correction terms for the different types of experimental measurement setups, including absorption, multiple scattering and inelastic recoil corrections.

The μ and ν terms are defined in Table 6.

Correction term	Definition
$A_{\mu,\nu}$	Absorption correction factor for scattering in μ absorbed by ν
M_ν	Multiple scattering correction factor for ν
$P_\nu^{\rm ie}$	Inelastic recoil correction factor for ν

Table 8
The different experimentally measured intensities required for a total scattering experiment.

Experimental term	Definition
$I_{\rm sca}^{\rm E}$	Experimentally measured intensity of sca
$I_{\rm ca}^{\rm E}$	Experimentally measured intensity of ca
$I_{\rm a}^{\rm E}$	Experimentally measured intensity of a
$I_{\rm e}^{\rm E}$	Experimentally measured intensity of e
Φ	Normalization term to account for detector efficiency, solid angle coverage *etc.*

properly measured when the incident and final energy of the probe are known. Integrating the energy yields the following set of equations:

$$\begin{aligned}
I_{\rm s}^{\rm E}(Q) &= \Phi\left\{A_{\rm s,sca}N_{\rm s}\left[\frac{{\rm d}\sigma^{\rm s}}{{\rm d}\Omega}(Q)+P_{\rm s}^{\rm ie}\right]+A_{\rm c,sca}N_{\rm c}\left[\frac{{\rm d}\sigma^{\rm c}}{{\rm d}\Omega}(Q)+P_{\rm c}^{\rm ie}\right]\right.\\
&\quad\left.+A_{\rm a,sca}N_{\rm a}\left[\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega}(Q)+P_{\rm a}^{\rm ie}\right]+M_{\rm sca}+\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}(Q)\right\},\\
I_{\rm c}^{\rm E}(Q) &= \Phi\left\{A_{\rm c,ca}N_{\rm c}\left[\frac{{\rm d}\sigma^{\rm c}}{{\rm d}\Omega}(Q)+P_{\rm c}^{\rm ie}\right]+A_{\rm a,ca}N_{\rm a}\left[\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega}(Q)+P_{\rm a}^{\rm ie}\right]\right.\\
&\quad\left.+M_{\rm ca}+\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}(Q)\right\},\\
I_{\rm a}^{\rm E}(Q) &= \Phi\left\{A_{\rm a,a}N_{\rm a}\left[\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega}(Q)+P_{\rm a}^{\rm ie}\right]+M_{\rm a}+\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}(Q)\right\},\\
I_{\rm e}^{\rm E}(Q) &= \Phi\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}(Q).
\end{aligned} \tag{45}$$

Solving each of the equations above for the respective DCS [and simplifying $({\rm d}\sigma^{i}/{\rm d}\Omega)(Q)$ to just ${\rm d}\sigma^{i}/{\rm d}\Omega$] provides

$$\begin{aligned}
\frac{{\rm d}\sigma^{\rm s}}{{\rm d}\Omega} &= \frac{1}{A_{\rm s,sca}N_{\rm s}}\left[\frac{1}{\Phi}\left(I_{\rm s}^{\rm E}-\Phi\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}\right)-M_{\rm sca}\right.\\
&\quad\left.-\left(A_{\rm c,sca}N_{\rm c}\frac{{\rm d}\sigma^{\rm c}}{{\rm d}\Omega}+P_{\rm c}^{\rm ie}\right)-\left(A_{\rm a,sca}N_{\rm a}\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega}+P_{\rm a}^{\rm ie}\right)\right]-P_{\rm s}^{\rm ie},\\
\frac{{\rm d}\sigma^{\rm c}}{{\rm d}\Omega} &= \frac{1}{A_{\rm c,ca}N_{\rm c}}\left\{\frac{1}{\Phi}\left(I_{\rm c}^{\rm E}-\Phi\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}\right)-M_{\rm ca}\right.\\
&\quad\left.-\left[A_{\rm a,ca}N_{\rm a}\left(\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega}+P_{\rm a}^{\rm ie}\right)\right]\right\}-P_{\rm c}^{\rm ie},\\
\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega} &= \frac{1}{A_{\rm a,a}N_{\rm a}}\left[\frac{1}{\Phi}\left(I_{\rm a}^{\rm E}-\Phi\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}\right)-M_{\rm a}\right]-P_{\rm a}^{\rm ie},\\
\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega} &= \frac{1}{\Phi}I_{\rm e}^{\rm E}.
\end{aligned} \tag{46}$$

