



# Orientalional order of liquids and glasses *via* fluctuation diffraction

Andrew V. Martin\*

ARC Centre of Excellence for Advanced Molecular Imaging, School of Physics, University of Melbourne, Parkville, Victoria 3010, Australia. \*Correspondence e-mail: andrew.martin@unimelb.edu.au

Received 28 June 2016

Accepted 19 October 2016

Edited by T. Ishikawa, Harima Institute, Japan

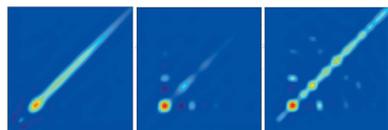
**Keywords:** correlated fluctuations; dynamical studies; XFEL; framework-structured solids and amorphous materials; structure prediction; coherent diffraction.

Liquids, glasses and other amorphous matter lack long-range order, which makes them notoriously difficult to study. Local atomic order is partially revealed by measuring the distribution of pairwise atomic distances, but this measurement is insensitive to orientational order and unable to provide a complete picture of diverse amorphous phenomena, such as supercooling and the glass transition. Fluctuation scattering with electrons and X-rays is able to provide this orientational sensitivity, but it is difficult to obtain clear structural interpretations of fluctuation data. Here we show that the interpretation of fluctuation diffraction data can be simplified by converting it into a real-space angular distribution function. We calculate this function from simulated diffraction of amorphous nickel, generated with a classical molecular dynamics simulation of the quenching of a high temperature liquid state. We compare the results of the amorphous case to the initial liquid state and to the ideal f.c.c. lattice structure of nickel. We show that the extracted angular distributions are rich in information about orientational order and bond angles. The diffraction fluctuations are potentially measurable with electron sources and also with the brightest X-ray sources, like X-ray free-electron lasers.

## 1. Introduction

Phases of matter that lack long-range order, such as liquids, glasses and other amorphous phases, derive their physical properties from short-range ( $<5 \text{ \AA}$ ) and medium-range order ( $5\text{--}20 \text{ \AA}$ ) (Elliott, 1991). However, at these length scales there is a daunting amount of structural variability and we are far from a full understanding of the macroscopic phenomena that arise from local order like the glass transition (Mauro, 2014). While short-range order is often a direct product of chemical bonds, like the well coordinated polyhedra of oxide glasses (Elliott, 1991), the medium-range order that is prevalent in chalcogenides (Salmon *et al.*, 2005) and metallic glasses (Sheng *et al.*, 2006; Wu *et al.*, 2015) is far more complex. Medium-range order is key to the optical, chemical and thermodynamic properties of glass-forming materials (Mauro, 2014) and the possibility of engineering them (Martin *et al.*, 2002; Mauro, 2014), which has made understanding topological order and its extraordinary diversity one of the outstanding challenges in condensed matter and materials science.

Studies of amorphous phases rely heavily on measurements of the pair-distribution function (PDF) (Elliott, 1983; Fischer *et al.*, 2006), which is related to the probability of finding an atom at a distance  $r$  from a given reference atom. The popularity of the PDF is due largely to its accessibility *via* X-ray, neutron or electron diffraction. It is used to validate structural models like molecular dynamics (MD) models (Car & Parrinello, 1985) and as an input to reverse Monte-Carlo techni-



OPEN ACCESS

ques (Fischer *et al.*, 2006). However, the pair distribution lacks information about orientational order and is frequently insufficient to experimentally measure hidden topological order predicted by MD simulations (Wu *et al.*, 2015). The deficiency of the PDF is partially addressed by complimentary techniques like nuclear magnetic resonance and Raman scattering (Elliott, 1983) that can be matched to structural models to infer information about short-range order, like bond angles. Further insight can also be gained by measuring the PDF with elemental specificity with anomalous X-ray diffraction (Elliott, 1983) or with neutrons *via* isotropic substitution (Fischer *et al.*, 2006).

Promisingly, fluctuation electron microscopy is developing into a powerful probe of medium-range order (Treacy *et al.*, 2005). It was used to find nanometre-sized polycrystalline regions in amorphous silicon that were previously undetected in PDF measurements (Treacy & Borisenko, 2012). Fluctuation diffraction microscopy has also been developed for X-rays at 100 nm resolution (Fan *et al.*, 2005) and it has been shown that coherent X-ray diffraction is sensitive to local rotational symmetries in disordered materials (Wochner *et al.*, 2009). Femtosecond X-ray free-electron lasers have been used to study supercooled water (Sellberg *et al.*, 2014) and could be extended to orientational order if X-ray fluctuation diffraction can be pushed to atomic resolution. However, fluctuation measurements typically take the form of statistical or correlation measures of scattered intensity that, aside from rotational symmetries, are difficult to interpret structurally. Detailed numerical forward models of the structure and the diffraction are usually required (Treacy *et al.*, 2005). This stands in stark contrast to the PDF which has a direct, invertible relationship to the mean diffraction signal. A similar invertible relationship to real-space statistical distributions has been lacking for fluctuation diffraction measurements.

Alongside the development of fluctuation microscopy, diffraction fluctuations have been proposed as a route to the structures of biological molecules, like proteins, without the need for crystallization (Kam, 1977). The idea is to measure the diffraction from multiple identical copies of the molecule in liquid suspension and perform a correlation analysis. It can be shown that information about the orientations and separations of individual molecules is lost in the analysis, leaving information that depends only on the internal atomic

structure of the molecule. For some time this research was hampered by an inability to measure proteins faster than their rotational diffusion, thereby washing out the diffraction fluctuations and this was solved by freezing samples (Kam *et al.*, 1981). Nevertheless, it did not become an established technique with synchrotron sources, most likely because the beam intensity required for sufficient signal-to-noise exceeded radiation dose limits. However, there has been an exciting resurgence within the X-ray free-electron laser community (Saldin *et al.*, 2009; Starodub *et al.*, 2012) because femtosecond pulses can take snapshot measurements effectively freezing the molecules in place and outrunning radiation damage processes. This allows potential applications to proteins in solution at room temperature and liquid samples that have rapid decoherence times. Structure determination methods were developed at first for molecules with rotational symmetries (Saldin *et al.*, 2011; Starodub *et al.*, 2012) and more recently for an arbitrary three-dimensional structure (Donatelli *et al.*, 2015).

Underlying fluctuation diffraction methods for biomolecules is a well developed theoretical analysis of intensity correlations in a spherical geometry. Our work here is to show that the same theory can be reapplied and extended in the context of amorphous systems (*e.g.* liquids and glasses) to make fluctuation diffraction data easier to interpret. Here we show how the fluctuations in kinematic far-field diffraction can be mapped into a three- and four-atom correlation function  $\Theta(r, r', \theta)$  that depends on two pairwise distances and one relative angle (as illustrated in Fig. 1). It is given by

$$\Theta(r, r', \theta) \equiv \int \int N_a \langle g^\alpha(\mathbf{r}) g^\alpha(\mathbf{r}') \rangle_\alpha \delta\left(\cos \theta - \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}| |\mathbf{r}'|}\right) d\Omega d\Omega', \quad (1)$$

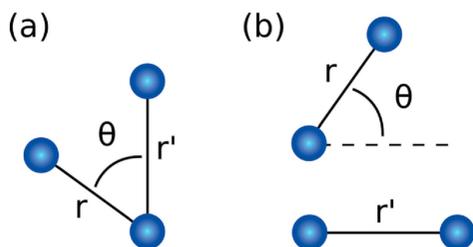
where the average  $\langle \rangle_\alpha$  is taken over the ensemble of possible sample configurations,  $g^\alpha(\mathbf{r})$  is a three-dimensional two-atom correlation function for a particular sample state  $\alpha$ , and  $N_a$  is the mean number of atoms illuminated per measurement.  $\Theta(r, r', \theta)$  is sensitive to orientation in short- and medium-range order that is absent in the PDF and at small values of  $r$  and  $r'$  the angular dependence of  $\Theta(r, r', \theta)$  is determined by bond angles.

$\Theta(r, r', \theta)$  is also related to higher order correlation functions from statistical physics as follows,

$$\Theta(r, r', \theta) = \tilde{g}^{(2)}(r, r', \theta) + \tilde{g}^{(3)}(r, r', \theta) + \tilde{g}^{(3)}(r, r', \pi - \theta) + \tilde{g}^{(4)}(r, r', \theta), \quad (2)$$

where  $\tilde{g}^{(3)}(r, r', \theta)$  and  $\tilde{g}^{(4)}(r, r', \theta)$  can be derived (see Appendix A) from the respective  $n$ -body correlation function  $g^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$  by integrating out the degrees of freedom that the measurement is insensitive to, such as the absolute position of each atomic pair and the absolute orientation of the sample. They are given by

$$\tilde{g}^{(2)}(r, r', \theta) = \int \tilde{g}^{(2)}(\mathbf{r}, \mathbf{r}') \delta\left(\cos \theta - \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}| |\mathbf{r}'|}\right) d\Omega d\Omega', \quad (3)$$



**Figure 1**  
Atom combinations that contribute to the real-space correlation function  $\Theta(r, r', \theta)$  with three atoms (a) and four atoms (b). The four-atom contributions are insensitive to the separation between the two pairs of atoms.

$$\tilde{g}^{(3)}(r, r', \theta) = \int g^{(3)}(\mathbf{r}_1, \mathbf{r}, \mathbf{r}') d\Omega d\phi' d\mathbf{r}_1 \quad (4)$$

and

$$\tilde{g}^{(4)}(r, r', \theta) = \int g^{(4)}(\mathbf{r}_1, \mathbf{r}, \mathbf{r}_3, \mathbf{r}') d\Omega d\phi' d\mathbf{r}_1 d\mathbf{r}_3, \quad (5)$$

where  $\mathbf{r}_1$  and  $\mathbf{r}_3$  are coordinates of the reference atoms in each pair that are integrated out, and  $\Omega$  and  $\phi'$  are angular coordinates that specify the absolute orientation of the three- or four-atom group. The  $n$ -body correlation function  $g^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$  can be described by an average over atomic configurations using delta functions:

$$g^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \frac{1}{\rho^n} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \dots \sum_{i_n \neq i_1, \dots, i_{n-1}} \delta(\mathbf{r}_1 - \mathbf{r}_{i_1}) \delta(\mathbf{r}_2 - \mathbf{r}_{i_2}) \dots \delta(\mathbf{r}_n - \mathbf{r}_{i_n}) \right\rangle, \quad (6)$$

where  $\mathbf{r}_{i_x}$  is the position vector for atom  $i_x$ ,  $N_a$  is the number of atoms contained in a sample volume  $V$ ,  $\delta(\mathbf{r})$  is the delta function and  $\rho$  is the number density ( $\rho = N_a/V$ ). The tilde symbol over the two-body term  $\tilde{g}^{(2)}(r, r', \theta)$  is to indicate that this is not equivalent to the PDF (see Appendix A), but it is effectively one-dimensional because it is only non-zero when  $r = r'$  and  $\theta = 0$  or  $\pi$ .

The function  $\Theta(r, r', \theta)$  contains information about orientational order through the angular dependence  $\theta$ , which is related to internal angles in three- and four-atom correlations as shown in Fig. 1. We use the term orientational order here to refer to non-uniform angular structure in  $\Theta(r, r', \theta)$ . This usage is analogous to the association between the term ‘local order’ and the presence of peaks in the PDF. We thus use the term ‘orientational order’ in broader sense than the term ‘bond orientational order’ (Steinhardt *et al.*, 1983), which was defined with respect to specific angular metrics for quantifying local structure. The relationship between  $\Theta(r, r', \theta)$  and bond orientational order is of interest for future study.

We note that the theory of local rotational symmetries present in the fluctuation X-ray diffraction of disordered systems developed by Altarelli *et al.* (2010) overlaps with the work presented here. The key difference is that they conduct the majority of their analysis in Fourier space and focus on symmetry, whereas here we have identified a transformation to real space that makes a direct connection to statistical physics. The connection with statistical physics has been made in the theory of fluctuation electron microscopy (Gibson *et al.*, 2000) involving correlation functions of the same order as in equation (2). There a forward model of diffraction is presented that is specific to their experimental geometry and they did not identify the inverse relationship that we present here.

## 2. Methods

### 2.1. Derivation of $\Theta(r, r', \theta)$

$\Theta(r, r', \theta)$  is obtained from fluctuation diffraction by first calculating the angular correlations of diffraction patterns

averaged over an ensemble of different states of the sample. We assume the sample can take an atomic configuration  $\alpha$  from a statistical ensemble of possible configurations. We follow the practice developed for pair-distribution analysis (Fischer *et al.*, 2006) and rescale the diffracted intensity to isolate the structural information. It is convenient to work with  $g^\alpha(q, \theta(q), \phi)$  [the Fourier transform of  $g^\alpha(\mathbf{r})$ ], which is obtained by rescaling the kinematic diffracted intensity

$$g^\alpha[q, \theta(q), \phi] = \frac{I[q, \theta(q), \phi]}{\rho_0 \phi_0 N_a |\bar{f}(q)|^2} - \frac{1}{\rho_0} - \delta(\mathbf{q}), \quad (7)$$

where  $\bar{f}(q)$  is the mean atomic scattering factor,  $N_a$  is the number of atoms in the beam,  $\phi_0$  depends on experimental parameters (see Appendix B) and  $\rho_0$  is the mean number density of the sample. We assume these parameters are known. The coordinate  $q$  is the magnitude of the scattering vector,  $\phi$  is an azimuthal angle around the beam axis and  $\theta(q)$  is a polar angle with respect to the beam axis (Saldin *et al.*, 2009). The last term,  $\delta(\mathbf{q})$ , represents low-angle scattering from the mean density which in practice is not measured, but which we include here for completeness. The function  $g^\alpha(\mathbf{q})$  is the Fourier transform of a three-dimensional pair correlation function  $g^\alpha(\mathbf{r})$  for the sample in state  $\alpha$ .

We construct an angular cross-correlation in a similar fashion to existing fluctuation diffraction methods (Kam, 1977; Wochner *et al.*, 2009; Saldin *et al.*, 2009) and take the average over  $N_\alpha$  measurements of the sample in different structural states  $\alpha$ :

$$C(q, q', \Delta\phi) = \frac{1}{N_\alpha} \sum_{\alpha=1}^{N_\alpha} \int g^\alpha[q, \theta(q), \phi] g^\alpha[q', \theta(q'), \phi + \Delta\phi] d\phi. \quad (8)$$

where  $\alpha$  is summed over the number of measurements (and, therefore, sample states) and  $N_\alpha$  is the number of samples that are measured.

Here we make a key assumption that there is no correlation between the orientation of the sample state and the beam axis. The average of  $C(q, q', \Delta\phi)$  over an ensemble of sample configurations leads to the loss of information about absolute orientation as the number of measurements  $N_\alpha$  increases. This assumption lets us use a powerful result from the fluctuation diffraction theory for biomolecules that establishes a relationship between  $C(q, q', \Delta\phi)$  and a series of mutual intensity matrices in a spherical harmonic representation (Saldin *et al.*, 2009) as follows

$$C(q, q', \Delta\phi) = \sum_l \frac{1}{4\pi} P_l \left( \frac{\mathbf{q} \cdot \mathbf{q}'}{|\mathbf{q}| |\mathbf{q}'|} \right) \langle B_l^\alpha(q, q') \rangle_\alpha, \quad (9)$$

where  $P_l(x)$  is a Legendre polynomial,  $\langle \rangle_\alpha$  represents an ensemble average and  $B_l^\alpha(q, q')$  is related to a spherical harmonic expansion coefficients  $I_{lm}^\alpha(q')$  of  $g^\alpha(\mathbf{q})$  [the Fourier transform of  $g^\alpha(\mathbf{r})$ ] by

$$B_l^\alpha(q, q') = \sum_{|m| \leq l} I_{lm}^{\alpha,*}(q) I_{lm}^\alpha(q'), \quad (10)$$

and the spherical harmonic expansion of  $g^\alpha(\mathbf{q})$  is written

$$g^\alpha(\mathbf{q}) = \sum_{l=0}^{\infty} \sum_{m \leq |l|} I_{lm}^\alpha(q) Y_{lm}(\theta, \phi), \quad (11)$$

where  $Y_{lm}(\theta, \phi)$  is a spherical harmonic function. Equation (9) is a linear system of equations that can be inverted with standard methods to obtain  $\langle B_l^\alpha(q, q') \rangle_\alpha$ , as is done in solution-based biomolecule scattering methods (Saldin *et al.*, 2009).

In the fluctuation diffraction theory for biomolecules (Kam, 1977; Saldin *et al.*, 2009), multiple molecules with identical structure are illuminated and the spherical harmonic expansion defined for the diffracted intensity of a single molecule. In that context, the extracted matrices  $B_l(q, q')$  are used to recover a real-space image of the particle or molecule *via* phase retrieval, either by first recovering a three-dimensional Fourier intensity (Saldin *et al.*, 2009) or by using them directly as constraints for the image recovery algorithm (Donatelli *et al.*, 2015). In the context of disordered phases of matter, however, every measured state of the sample has a different atomic structure and an imaging analysis is not appropriate. Instead, we diverge from the biomolecular fluctuation diffraction methods by converting the mutual intensity matrices into  $\Theta(r, r', \theta)$  *via* a series of linear transformations and thereby extracting statistical information about short- and medium-range order.