Reducing the equations into terms that only involve the corrections and the measured intensities, the sample DCS is

$$\begin{aligned}
\frac{{\rm d}\sigma^{\rm s}}{{\rm d}\Omega} &= \frac{1}{A_{\rm s,sca}N_{\rm s}}\left(\frac{1}{\Phi}(I_{\rm s}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm sca}\right.\\
&\quad-\frac{A_{\rm c,sca}}{A_{\rm c,ca}}\left\{\frac{1}{\Phi}(I_{\rm c}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm ca}-\frac{A_{\rm a,ca}}{A_{\rm a,a}}\left[\frac{1}{\Phi}(I_{\rm a}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm a}\right]\right\}\\
&\quad\left.-\frac{A_{\rm a,sca}}{A_{\rm a,a}}\left[\frac{1}{\Phi}(I_{\rm a}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm a}\right]\right)-P_{\rm s}^{\rm ie},
\end{aligned} \tag{47}$$

the container DCS is

$$\begin{aligned}
\frac{{\rm d}\sigma^{\rm c}}{{\rm d}\Omega} &= \frac{1}{A_{\rm c,ca}N_{\rm c}}\left\{\frac{1}{\Phi}(I_{\rm c}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm ca}\right.\\
&\quad\left.-\frac{A_{\rm a,ca}}{A_{\rm a,a}}\left[\frac{1}{\Phi}(I_{\rm a}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm a}\right]\right\}-P_{\rm c}^{\rm ie}
\end{aligned} \tag{48}$$

and the sample environment apparatus DCS is

$$\frac{{\rm d}\sigma^{\rm a}}{{\rm d}\Omega} = \frac{1}{A_{\rm a,a}N_{\rm a}}\left[\frac{1}{\Phi}(I_{\rm a}^{\rm E}-I_{\rm e}^{\rm E})-M_{\rm a}\right]-P_{\rm a}^{\rm ie}. \tag{49}$$

Note that here the normalization term, Φ, has not been made explicit. Thus, an additional experimental measurement is needed to characterize the normalization term. Using our definitions in equations (45), we can define this normalization measurement as

$$I_{\rm n}^{\rm E} = \Phi\left[A_{\rm n,n}N_{\rm n}\left(\frac{{\rm d}\sigma^{\rm n}}{{\rm d}\Omega}+P_{\rm n}^{\rm ie}\right)+M_{\rm n}+\frac{{\rm d}\sigma^{\rm e}}{{\rm d}\Omega}\right]. \tag{50}$$

The normalization measurement performed will differ according to the probe used, X-rays or neutrons. Yet, both serve a similar purpose in that they factor out the dependence of detector efficiencies, solid angle coverage *etc.* As a specific example, vanadium is often utilized for neutron scattering, mainly because it has a small coherent scattering length, implying that the distinct scattering signal is small compared with the self-scattering and thus it exhibits a relatively smooth diffraction pattern. Since self-scattering does not vary much with Q or 2θ, the differential cross section is essentially ${\rm d}\sigma^{\rm v}/{\rm d}\Omega \simeq \sigma_{\rm tot}^{\rm v}/4\pi = b_{\rm tot,v}^2$, where $b_{\rm tot,v}$ is the total scattering length of vanadium from equation (53). For more detail on reasons to use vanadium, we refer the reader to Section 3.8.1 of the *GUDRUN* manual (Soper, 2010) and for more information regarding the details of the normalization, we refer the reader to both Section A5 of the text by Egami & Billinge (2012) and Section 3.8 of the *GUDRUN* manual (Soper, 2010). By substituting $\langle b_{\rm tot,v}^2\rangle$ into equation (50) for ${\rm d}\sigma^{\rm n}/{\rm d}\Omega$ and the $I_{\rm e}^{\rm E}$ term from equation (45), we can derive the following:

$$I_{\rm n}^{\rm E} = \Phi\left[A_{\rm n,n}N_{\rm n}(\langle b_{\rm tot,v}^2\rangle+P_{\rm n}^{\rm ie})+M_{\rm n}\right]+I_{\rm e}^{\rm E}. \tag{51}$$