First we use a spherical Bessel transform to map  $B_l(q, q')$  into a function of two real space variables  $r$  and  $r'$ . We write the transform as an operator  $\hat{\mathcal{S}}_{qr}$  that is given by

$$\hat{\mathcal{S}}_{qr}[f(q)] = 4\pi \int_0^{q_{\max}} f(q) j_l(2\pi qr) q^2 dq. \quad (12)$$

Applying the transform twice, we obtain real-space matrices given by

$$\begin{aligned} \langle B_l^\alpha(r, r') \rangle_\alpha &\equiv (-1)^l \hat{\mathcal{S}}_{q'r'} \left\{ \hat{\mathcal{S}}_{qr} [\langle B_l^\alpha(q, q') \rangle_\alpha] \right\} \\ &= \sum_{|m| \leq l} \langle I_{lm}^{\alpha*}(r) I_{lm}^\alpha(r') \rangle_\alpha, \end{aligned} \quad (13)$$

where  $I_{lm}^\alpha(r)$  are the spherical harmonic coefficients of  $g^\alpha(\mathbf{r})$ . Writing out  $I_{lm}^\alpha(r)$  explicitly as a projection of  $g^\alpha(\mathbf{r})$  onto the spherical harmonic basis  $Y_{lm}(\theta, \phi)$ , we can evaluate

$$\begin{aligned} \langle B_l^\alpha(r, r') \rangle_\alpha &= \iint \langle g^\alpha(\mathbf{r}) g^\alpha(\mathbf{r}') \rangle_\alpha \sum_m Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') d\Omega_r d\Omega_{r'} \\ &= \iint \langle g^\alpha(\mathbf{r}) g^\alpha(\mathbf{r}') \rangle_\alpha P_l \left( \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}| |\mathbf{r}'|} \right) \frac{2l+1}{4\pi} d\Omega_r d\Omega_{r'}, \end{aligned} \quad (14)$$

where we have used the following relation to derive the second line

$$\sum_m Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = P_l \left( \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}| |\mathbf{r}'|} \right) \frac{2l+1}{4\pi}. \quad (15)$$

We can construct the following weighted sum to recover  $\Theta(r, r', \theta)$ :

$$\begin{aligned} \Theta(r, r', \theta) &\equiv 2\pi N_a \sum_l P_l(\cos \theta) \langle B_l^\alpha(r, r') \rangle_\alpha \\ &= \iint N_a \langle g^\alpha(\mathbf{r}) g^\alpha(\mathbf{r}') \rangle_\alpha \delta \left( \cos \theta - \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}| |\mathbf{r}'|} \right) d\Omega_r d\Omega_{r'}, \end{aligned} \quad (16)$$

where we have used a known relation for Legendre polynomials:

$$\frac{1}{2} \sum_l (2l+1) P_l(x) P_l(y) = \delta(x-y). \quad (17)$$

The inclusion of the  $N_a$  term in equation (16) corrects for the fact that diffraction fluctuations scale as  $N_a^{1/2}$ , not  $N_a$ , and makes  $\Theta(r, r', \theta)$  independent of the number of atoms in the sample.

In practice, to recover  $\Theta(r, r', \theta)$  from diffraction measurements  $I(\mathbf{q})$ , we need to perform the following steps: (i) calculate the angular correlation of each diffraction pattern and average them [equation (8)]; (ii) invert a system of linear equations to recover  $\langle B_l^\alpha(q, q') \rangle_\alpha$  [equation (9)]; (iii) map  $\langle B_l^\alpha(q, q') \rangle_\alpha$  into real space by numerically applying the spherical Bessel transform for both  $q$  and  $q'$  variables at each value of  $l$  [equation (13)]; and (iv) sum the resulting real-space functions weighted by the Legendre polynomials [equation (16)].

It turns out that ignoring the  $l=0$  contribution to equation (16), produces  $\Theta(r, r', \theta) - g(r)g(r)$ , which provides a clearer representation of the angular information as shown in Fig. 3. As a by-product of ignoring  $l=0$  we do not need to explicitly subtract the  $1/\rho_0$  term when rescaling the intensity using equation (7).

## 2.2. Angular symmetries of $\Theta(r, r', \theta)$

$\Theta(r, r', \theta)$  has angular symmetries which make it unique only in the range  $0 < \theta < \pi/2$ . Averaging over absolute orientation has the effect of making  $\Theta(r, r', \theta)$  an even function of  $\theta$ , *i.e.*  $\Theta(r, r', -\theta) = \Theta(r, r', \theta)$ . Mathematically, this is because angular information is captured by  $\Theta(r, r', \theta)$  *via* a dot product between the displacement vectors between atomic pairs,  $\mathbf{r} \cdot \mathbf{r}'$ , which is proportional to  $\cos \theta$ , which is an even function of  $\theta$ .

A second angular symmetry arises because  $g^\alpha(\mathbf{r})$  is centrosymmetric, *i.e.*  $g^\alpha(-\mathbf{r}) = g^\alpha(\mathbf{r})$ . We can write  $g^\alpha(\mathbf{r})$  in terms of displacement vectors  $\mathbf{R}_{ij}^\alpha$  between atoms  $i$  and  $j$  in sample configuration  $\alpha$ ,

$$g^\alpha(\mathbf{r}) = \frac{1}{\rho} \sum_i \sum_{j \neq i} \delta(\mathbf{r} - \mathbf{R}_{ij}^\alpha). \quad (18)$$

In the above sum, every pair of non-identical atoms contributes two terms with displacement vectors  $\mathbf{R}_{12}$  and  $\mathbf{R}_{21}$ , which leads to centrosymmetry. A consequence of this symmetry is that  $\Theta(r, r', \pi - \theta) = \Theta(r, r', \theta)$ .

### 3. Simulations

#### 3.1. Simulation of the extraction of $\Theta(r, r', \theta)$

We extracted  $\Theta(r, r', \theta)$  from simulated kinematic diffraction data from nickel in amorphous, liquid and crystalline states. Amorphous nickel is a metallic glass with known medium-range order that can be simulated with classical MD simulations by rapidly quenching from a liquid state. As a by-product, we also generate high-temperature liquid states (3000 K).

The MD simulations were performed with the *LAMMPS* package (Plimpton, 1995), using the embedded-atom potential and the parameters from Lu & Szpunar (1997). The supercell size was  $11 \times 11 \times 11$  f.c.c. unit cells with lattice parameter  $3.52 \text{ \AA}$ , which contains 5324 atoms. Atoms were randomly displaced from equilibrium positions up to a maximum distance of  $0.17 \text{ \AA}$ , then the system was equilibrated at 3000 K for 40000 time steps using the isothermal-isobaric ensemble (NPT). Each time step was 2.5 fs. The system was then cooled at a rate of  $248 \text{ K ps}^{-1}$  to 300 K and then run for a further 40000 time steps at 300 K. One thousand different amorphous configurations were generated by the same procedure. The liquid states were recorded after the first 40000 time steps at 3000 K. The PDFs generated by the simulation are shown in Fig. 2 (bottom right) and the bond angle distribution is given by the dashed lines for the distance  $r = 2.5 \text{ \AA}$  shown in Fig. 3 (bottom left). Coordination statistics for the amorphous state are: 25.7% 12-fold coordinated, 52.5% 13-fold coordinated and 17.8% 14-fold coordinated. These were calculated including neighbours within the first minimum of the PDF located at  $3 \text{ \AA}$ .

For a simple illustration of the principle, we have used small sample volumes ( $<40 \text{ \AA}$ ) and not modelled noise on the diffraction patterns. The experimental implications of noise and larger sample volumes are discussed in Section 4.3.

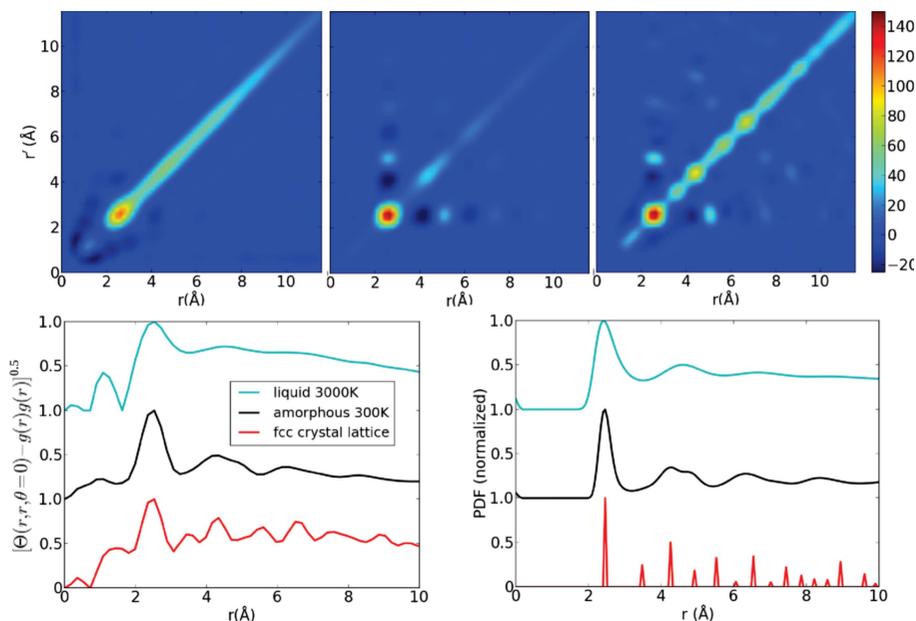
X-ray kinematic diffraction simulations were performed for each liquid and amorphous sample configuration at a wavelength of  $0.5 \text{ \AA}$  and a maximum scattering amplitude of  $1.38 \text{ \AA}^{-1}$ . Atomic scattering factors were taken from Waasmaier & Kirfel (1995). Atoms within a  $40 \text{ \AA}$  diameter sphere were used in the diffraction calculation, ensuring adequate sampling on a  $128 \times 128$   $q$ -space grid. This was purely for computational convenience, as larger sample volumes and finer grid sizes do not change the final result, but are slower to calculate and require averaging over more configurations to converge. No absorption was modelled.

The inversion of equation (9) was performed with singular value decom-

position (SVD) using a maximum  $l$  value of 40. The elements of the matrix to be inverted are given by

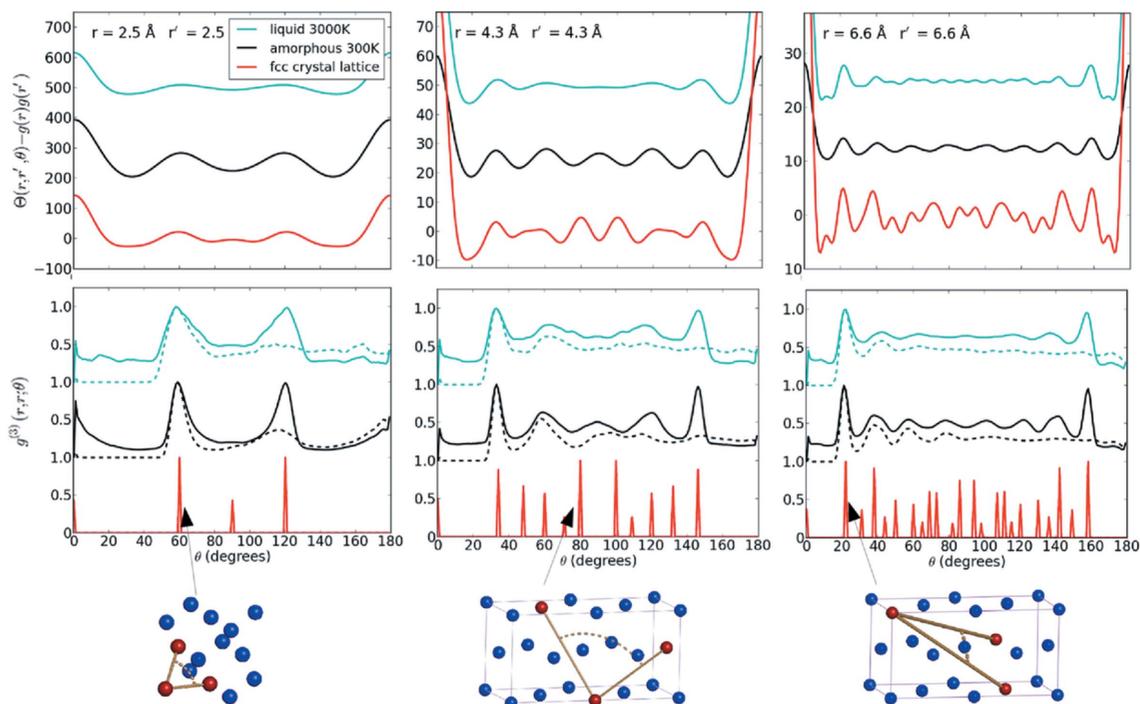
$$M_{lk} = \frac{1}{4\pi} P_l[\sin \theta(q) \sin \theta(q') + \cos \theta(q) \cos \theta(q') + \cos(\Delta\phi_k)], \quad (19)$$

where  $\Delta\phi_k$  is a discrete sample of the internal angle between two pixel coordinates  $\mathbf{q}$  and  $\mathbf{q}'$ , and  $\theta(q)$  is a polar angle that depends on the magnitude of a pixel coordinate vector [see equation (9) for further details on the geometry]. A separate matrix inversion is thus required for each value of  $q$  and  $q'$ . We use 402 angular sampling points ( $k$  points). Friedel symmetry was applied by setting  $I_{lm}(q) = 0$  for odd values of  $l$ , and excluding these from the SVD analysis greatly improved numerical stability. The condition number of the inversion is around 10 when  $|q - q'| \sim 0$ , and worsens as  $|q - q'|$  increases. A cutoff was applied to exclude all singular values below 5% of the largest singular value. The discrete spherical Bessel transform (DSBT) (Lanusse *et al.*, 2012) was used to map the  $\mathcal{B}_l(q, q')$  matrices to real space [equation (13)]. The DSBT has a boundary condition that requires  $I_{lm}(q_{\max}) = 0$ . In our simulations,  $q_{\max}$  is set to the value measured at the edge of the detector and a Gaussian filter was applied to each diffraction pattern with a width of  $q_{\max}/4$  to ensure the boundary condition was met, thereby minimizing numerical errors for a reduction in the effective resolution by a factor of 4. The reduced resolution is likely to contribute to the unphysical non-zero correlation below  $2.5 \text{ \AA}$  shown in Fig. 2.



**Figure 2**

When evaluated at  $\theta = 0$ ,  $\Theta(r, r', \theta)$  provides information about radial correlations. The top row shows the  $\Theta(r, r', \theta = 0)$  extracted from diffraction simulations of liquid Ni at 3000 K (left), amorphous Ni at 300 K (middle) and an f.c.c. Ni crystal (right). On the bottom left we plot the  $r = r'$  diagonal which contains peak structure that is similar qualitatively to the pair distribution function (bottom right). All line plots have been normalized to have a maximum value of 1 to aid comparison. The information less than  $2.5 \text{ \AA}$  is a numerical artifact related to the finite resolution of the simulation.


**Figure 3**

The angular dependence of  $\Theta(r, r', \theta)$  for three different radial shells:  $r = r' = 2.5 \text{ \AA}$  (top left),  $r = r' = 4.3 \text{ \AA}$  (top middle) and  $r = r' = 6.6 \text{ \AA}$  (top right). The peak positions show good agreement with the sum  $\tilde{g}^{(3)}(r, r', \theta) + \tilde{g}^{(3)}(r, r', \pi - \theta)$  shown by the solid lines in the second row that are calculated directly from the simulated atomic structures. The dashed lines in the second row show the asymmetric correlation  $\tilde{g}^{(3)}(r, r', \theta)$  for reference. All  $\tilde{g}^{(3)}(r, r', \theta)$  plots have been normalized to a maximum value of unity. The lattice diagrams are examples of the atoms from the f.c.c. lattice (in red) that contribute to peaks in  $\tilde{g}^{(3)}(r, r', \theta)$ . To aid comparison, the plots have been offset and  $\Theta(r, r', \theta)$  has been scaled by a factor of  $7 \times 10^{-5}$  for the crystalline case and by factor of 5 for the liquid case.

The  $\Theta(r, r', \theta)$  extracted from our simulation displays a peak structure that is dominated by two- and three-atom correlations. Two-atom correlations are prominent along the line  $r = r'$  and  $\theta = 0$  as shown in Fig. 2, which displays peaks in the same positions as the pair-distribution function for all three phases of the sample. Three-atom correlations appear when we plot the  $\Theta(r, r', \theta)$  as a function of  $\theta$ , revealing the orientational order. The majority of angular peaks shown in the top row of Fig. 3 correspond well to  $\tilde{g}^{(3)}(r, r', \theta)$  calculated directly from the atomic structures, as shown in the bottom row of Fig. 3.

## 4. Prospects for experiment

### 4.1. Electrons

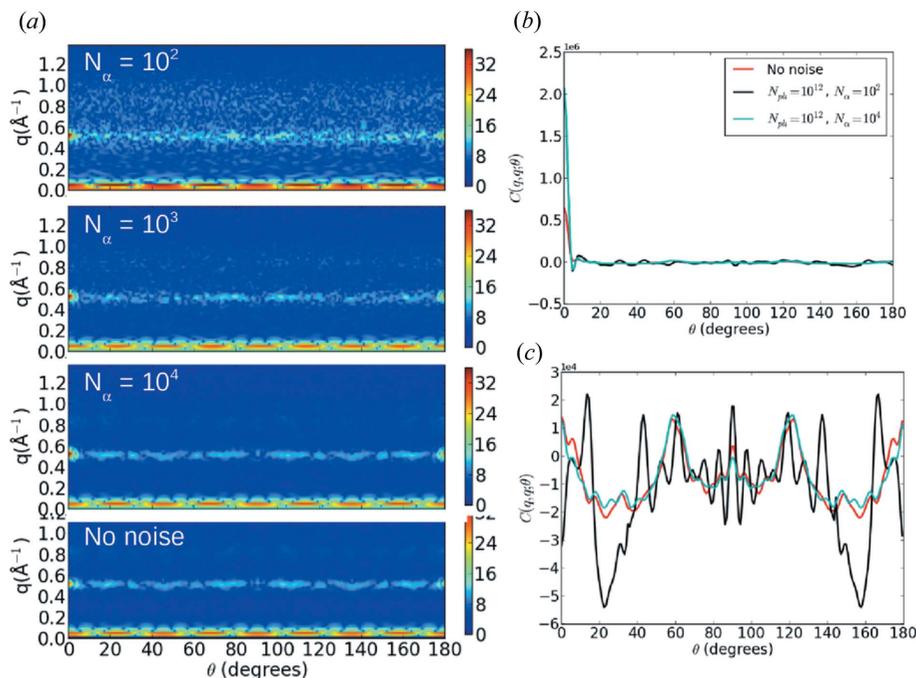
Measurements of  $\Theta(r, r', \theta)$  should be possible *via* fluctuation electron diffraction. Cross-correlation functions similar to equation (8) have already been measured with electrons (Gibson *et al.*, 2010) and could be converted into a measurement of  $\Theta(r, r', \theta)$  using the methods described here. For electrons, multiple scattering will place limits on sample thickness of around 100–200 nm and further work is needed to establish these limits more precisely. Effects of multiple scattering can be corrected in PDF analysis (Anstis *et al.*, 1988) and it may be possible to extend these techniques to diffraction fluctuation measurements. Beam profile effects are

another known issue (Gibson *et al.*, 2000) that are important if there is non-uniform illumination on length scales below the correlation length of the sample. A further issue is induced sample dynamics that have been invoked to account for the observation of lower than expected contrast in electron intensity correlation measurements (Rezikyan *et al.*, 2015).