Solving for Φ gives

$$\Phi = \frac{I_{\rm n}^{\rm E}-I_{\rm e}^{\rm E}}{[A_{\rm n,n}N_{\rm n}(\langle b_{\rm tot,v}^2\rangle+P_{\rm n}^{\rm ie})+M_{\rm n}]}, \tag{52}$$

which defines the final Φ term needed in equations (47), (48) and (49).

Solving equation (47) from the experimentally measured intensities then allows one to arrive at the differential cross section, also commonly defined as $I(Q)/N$ in the total scattering community. This is the same as that defined in equation (1) and throughout the rest of this work. For further details on the data reduction described briefly herein and all necessary corrections, the reader is referred to more complete discussions in the literature (Howe *et al.*, 1989; Hannon *et al.*, 1990; Soper, 2010; Egami & Billinge, 2012). There are also other excellent references that deviate in detail from the reduction described herein, specifically in the steps where corrections are applied. (Windsor, 1981) Naturally, refinement of the methodology has occurred with time.

The derivations herein integrated the scattering intensities over all measurement energies, where inelastic and multiple scattering effects are handled through data corrections. Several authors have noted the implications of this simplification (Page *et al.*, 2011). Dynamic PDF has emerged as an extension of total scattering, making explicit use of the energy

Table 9
Table of scattering lengths for the example materials taken from Sears (1992).

		b_{coh} (fm)		
Atom	b_{tot} (fm)	Real	Imaginary	Length
Ar	2.3313	1.909	0.0	1.909
Mn	4.1363	−3.73	0.0	3.73
O	5.8032	5.803	0.0	5.803

Table 10
Calculated values for the normalized Laue term shown with fixed precision.

Note that the normalized Laue term is not necessarily zero for monatomic materials.

Material	$\langle b_{tot}^2 \rangle$ (fm^2)	$\langle b_{coh} \rangle^2$ (fm^2)	L
Ar	5.435	3.644	0.491
MnO	25.393	1.074	22.636

dependence (Egami & Billinge, 2012). This falls outside the scope of the present work.

APPENDIX *D*
Calculating the normalized Laue term

There is often confusion when determining the normalized Laue term of equation (5). Lovesey (1986) introduced quantities related to the total cross section of the material,

$$\sigma_{tot} = 4\pi \langle b_{tot}^2 \rangle, \tag{53}$$

and the coherent cross section,

$$\sigma_{coh} = 4\pi \langle b_{coh} \rangle^2. \tag{54}$$

Here, the subscript coh stands for coherent and tot stands for total, where $\langle b_{tot}^2 \rangle$ must be inferred from the total cross section (Sears, 1992). In the Faber–Ziman scheme, $\langle b_{coh} \rangle^2$ comes from the sum of the partials and the pupose of the $\langle b_{tot}^2 \rangle$ term is to remove the self-scattering from the differential cross section (Faber & Ziman, 1965). These are the terms that appear in the normalized Laue term with the factor of 4π canceled out. Note that $\langle \sigma_{coh} \rangle \neq 4\pi \langle b_{coh} \rangle^2$, but Lovesey's notation obscures this fact. Being more explicit, it is straightforward to calculate the terms as originally intended. First we introduce a normalized concentration

$$c_\alpha = N_\alpha / N, \tag{55}$$

where N_α is the number of atoms of type α and $N = \sum_\alpha N_\alpha$. This provides for the normalization that $\sum_\alpha c_\alpha = 1$. The two quantities needed in the normalized Laue term are simply

$$\langle b_{tot}^2 \rangle = \sum_\alpha c_\alpha b_{tot,\alpha}^2 = (1/4\pi) \sum_\alpha c_\alpha \sigma_{tot,\alpha} \tag{56}$$

and

$$\langle b_{coh} \rangle^2 = \left(\sum_\alpha c_\alpha b_{coh,\alpha} \right) \left(\sum_\alpha c_\alpha b_{coh,\alpha}^* \right). \tag{57}$$