### 4.2. X-rays

An advantage for X-rays over electrons is that the kinematic diffraction approximation has greater validity. However, it is more difficult to focus X-ray beams and diffraction fluctuations diminish unfavourably relative to the total scattering as the number of illuminated atoms increases. The best nanofocus X-ray beams are 15 nm or better in diameter, but more commonly they are greater than 25 nm (Sakdinawat & Attwood, 2010), which is at least an order of magnitude larger than the length scales of short- and medium-range order in amorphous matter. Since shot noise increases with the total scattering with the illuminated volume, we would expect that larger illuminated sample volumes require more diffraction measurements for convergence. Shot noise is investigated in detail in the next section.

It is known from fluctuation diffraction of biomolecules that the contribution of random correlations of the diffraction from different molecules produce a noise in the intensity correlation that scales linearly with the number of molecules.


**Figure 4**

(a) The effect of noise on a cross-section of the intensity correlation function  $C(q = 0.5 \text{ \AA}^{-1}, q', \theta)$  varying the number of patterns. Noise is clearly visible for  $N_\alpha = 10^2$ , but is suppressed by  $N_\alpha = 10^4$  except for a large peak at  $\theta = 0$  which arises from the self-correlation of the shot noise with itself. (b) A line plot of  $q = q' = 0.5 \text{ \AA}^{-1}$  which shows that the peak at  $\theta = 0$  varies with the noise level. This peak can be removed with a centrosymmetric filter that replaces information in the vicinity of  $\theta = 0^\circ$  with the information near  $\theta = 180^\circ$ . (c) After applying the centrosymmetric filter, the  $N_\alpha = 10^4$  result agrees well with the noise-free simulation for the full range of  $\theta$ .

As the signal from the autocorrelation of diffraction from each molecule also scales linearly, the ratio of signal to this source of noise is independent of the number of molecules and bigger than one. An equivalent result is expected to hold for amorphous materials with respect to the number of atoms. Hence, we anticipate that noise from uncorrelated atoms will not be as important as shot noise for the feasibility of fluctuation X-ray experiments of amorphous materials.

One key difference between measuring amorphous materials and biomolecules is that biomolecules are typically delivered to the beam in a liquid environment, which generates background scattering that increases noise. It has been estimated to overcome background scattering that around  $10^{13}$  photons per pulse would be required to image protein structures (Kirian *et al.*, 2011; Kirian, 2012), which is just beyond the reach of current XFEL facilities (Emma *et al.*, 2010). However, this issue is not present if we are studying the liquid itself or an amorphous solid, which can be placed in an X-ray beam effectively in isolation from other scattering material. In both experiments there are stray background signals from the X-ray beamline, but methods to reduce these to single photon level are under development for single molecule imaging (Aquila *et al.*, 2015), which if necessary could be used to for fluctuation diffraction measurements.

These measurements do not require greater beam coherence than typical small-angle X-ray scattering measurements, because the structural correlation length of the sample does not typically extend beyond 2 nm. Although high beam coherence has been a feature of fluctuation X-ray measure-

ments at 100 nm length scales, it becomes less critical as the structural features under investigation become smaller.

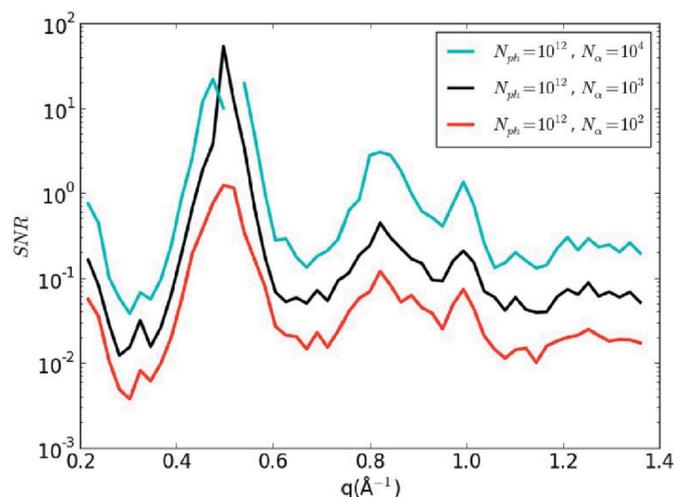
### 4.3. Investigation of the effects of noise

We have investigated the number of patterns required with a statistical model, which is based on an observation from simulation that the diffraction fluctuations scale with  $N_a^{1/2}$ .

This simulation and the statistical model are detailed in Appendix B. We can use these statistical properties of the diffraction to derive the number of diffraction patterns required to measure the diffraction fluctuations as a function of experimental parameters. It turns out that the required number of patterns is independent of  $N_a$ , because both the diffraction fluctuations and shot noise have the same dependence on this parameter. The required number of patterns is sensitive to flux and therefore to the focal spot size. In fact, we found that the number of patterns  $N_\alpha$  required to measure the diffracted intensity correlation has the following relationship to the beam area  $A$ , the number of incident photons  $N_{\text{ph}}$  and the mean atomic scattering factor  $\bar{f}$ :

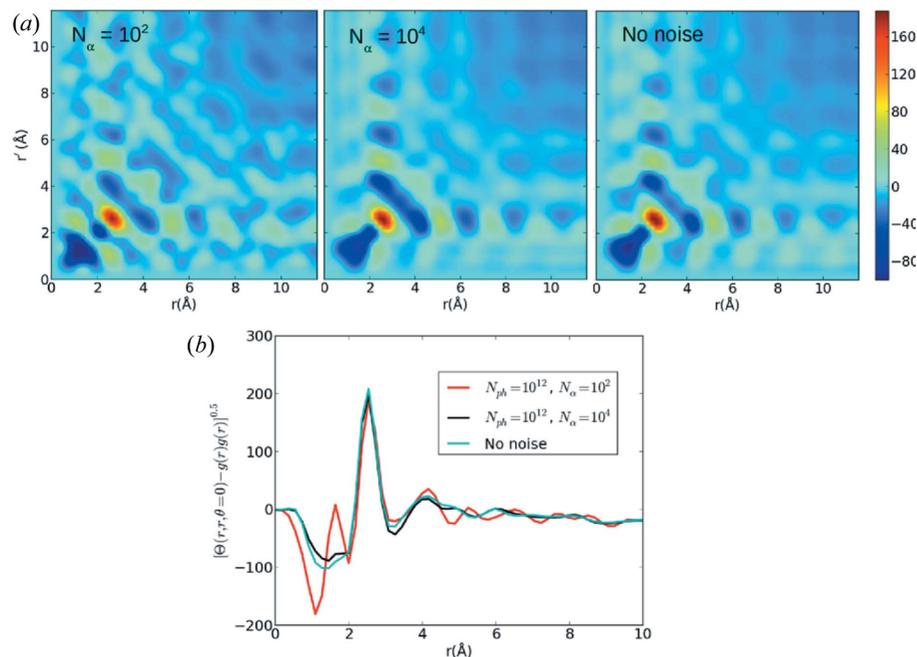
$$N_\alpha \propto A^2 N_{\text{ph}}^{-2} \bar{f}^{-4}. \quad (20)$$

The quadratic and quartic powers indicate that small changes in beam parameters or sample composition can have a big impact on the required number of patterns and place practical limits on the accuracy of the measurement. It will be more challenging to measure lighter elements as  $\bar{f} \sim Z$ . High repetition rate X-ray sources ( $>100$  Hz) like X-ray lasers can


**Figure 5**

The numerically calculated SNR as function of the number of patterns. The signal is the standard deviation of the noise-free calculation and the noise level is calculated from the difference between the variance of each noisy simulation and that of the noise-free simulation. The increase of the SNR is proportional to  $N_\alpha^{1/2}$ , as predicted by the theory in Appendix B.

produce of the order of  $10^7$ – $10^8$  measurements in a 24 h period, so we regard this as the upper limit on the number of patterns available in a single experiment. Assuming parameter values available at the Linac Coherent Light Source (Emma *et al.*, 2010) (100 nm diameter beam and  $10^{12}$  incident photons at 1.5 Å wavelength), we estimate that the correlation function from equation (8) for amorphous nickel could be measured


**Figure 6**

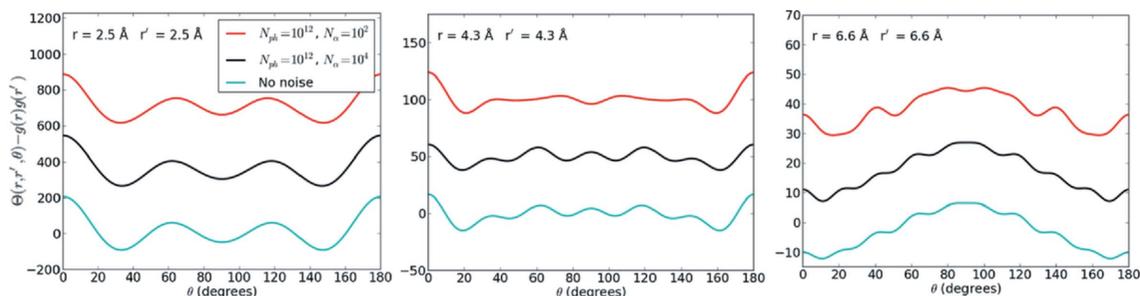
(a) The effect of noise on  $\Theta(r, r', \theta = 0)$  for amorphous nickel varying the number of patterns. The centrosymmetric filter has been applied to the intensity correlations. For  $N_\alpha = 10^2$ , noise manifests in the reconstruction, but  $N_\alpha = 10^4$  agrees well with the noise-free simulation. *b* The  $r = r'$  diagonal shows that  $N_\alpha = 10^2$  agrees with the noise-free simulation at the first-nearest-neighbour peak (2.5 Å), but disagrees elsewhere.

with a signal-to-noise ratio (SNR) of 5 by collecting  $10^5$  diffraction patterns, which requires less than 20 min of data collection at 100 Hz.

We have verified these estimates for the required number of patterns with a numerical implementation of the statistical model. To avoid intensive MD simulations, we took the diffraction pattern calculated for a small sample volume (40 Å diameter; 2900 atoms), then scaled the mean scattering signal (as a function of  $q$ ) by the number of atoms and scaled the fluctuations by the square root of the number of atoms, then combined both to create a new diffraction pattern. Shot noise was then calculated for the scaled diffraction pattern. The new pattern will have the correct ratio between shot noise and the interference terms between atoms with correlated positions. It does not model the noise from atoms with uncorrelated positions, which is a smaller effect than shot noise. To limit the number of MD simulations required, we randomly selected a sample state from 1000 different MD results and randomly rotated each selected structure. We fixed the experimental parameters at 100 nm diameter spot size,  $10^{12}$  incident photons and  $8.6 \times 10^7$  illuminated atoms. The number of atoms was calculated for a 100 nm sample thickness and a number density of 85.9 atoms  $\text{nm}^{-3}$  which was taken from the MD simulation. Fig. 4(a) shows cross-sections of  $C(q, q', \theta)$  as  $N_\alpha$  is varied. We see that the correlation function converges to the noise free calculation by  $10^4$  patterns, except for the self-correlation of the shot noise that generates a large peak when  $q = q'$  and  $\theta = 0^\circ$ , as shown in Fig. 4(b). This peak can be removed by applying centrosymmetry to replace the data near

$\theta = 0^\circ$  with the data measured around  $\theta = 180^\circ$ , which is appropriate if we ignore absorption. In our previous simulations centrosymmetry was already applied during the reconstruction of  $\Theta(q, q', \theta)$ . After filtering the noise peak, the  $N_\alpha = 10^4$  simulation agrees well with the noise-free simulation over the full range of  $\theta$  as shown in Fig. 4(c). The numerically computed signal-to-noise estimates shown in Fig. 5 agree well with the analytic estimates, as the conservative limit of SNR = 5 has not yet been reached at most scattering angles for  $N_\alpha = 10^4$ .

Fig. 6 shows how the radial information contained in  $\Theta(r, r', \theta)$  converges as the number of patterns increases. At  $N_\alpha = 10^2$  noise artefacts are apparent over the full range of  $r$  values, but disappear as the number of patterns increases to  $10^4$ , which shows good agreement with the noise-free simulation. Interestingly, even at  $N_\alpha = 10^2$  when shot noise exceeds the fluctuation signal for all scattering angles (the SNR of the intensity correlation is equal or below one),  $\Theta(r, r', \theta)$  is still reasonably



**Figure 7**  
Angular distributions for amorphous nickel varying the number of patterns. The centrosymmetric filter has been applied to the intensity correlations. For  $N_\alpha = 10^2$ , noise significantly alters the peak structure beyond the first-nearest-neighbour distance ( $>2.5$  Å). For  $N_\alpha = 10^4$ , there is good agreement with the noise-free calculation. The plots have been offset to aid comparison.

accurate in the strong first- and second-nearest-neighbour peaks as shown in Fig. 6(b). This may be because the extraction of  $\Theta(r, r', \theta)$  effectively applies an angular bandwidth limit and also because the calculation of the  $B_i(q, q')$  matrices is regularized. In other words, the diffraction fluctuations have a correlation on the detector that spans many pixels (*i.e.* oversampling) which are exploited in the conversion to real space to suppress noise. The angular distributions shown in Fig. 7 indicate that  $N_\alpha = 10^2$  noise changes the angular peak structure of  $\Theta(r, r', \theta)$  beyond the first-nearest neighbours, but these effects are largely gone by  $N_\alpha = 10^4$ .

Although we have taken indicative parameters from an XFEL for our noise analysis, it would be very interesting to explore whether these measurements can be made at nanofocus synchrotron beams, which are more accessible than XFELs. The number of X-rays available at a synchrotron source per second is comparable to the number in a single XFEL pulse. However, the efficiency of the X-ray focusing optics will be critical to delivering a high number of incident photons per measurement. For a continuous source, equation (20) shows that it is more advantageous to increase exposure time to reduce the number of measurements required. However, the maximum possible exposure time will be limited by other factors like radiation damage and instrument stability.

## 5. Conclusion

We have shown that kinematic diffraction fluctuations can be mapped into a real-space correlation function that provides distributions of bond angles and orientational order. The correlations should be measurable with electrons for samples thin enough to avoid multiple scattering and with X-rays provided sufficiently high intensity and data rates can be achieved to overcome noise. Our analysis of the latter issue indicates that measurements of orientational order in metallic glasses, like nickel, should be within reach of current XFEL facilities and possibly nanofocus synchrotron beams by taking of the order of  $10^4$ – $10^5$  measurements. Our noise model predicts that lighter elements such as carbon or oxygen will be harder to measure and may require at least two orders of magnitude more measurements. The situation improves for organic molecules as then it is the number of molecules not the

number of atoms that is relevant, as is known from the existing theory for the fluctuation diffraction of biomolecules.

There are a wide range of current applications for pair distribution analysis that can potentially be extended to the orientational analysis proposed here. Aside from glasses and amorphous solids, ultrafast X-ray or electron sources could be used to probe orientational order in liquid or gas phases (*e.g.* airborne particulate matter (Loh *et al.*, 2012)) because the ultrafast pulses outrun the translational and rotational motion of the sample. Ultrafast pulses could extend time-resolved small-angle or wide-angle X-ray scattering to orientational order. Real-space correlations could be used to add a new dimension to solution-scattering methods for biological structure, which were a key inspiration for our work. Further applications can be envisioned to heterogeneous biological systems such as unfolded or partially folded conformational ensembles (Lipfert & Doniach, 2007) and also to study the dynamics of these systems.

## APPENDIX A Relationship between $\Theta(r, r', \theta)$ and correlation functions

In this appendix, we describe how  $\Theta(r, r', \theta)$  can be written in terms of correlation functions from statistical physics.

The  $n$ th-order atomic correlation function can be written as a formal count over the atoms in the sample using delta functions

$$g^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \frac{1}{\rho^n} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \dots \sum_{i_n \neq i_1, \dots, i_{n-1}} \delta(\mathbf{r}_1 - \mathbf{r}_{i_1}) \delta(\mathbf{r}_2 - \mathbf{r}_{i_2}) \dots \delta(\mathbf{r}_n - \mathbf{r}_{i_n}) \right\rangle, \quad (21)$$

where  $\mathbf{r}_{i_x}$  is the position vector for atom  $i_x$ ,  $N_a$  is the number of atoms contained in a sample volume  $V$ ,  $\delta(\mathbf{r})$  is the delta function and  $\rho$  is the number density ( $\rho = N_a/V$ ). The ensemble average is denoted by  $\langle \rangle$ .