This formulation of $\langle b_{coh} \rangle^2$ makes more explicit that the complex scattering lengths, with sign, are averaged. Sears (1992) offers a useful discussion of calculating average cross sections. Unique to neutron measurements, this can lead to materials where the normalized Laue term becomes infinite because of atoms with a negative scattering length. $Ti_{2.08}Zr$ is a classic example of $\langle b_{coh} \rangle^2 = 0$ and $\langle b_{tot}^2 \rangle \neq 0$.

An additional complication is found when $\langle b_{tot}^2 \rangle$ is calculated. Since 1 barn = 100 fm^2, people often calculate $\langle b_{tot}^2 \rangle$ by simply dividing by 4π and ignoring the units. While the factor of 100 will cancel out in the normalized Laue term, the values of the individual terms will be listed in units of 10 fm (or dekafemtometres) rather than fm. Tables 9 and 10 show calculations for the example materials.

APPENDIX *E*
Determining number density

Often when processing or analyzing data, the effective number density, or packing fraction, is treated as an adjustable parameter. Here we demonstrate how to calculate it and give guidance on reasonable limits. The effective number density affects absorption and multiple scattering corrections as well as the total number of illuminated atoms as described in Appendix *C*. Since it is difficult to directly measure the packing fraction precisely, one can see why it is often tuned during data reduction.

First we introduce the concept of effective number density,

$$\rho_{0,eff} = f \rho_0, \tag{58}$$

which is based on the crystallographic number density, ρ_0, using a packing fraction $f \in (0, 1]$. For measurements of crystalline powders, the packing fraction rarely exceeds $f = 0.5$. The number density for a crystalline sample is often phrased in terms of the number of atoms per unit cell, called the Z parameter, and the unit-cell volume. Using standard crystallographic conventions (Giacovazzo, 1992)

$$\rho_0 = \frac{Z}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}. \tag{59}$$

For liquids, ρ_0 is not normally known. The technique for establishing it is to calculate $\rho_{0,eff}$ from the mass density, ρ_m. This is done by weighing the empty sample container, adding the sample, and then weighing the filled container and measuring how full the container is. This measurement will give a mass density with the packing fraction correctly accounted for. Then the effective number density is

$$\rho_{0,eff} = f \rho_0 = \frac{\rho_m N_A}{\sum_\alpha c_\alpha m_\alpha}, \tag{60}$$

where N_A is Avogadro's number, m_α is the atomic mass, and c_α is the normalized concentration of element α or the atomic fraction as defined in Appendix *D*. This procedure is often used for crystalline samples as well. An alternative to this approximation is to measure the volume by displacement in a fluid that the material does not react with and is insoluble within. However, it is likely to be difficult to recover a sample

from this type of measurement so it is recommended that it is done after the scattering measurement is completed.

APPENDIX *F*
Partial structure functions

To use the prevalent Faber–Ziman partial structure functions, one must first define weights (Faber & Ziman, 1965; Suck *et al.*, 1993; Egami & Billinge, 2012)

$$W_{\nu\mu} = c_\nu c_\mu \frac{b_{\mathrm{coh},\nu}\, b_{\mathrm{coh},\mu}}{\langle b_{\mathrm{coh}}\rangle^2} = \frac{N_\nu N_\mu}{N^2} \frac{b_{\mathrm{coh},\nu}\, b_{\mathrm{coh},\mu}}{\langle b_{\mathrm{coh}}\rangle^2}, \tag{61}$$

which are normalized such that $\sum_{\nu\mu} W_{\nu\mu} = 1$. Note that the same symbol is used with an alternative normalization in some communities with $\sum_{\nu\mu} W_{\nu\mu} = \langle b_{\mathrm{coh}}\rangle^2$. The Faber–Ziman partial structure functions are related to the total scattering structure factor

$$S(Q) = \sum_{\nu\mu} W_{\nu\mu} A_{\nu\mu}(Q) \tag{62}$$

with $A_{\nu\mu}(Q) = S_{\nu\mu}(Q)$ in the nomenclature that Egami uses. The benefit of this normalization of $S_{\nu\mu}(Q)$ is that it has the same asymptotes as $S(Q)$ with

$$\lim_{Q\to\infty} A_{\nu\mu}(Q) = 1 \tag{63}$$

and

$$\lim_{Q\to 0} A_{\nu\mu}(Q) = \eta. \tag{64}$$

In equation (61), $\langle b_{\mathrm{coh}}\rangle^2$ is the average scattering length for the material and the numerator is the average scattering length for the atoms contributing to the partial.

Then it is straightforward to define a partial reduced pair distribution function as

$$G_{\nu\mu}(r) = \frac{2}{\pi}\int_0^\infty Q[A_{\nu\mu}(Q) - 1]\sin(Qr)\,\mathrm{d}Q, \tag{65}$$

which observes a summation rule of

$$G(r) = \sum_\nu \sum_\mu W_{\nu\mu} G_{\nu\mu}(r). \tag{66}$$

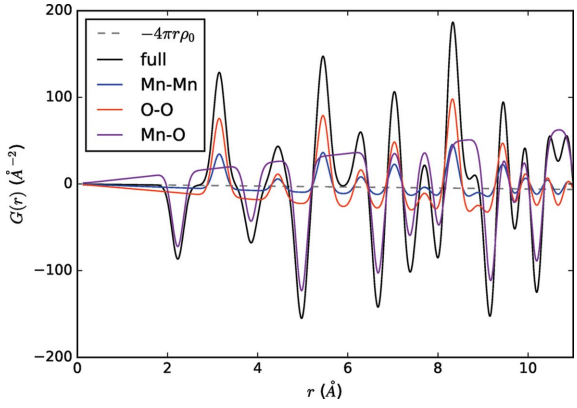


Figure 6
Partial reduced pair distribution functions for MnO.

Table 11
Weights, $W_{\nu\mu}$, for partials for the materials chosen.

The atom listed first in the chemical formula is α and the second is β.

Material	$W_{\alpha\alpha}$	$W_{\alpha\beta}$	$W_{\beta\beta}$
Ar	1.000	N/A	N/A
MnO	3.238	5.037	7.836

The other real-space correlation functions follow analogous forms.

To clarify a related point, difference correlation functions are a distinct concept. Instead of being the correlation of two atomic species, it is all correlations with an atom of type ν at the origin. In other words

$$G_\nu(r) = \sum_\mu \frac{c_\mu b_{\mathrm{coh},\mu}}{\langle b_{\mathrm{coh}}\rangle^2} G_{\nu\mu}(r), \tag{67}$$

which follows the summation rule

$$G(r) = \sum_\nu \frac{c_\nu b_{\mathrm{coh},\nu}}{\langle b_{\mathrm{coh}}\rangle^2} G_\nu(r). \tag{68}$$

A more detailed description is given by Egami & Billinge (2012, Sections 3.1 and 3.2).

For a two-atom system with atoms labeled α and β, the summation rule for the weights in equation (61) is

$$W_{\alpha\alpha} + 2W_{\alpha\beta} + W_{\beta\beta} = 1. \tag{69}$$

For a monatomic system, there are no partial functions. See Table 11 for explicit calculation of the values. Fig. 6 shows the partial $G(r)$.

Acknowledgements

Much of the diffraction group of ORNL's Neutron Scattering Division gave advice during the production of this document. The authors acknowledge and express their thanks for the consultation of Matt G. Tucker, Joerg C. Neuefeind and Andrei T. Savici.

Funding information

This material is based upon work supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract No. DE-AC05-00OR22725.