Kinematic scattering is sensitive to a reduced form of  $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ , because the absolute position of each atom pair is not measured, only the relative distance. We can express this

by explicitly integrating out the unmeasured degrees of freedom. First we make a change of coordinates to place one atom at the origin (the reference atom), so that  $\mathbf{r}_{i_1} = 0$  and  $\mathbf{r}_{i_2}$  is the relative displacement between the two atoms, which for clarity we denote  $\mathbf{r}_{i_2}$ . We then integrate out  $\mathbf{r}_1$  as follows

$$\begin{aligned} g^{(2)}(\mathbf{r}) &\equiv \frac{1}{V} \int g^{(2)}(\mathbf{r}_1, \mathbf{r}) d\mathbf{r}_1 \\ &= \frac{1}{V\rho^2} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \left[ \int \delta(\mathbf{r}_1) d\mathbf{r}_1 \right] \delta(\mathbf{r} - \mathbf{r}_{i_2}) \right\rangle \\ &= \frac{1}{V\rho^2} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \delta(\mathbf{r} - \mathbf{r}_{i_2}) \right\rangle \end{aligned} \quad (22)$$

To evaluate  $\Theta(r, r', \theta)$  we need to consider a two-body correlation function for a single instance of the sample  $g^{(2,\alpha)}(\mathbf{r})$  which has the same form as equation (21) but is not averaged over the ensemble. We then evaluate

$$N_a \langle g^{(2,\alpha)}(\mathbf{r}) g^{(2,\alpha)}(\mathbf{r}') \rangle = \frac{N_a}{\rho^4 V^2} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \sum_{i_3=1}^N \sum_{i_4 \neq i_3} \delta(r - r_{i_2}) \delta(r' - r_{i_4}) \right\rangle. \quad (23)$$

Equation (23) resembles the fourth-order correlation function  $g^{(4)}$  with two reference atoms and two coordinates have been integrated out (the absolute position and the relative displacement between the two pairs), except that the limits of the sums over atoms include cases where atoms are identical, whereas equation (21) does not. In fact, the terms where two pairs of indices are equal (*e.g.*  $i_1 = i_3$  and  $i_2 = i_4$ ) generate two-body correlations and terms where one pair of indices are equal generate three-body terms. Explicitly writing out the sums over identical atoms we find

$$\begin{aligned} N_a \langle g^{(2,\alpha)}(\mathbf{r}) g^{(2,\alpha)}(\mathbf{r}') \rangle &= \frac{1}{\rho^3 V} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' - \mathbf{r}_{i_2}) + \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' + \mathbf{r}_{i_2}) \right\rangle \\ &\quad + \frac{1}{\rho^3 V} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \sum_{i_3 \neq i_1, i_2} \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' - \mathbf{r}_{i_3}) \right. \\ &\quad \left. + \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' + \mathbf{r}_{i_3}) \right\rangle \\ &\quad + \frac{N_a}{\rho^4 V^2} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \sum_{i_3=i_1, i_2} \sum_{i_4 \neq i_1, i_2, i_3} \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' - \mathbf{r}_{i_4}) \right\rangle \\ &= \tilde{g}^{(2)}(\mathbf{r}, \mathbf{r}') + \frac{1}{V} \int [g^{(3)}(\mathbf{r}_1, \mathbf{r}, \mathbf{r}') + g^{(3)}(\mathbf{r}_1, \mathbf{r}, -\mathbf{r}')] d\mathbf{r}_1 \\ &\quad + N_a \frac{1}{V^2} \int \int g^{(4)}(\mathbf{r}_1, \mathbf{r}, \mathbf{r}_3, \mathbf{r}') d\mathbf{r}_1 d\mathbf{r}_3 \\ &= \tilde{g}^{(2)}(\mathbf{r}, \mathbf{r}') + g^{(3)}(\mathbf{r}, \mathbf{r}') + g^{(3)}(\mathbf{r}, -\mathbf{r}') + g^{(4)}(\mathbf{r}, \mathbf{r}'). \end{aligned} \quad (24)$$

where we have defined

$$\begin{aligned} \tilde{g}^{(2)}(\mathbf{r}, \mathbf{r}') &= \frac{1}{\rho^3 V} \left\langle \sum_{i_1=1}^N \sum_{i_2 \neq i_1} [\delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' - \mathbf{r}_{i_2}) \right. \\ &\quad \left. + \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' + \mathbf{r}_{i_2})] \right\rangle, \end{aligned} \quad (25)$$

$$g^{(3)}(\mathbf{r}, \mathbf{r}') \equiv \frac{1}{V} \int g^{(3)}(\mathbf{r}_1, \mathbf{r}, \mathbf{r}') d\mathbf{r}_1 \quad (26)$$

and

$$g^{(4)}(\mathbf{r}, \mathbf{r}') = \frac{N_a}{V^2} \int \int g^{(4)}(\mathbf{r}_1, \mathbf{r}, \mathbf{r}_3, \mathbf{r}') d\mathbf{r}_1 d\mathbf{r}_3. \quad (27)$$

We substitute this result into the expression for  $\Theta(r, r', \theta)$  and perform the angular integrations to obtain

$$\begin{aligned} \Theta(r, r', \theta) &= \tilde{g}^{(2)}(r, r', \theta) + \tilde{g}^{(3)}(r, r', \theta) \\ &\quad + \tilde{g}^{(3)}(r, r, \pi - \theta) + \tilde{g}^{(4)}(r, r', \theta). \end{aligned} \quad (28)$$

where

$$\tilde{g}^{(2)}(r, r', \theta) = \int \tilde{g}^{(2)}(\mathbf{r}, \mathbf{r}') \delta(\cos \theta - \mathbf{n} \cdot \mathbf{n}') d\Omega d\Omega', \quad (29)$$

$$\tilde{g}^{(3)}(r, r', \theta) = \int g^{(3)}(\mathbf{r}, \mathbf{r}') d\Omega d\phi' \quad (30)$$

and

$$\tilde{g}^{(4)}(r, r', \theta) = \int g^{(4)}(\mathbf{r}, \mathbf{r}') d\Omega d\phi', \quad (31)$$

where we have defined  $\mathbf{n} = \mathbf{r}/|\mathbf{r}|$ . The angular integrations can be performed by representing the delta functions in spherical coordinates. We choose relative coordinates for the second pair ( $\mathbf{r}_{i_{13}}$  or  $\mathbf{r}_{i_{34}}$ ) such that the first atom pair  $\mathbf{r}_{i_{12}}$  lies along the zenith direction. We then integrate over absolute orientation by integrating over the orientation of the first pair  $d\Omega$  and by integrating the relative azimuthal angle of the second pair  $d\phi'$ . For example, the angular integration of a three-body correlation term,  $g^{(3)}(\mathbf{r}, \mathbf{r}')$ , is given by

$$\begin{aligned} \tilde{g}^{(3)}(r, r', \theta') &= \int g^{(3)}(\mathbf{r}, \mathbf{r}') d\Omega d\phi' \\ &= \frac{1}{\rho^3 V} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \sum_{i_3 \neq i_1, i_2} \int \delta(\mathbf{r} - \mathbf{r}_{i_2}) \delta(\mathbf{r}' - \mathbf{r}_{i_3}) d\Omega d\phi' \right\rangle \\ &= \frac{1}{\rho^3 V} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \sum_{i_3 \neq i_1, i_2} \frac{1}{r^2} \delta(r - r_{i_2}) \int \frac{1}{\sin \theta} \delta(\theta) \delta(\phi) d\Omega \right. \\ &\quad \left. \times \frac{1}{r'^2 \sin \theta'} \delta(r' - r_{i_3}) \delta(\theta' - \theta_{i_{23}}) \int \delta(\phi') d\phi' \right\rangle \\ &= \frac{1}{\rho^3 V} \left\langle \sum_{i_1=1}^{N_a} \sum_{i_2 \neq i_1} \sum_{i_3 \neq i_1, i_2} \frac{1}{r^2 r'^2 \sin \theta'} \delta(r - r_{i_2}) \delta(r' - r_{i_3}) \delta(\theta' - \theta_{i_{23}}) \right\rangle \end{aligned} \quad (32)$$

APPENDIX B

Estimation of the required number of X-ray diffraction patterns

X-ray fluctuation diffraction has been demonstrated on 100 nm length scales, but to push to atomic length scales, we must understand how experimental parameters affect the size of the diffraction fluctuations and the feasibility of measuring them. Here we will use statistical arguments to determine the required number of patterns as a function of the other relevant experimental parameters. The arguments presented here are similar to those developed for the fluctuation diffraction of biomolecules (Kirian, 2012).

The intensity at a pixel can be modelled by

$$I(\mathbf{q}) = I_0(q) + I_f(\mathbf{q}) + \epsilon(\mathbf{q}). \quad (33)$$

where  $I_0(q)$  is the angular average of the intensity and  $I_f(\mathbf{q})$  is fluctuation of the diffraction signal from the mean. The term  $\epsilon(\mathbf{q})$  is a random variable that accounts for shot noise, which has a mean of zero. The standard deviation of  $\epsilon(\mathbf{q})$  is equal to the square root of the intensity:

$$\sigma(\epsilon) = [I_0(q)]^{1/2} = [\phi_0 N_a \bar{f}^2(q)]^{1/2} \quad (34)$$

where  $\bar{f}(q)$  is the mean atomic structure factor and  $\phi_0$  depends on experimental parameters:

$$\phi_0 = \frac{N_{\text{ph}}}{A} r_e^2 d\Omega_q, \quad (35)$$

where  $N_{\text{ph}}$  is the number of incident photons per measurement,  $A$  is the beam area,  $r_e$  is the classical electron radius and  $d\Omega_q$  is the solid angle of subtended by a pixel.