References

Ashcroft, N. & Langreth, D. (1967). *Phys. Rev.* **156**, 685.
Azaroff, L. V. (1968). *Elements of X-ray Crystallography*. New York: McGraw-Hill.
Benmore, C. J. (2012). *ISRN Mater. Sci.* **2012**, 852905.
Bhatia, A. & Thornton, D. (1970). *Phys. Rev. B*, **2**, 3004.
Billinge, S. J. L. (1992). PhD thesis, University of Pennsylvania, USA.
Billinge, S. J. (2004). *Z. Kristallogr. Cryst. Mater.* **219**, 117–121.
Billinge, S. J. L. & Egami, T. (1993). *Phys. Rev. B*, **47**(21), 14386.
Billinge, S. J. & Farrow, C. L. (2013). *J. Phys. Condens. Matter*, **25**, 454202.
Billinge, S. J. & Kanatzidis, M. G. (2004). *Chem. Commun.* **2004**, 749–760.

Billinge, S. J. L. & Levin, I. (2007). *Science*, **316**, 561–565.
Debye, P. (1913). *Ann. Phys.* **348**, 49–92.
Debye, P. (1915). *Ann. Phys.* **351**, 809–823.
Debye, P. (1930). *Phys. Z.* **31**, 419–428.
Dinnebier, R. E. & Billinge, S. J. L. (2008). *Powder Diffraction: Theory and Practice.* Cambridge: Royal Society of Chemistry Publishing.
Dunitz, J. D., Schomaker, V. & Trueblood, K. N. (1988). *J. Phys. Chem.* **92**, 856–867.
Egami, T. (2007). *Annu. Rev. Mater. Res.* **37**, 297–315.
Egami, T. & Billinge, S. J. L. (2012). *Underneath the Bragg peaks: Structural Analysis of Complex Materials.* San Diego: Pergamon.
Egelstaff, P. (1967). *Adv. Phys.* **16**, 147–169.
Egelstaff, P. (2003). *Phys. Chem. Liq.* **41**(2), 109–121.
Egelstaff, P. A. (1992). *An Introduction to the Liquid State.* Oxford, New York: Clarendon Press/Oxford University Press.
Ellison, A., Crawford, R., Montague, D., Volin, K. & Price, D. (1993). *J. Neutron Res.* **1**(4), 61–70.
Faber, T. & Ziman, J. (1965). *Philos. Mag.* **11**, 153–173.
Farrow, C. L. & Billinge, S. J. L. (2009). *Acta Cryst.* A**65**, 232–239.
Farrow, C., Juhas, P., Liu, J., Bryndin, D., Božin, E., Bloch, J., Proffen, T. & Billinge, S. (2007). *J. Phys. Condens. Matter*, **19**, 335219.
Farrow, C., Juhás, P., Liu, J., Bryndin, D., Božin, E., Bloch, J., Proffen, T. & Billinge, S. (2016). *PDFgui User Guide*, http://www.diffpy.org/doc/pdfgui/pdfgui.pdf.
Fischer, H. E., Barnes, A. C. & Salmon, P. S. (2005). *Rep. Prog. Phys.* **69**, 233.
Gereben, O., Jóvári, P., Temleitner, L. & Pusztai, L. (2007). *J. Optoelectron. Adv. Mater.* **9**, 3021.
Giacovazzo, C. (1992). *Fundamentals of Crystallography.* Chester, Oxford, New York: International Union of Crystallography/Oxford University Press.
Guinier, A. & Fournet, G. (1955). *Small-Angle Scattering of X-rays.* New York: Wiley & Sons.
Hannon, A., Howells, W. & Soper, A. (1990). *Inst. Phys. Conf. Ser.* **107**, 193–211.
Hoover, W. G. (1985). *Phys. Rev. A*, **31**, 1695–1697.
Hou, D., Zhao, C., Paterson, A. R., Li, S. & Jones, J. L. (2018). *J. Eur. Ceram. Soc.* **38**, 971–987.
Howe, M. A., McGreevy, R. L. & Howells, W. S. (1989). *J. Phys. Condens. Matter*, **1**(22), 3433.
Jeong, I.-K., Heffner, R., Graf, M. & Billinge, S. (2003). *Phys. Rev. B*, **67**, 104301.
Jeong, I.-K., Proffen, T., Mohiuddin-Jacobs, F. & Billinge, S. J. (1999). *J. Phys. Chem. A*, **103**, 921–924.
Jeong, I.-K., Thompson, J., Proffen, Th., Turner, A. M. P. & Billinge, S. J. L. (2001). *J. Appl. Cryst.* **34**, 536.
Juhás, P., Davis, T., Farrow, C. L. & Billinge, S. J. L. (2013). *J. Appl. Cryst.* **46**, 560–566.
Keen, D. A. (2001). *J. Appl. Cryst.* **34**, 172–177.
Kohara, S., Itou, M., Suzuya, K., Inamura, Y., Sakurai, Y., Ohishi, Y. & Takata, M. (2007). *J. Phys. Condens. Matter*, **19**, 506101.
Kohara, S. & Suzuya, K. (2003). *Nucl. Instrum. Methods Phys. Res. B*, **199**, 23–28.
Lovesey, S. W. (1986). *Theory of Neutron Scattering from Condensed Matter.* Oxford: Clarendon Press.
McGreevy, R. L. (1995). *Nucl. Instrum. Methods Phys. Res. A*, **354**, 1–16.
McGreevy, R. L. (2001). *J. Phys. Condens. Matter*, **13**, R877–R913.
McGreevy, R. L. & Pusztai, L. (1988). *Mol. Simul.* **1**, 359–367.
McQuarrie, D. A. (2000). *Statistical Mechanics.* Mill Valley: University Science Books.
Mildner, D. & Carpenter, J. (1984). *J. Non-Cryst. Solids*, **69**, 27–37.
Narten, A. H. (1972). *J. Chem. Phys.* **56**, 1905–1909.
Neuefeind, J. (2002). *J. Mol. Liq.* **98**, 87–95.
Neuefeind, J., Feygenson, M., Carruth, J., Hoffmann, R. & Chipley, K. K. (2012). *Nucl. Instrum. Methods Phys. Res. B*, **287**, 68–75.
Neuefeind, J. & Poulsen, H. (1995). *Phys. Scr.* **1995**(T57), 112–116.
Neuefeind, J., Zeidler, M. & Poulsen, H. (1996). *Mol. Phys.* **87**, 189–201.
Nosé, S. (1984). *J. Chem. Phys.* **81**, 511–519.
Olds, D., Saunders, C. N., Peters, M., Proffen, T., Neuefeind, J. & Page, K. (2018). *Acta Cryst.* A**74**, 293–307.
Olds, D., Wang, H.-W. & Page, K. (2015). *J. Appl. Cryst.* **48**, 1651–1659.
Paalman, H. H. & Pings, C. J. (1962). *J. Appl. Phys.* **33**, 2635–2639.
Page, K., White, C. E., Estell, E. G., Neder, R. B., Llobet, A. & Proffen, Th. (2011). *J. Appl. Cryst.* **44**, 532–539.
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.
Playford, H. Y., Owen, L. R., Levin, I. & Tucker, M. G. (2014). *Annu. Rev. Mater. Res.* **44**, 429–449.
Plimpton, S. (1995). *J. Comput. Phys.* **117**, 1–19.
Plimpton, S. J. (2018). *Large-Scale Atomistic/Molecular Massively Parallel sSimulator (LAMMPS)*, https://lammps.sandia.gov/.
Powles, J. (1973). *Mol. Phys.* **26**, 1325–1350.
Powles, J. (1978*a*). *Mol. Phys.* **36**, 1161–1180.
Powles, J. (1978*b*). *Mol. Phys.* **36**, 1181–1198.
Proffen, T. & Billinge, S. J. L. (1999). *J. Appl. Cryst.* **32**, 572–575.
Qiu, X., Božin, E. S., Juhas, P., Proffen, T. & Billinge, S. J. L. (2004). *J. Appl. Cryst.* **37**, 110–116.
Qiu, X., Thompson, J. W. & Billinge, S. J. L. (2004). *J. Appl. Cryst.* **37**, 678.
Rietveld, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
Root, J., Egelstaff, P. & Hime, A. (1986). *Chem. Phys.* **109**, 437–453.
Sasaki, S., Fujino, K. & Takéuchi, Y. (1979). *Proc. Jpn Acad. Ser. B*, **55**, 43–48.
Sears, V. F. (1992). *Neutron News*, **3**(3), 26–37.
Soper, A. K. (1989). *ATLAS Manual*. Neutron Science Division, Rutherford Appleton Laboratory, Chilton, United Kingdom.
Soper, A. K. (2009). *Mol. Phys.* **107**, 1667–1684.
Soper, A. K. (2010). *GudrunN and GudrunX: Programs For Correcting Raw Neutron and X-ray Diffraction Data To Differential Scattering Cross Section*, https://www.isis.stfc.ac.uk/OtherFiles/Disordered%20Materials/Gudrun-Manual-2017-10.pdf.
Squires, G. L. (2012). *Introduction to the Theory of Thermal Neutron Scattering.* Cambridge University Press.
Srolovitz, D., Egami, T. & Vitek, V. (1981). *Phys. Rev. B*, **24**, 6936.
Suck, J. B., Raoux, D., Chieux, P. & Riekel, C. (1993). *Proceedings of the ILL/ESRF Workshop on Methods in the Determination of Partial Structure Factors of Disordered Matter by Neutron and Anomalous X-ray Diffraction.* Singapore River Edge: World Scientific.
Thorpe, M., Chung, J., Billinge, S. & Mohiuddin-Jacobs, F. (1998). *Local Structure from Diffraction*, pp. 157–174. New York: Plenum Press.
Tucker, M., Dove, M., Goodwin, A., Keen, D. A., Playford, H. & Slawinski, W. A. (2017). *RMCProfile User Manual, Code Version 6.7.0*, http://www.rmcprofile.org/imagesFhj/d/d5/Rmcprofile_v6.7_manual.pdf.
Tucker, M. G., Keen, D. A., Dove, M. T., Goodwin, A. L. & Hui, Q. (2007). *J. Phys. Condens. Matter*, **19**, 335218.
Verlet, L. (1967). *Phys. Rev.* **159**, 98–103.
Wagner, C. (1985). *J. Non-Cryst. Solids*, **76**, 29–42.
Waller, I. (1923). *Z. Phys.* **17**, 398–408.
Wang, H.-W., Fanelli, V. R., Reiche, H. M., Larson, E., Taylor, M. A., Xu, H., Zhu, J., Siewenie, J. & Page, K. (2014). *Rev. Sci. Instrum.* **85**, 125116.
Warren, B. E. (1934). *Phys. Rev.* **45**, 657–661.
Warren, B. E. (1937). *J. Appl. Phys.* **8**, 645–654.
Warren, B. E. (1990). *X-ray Diffraction.* New York: Dover Publications.
Windsor, C. G. C. G. (1981). *Pulsed Neutron Scattering.* London, New York: Taylor & Francis/Halsted Press.
Wright, A. C. (1990). *J. Non-Cryst. Solids*, **123**, 129–148.
Wright, A. C. (1994). *J. Non-Cryst. Solids*, **179**, 84–115.

Wright, A. C., Clare, A. G., Grimley, D. I. & Sinclair, R. N. (1989). *J. Non-Cryst. Solids*, **112**, 33–47.
Yarnell, J., Katz, M., Wenzel, R. G. & Koenig, S. (1973). *Phys. Rev. A*, **7**, 2130.
Young, C. A. & Goodwin, A. L. (2011). *J. Mater. Chem.* **21**, 6464–6476.
Zernike, V. & Prins, J. (1927). *Z. Phys.* **41**, 184.
Zhang, J. (1999). *Phys. Chem. Miner.* **26**, 644–648.