We assume that  $I_0(q)$  can be measured accurately and subtracted prior to calculating the correlation function [equation (8)]. Evaluating an intensity correlation between a pixel at  $\mathbf{q} = [q, \theta(q), \phi]$  and a pixel at  $\mathbf{q}' = [q', \theta(q'), \phi + \Delta\phi]$  by summing over sample states  $\alpha$  to obtain

$$\begin{aligned} \sum_{\alpha} \int I_{\alpha}(\mathbf{q}) I_{\alpha}(\mathbf{q}') d\phi &= \sum_{\alpha} \int I_{f,\alpha}(\mathbf{q}) I_{f,\alpha}(\mathbf{q}') d\phi \\ &+ 2 \sum_{\alpha} \int I_{f,\alpha}(\mathbf{q}) \epsilon_{\alpha}(\mathbf{q}') d\phi \\ &+ \sum_{\alpha} \int \epsilon_{\alpha}(\mathbf{q}) \epsilon_{\alpha}(\mathbf{q}') d\phi. \end{aligned} \quad (36)$$

The term  $I_f(\mathbf{q}) I_f(\mathbf{q}')$  contains contributions from both correlated pairs, which we want to measure, and uncorrelated atomic pairs that are effectively an additional source of noise. Both  $I_f(\mathbf{q}) \epsilon(\mathbf{q}')$  and  $\epsilon(\mathbf{q}) \epsilon(\mathbf{q}')$  are additional sources of noise. It turns out that  $\epsilon(\mathbf{q}) \epsilon(\mathbf{q}')$  is the dominant source of noise and to simplify this argument, this is the only source of noise we treat here. This is consistent with the analysis of Kirian *et al.* (2011) and we provide some further comment about other noise sources below.

We need to estimate the magnitude of the interference between correlated atom pairs and how it scales with the number of atoms. This was calculated directly from an MD simulation of amorphous nickel with  $41 \times 41 \times 41$  unit cells. For  $q = 0.25 \text{ \AA}^{-1}$ , we used the simulated atomic structure to directly evaluate the diagonal terms of the matrices  $B_l(q, q)$  (as defined in the *Methods* section) and estimated  $\langle I_f I_f' \rangle = \sum_{l=1}^{20} B_l(q, q)$ . We changed the number of atoms by selecting a spherical volume with a radius in the range 30–60 Å, which changed the number of atoms from 8517 up to 32457. Fig. 8 shows that  $[\langle I_f(\mathbf{q}) I_f(\mathbf{q}') \rangle]^{1/2}$  scales as the square root of the number of atoms  $N_a^{1/2}$ , while  $I_0(q)$  scales with  $N_a$ . Dividing  $[\langle I_f(\mathbf{q}) I_f(\mathbf{q}') \rangle]^{1/2}$  by  $N_a$  gives the contribution to  $I_f(\mathbf{q})$  per atom [denoted  $I_{fa}(\mathbf{q})$ ], which we can use to estimate the fluctuation signal for larger numbers of atoms than it is practical to simulate.

We then have that

$$\langle I_f I_f' \rangle = N_a \phi_0^2 I_{fa}^2(q) \quad \sigma(\epsilon \epsilon') = N_a \phi_0 |\bar{f}(q)|^2, \quad (37)$$

where  $\sigma(\epsilon \epsilon')$  is the standard deviation of the noise term. We note that  $I_f(\mathbf{q})$  is correlated over an angular and radial range that is determined by the correlation length of the sample, whereas  $\epsilon$  is uncorrelated between pixels. It is thus important to rebin the diffraction pattern before taking the correlation to maximize  $I_f(\mathbf{q})$  and reduce noise. Denoting the number of pixels in the bin by  $M_b$ , we find that  $\langle I_f I_f' \rangle$  scales as  $M_b^2$ , while  $\sigma(\epsilon \epsilon')$  scales as  $M_b$ . The correlation calculation also involves an integration over  $M_{\phi}$  angular bins that will further reduce noise. Finally the sum of the correlation over diffraction measurements  $N_{\alpha}$  must be taken into account. When all of these operations are combined, we obtain the following expressions for the signal and the dominant error term:

$$\text{signal} = M_b^2 M_{\phi} N_{\alpha} \langle I_f I_f' \rangle \quad \text{and} \quad \text{noise} = (N_{\alpha} M_{\phi})^{1/2} M_b \sigma(\epsilon \epsilon'). \quad (38)$$

We can then define a SNR,  $R$ , given by

$$R = \frac{\text{signal}}{\text{noise}} = \frac{M_b (M_{\phi})^{1/2} (N_{\alpha})^{1/2} \phi_0 I_{fa}^2(q)}{|\bar{f}(q)|^2}, \quad (39)$$

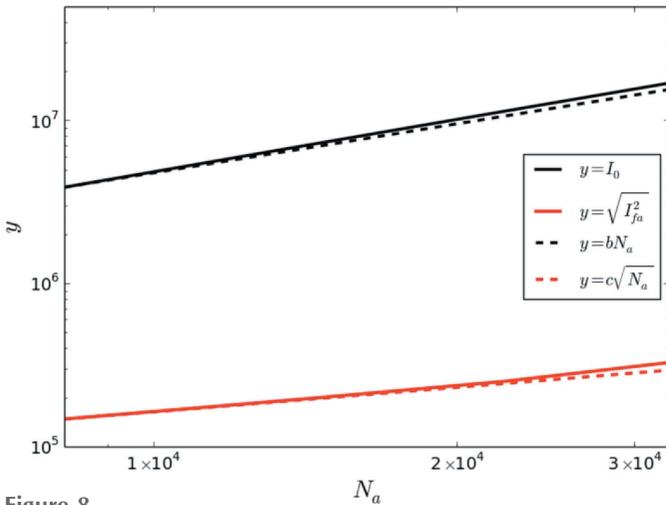


Figure 8 The scaling of  $I_0$  and  $I_f$  for nickel. The lines  $y = bN_a$  and  $y = c\sqrt{N_a}$  are given to indicate the scaling, where the parameters  $b$  and  $c$  are fixed so that the plots are equal to the simulations at  $N_a = 8517$  (corresponding to a radius of 30 Å).

which, after using equation (7) to replace  $\phi_0$ , we rearrange in terms of the number of patterns,

$$N_\alpha = \frac{R^2 |\bar{f}(q)|^4}{M_b^2 M_\phi I_{fa}^4(q)} \frac{A^2}{N_{ph}^2 r_e^4 d \Omega_q^2}. \quad (40)$$

Take the example of a  $2000 \times 2000$  pixel detector positioned to measure 1 Å resolution at the detector edge. We choose an angular bin width of  $2\pi/40$ , so  $M_\phi = 40$ . The  $q$ -width of a pixel is around  $100 \text{ nm}^{-1}$ , and we choose a radial bin of 40 pixels (sufficient to measure correlations between atom pairs up to 2.5 nm apart). Therefore, at 4 Å resolution which corresponds to 250 pixels distance from the beam center, a bin would contain close to 1560 pixels ( $= M_b$ ). For our simulation of amorphous nickel,  $I_{fa}(\mathbf{q}) = 1640$  and  $|f(q)|^2 = 500$  at 4 Å resolution. We choose beam parameters with values achievable at the Linac Coherent Light Source (LCLS) (Emma *et al.*, 2010):  $A = 100 \times 100 \text{ nm}^2$  and  $N_{ph} = 10^{12}$ . For a wavelength of 1.5 Å, the solid angle of a pixel is  $d\Omega_q = 8.1 \times 10^{-7}$ . To achieve  $R = 5$  we find that  $N_\alpha = 2.8 \times 10^4$ , which to be conservative we round up to  $10^5$ .

We note a few consequences of equation (40). The beam area  $A$  and the number of photons  $N_{ph}$  are squared, which suggests that both of these parameters are critical for measuring the fluctuations in a feasible number of measurements. Interestingly equation (40) is independent of the number of atoms, which is because the noise and signal have the same dependence on this parameter. The number of patterns required is sensitive to elemental composition because  $|\bar{f}(q)|^4 I_{fa}^{-4}(q) \propto 1/\bar{f}^4(q)$ . This means that lighter elements will be harder to measure, *i.e.* carbon would require around two orders of magnitude more measurements than nickel. The situation is more favourable for molecules, like proteins targeted by solution scattering methods, where  $I_f(\mathbf{q})$  is proportional to the square of the scattering factor of the whole molecule, because all of the atoms within a rigid molecule are correlated.

We return now to make a few remarks about the noise generated by random correlations between local environments centred on different atoms. In the case of fluctuation diffraction of biomolecules the equivalent noise source is correlations between the diffraction of different molecules, and it well known that the signal-to-noise for this case is independent of the number of molecules and the beam intensity. It was suggested by Kirian *et al.* (2011) that this result may carry over to densely packed systems by taking a local arrangement of atoms as analogous to a molecule. Some care must be taken because in a densely packed system local structures are interconnected, whilst molecules are distinct objects. Nevertheless, this is confirmed by our simulations on volumes a few times the width of the sample's structural correlation length. Larger sample volumes only increase the proportion of distinct local structures for which the assumptions of the biomolecule noise analysis are valid.

Our analysis has assumed a uniform beam profile, which is partially justified for X-ray beams because the spatial variation in the beam is typically on larger length scales that the

correlation length of amorphous materials ( $>2 \text{ nm}$ ). The overall intensity distribution can affect our signal-to-noise analysis. For example, some X-ray beams have broad 'tails' that, although weak, can spread over a large area and account for a significant fraction of the total photons. Thus, an extension of our noise analysis to non-uniform beam profiles will be very important for X-ray experiments and forms part of ongoing work. Our current understanding is that the signal-to-noise varies with the ratio of the variance of the beam intensity to the mean beam intensity, which means the bright focus of the beam is expected to make a much greater contribution to intensity correlation measurements than the beam tails.

Finally, we note that rebinning the data to an effective pixel size induces an effective traverse coherence length. When matched to the structural coherence length of the sample it effectively suppresses the noise from the interference between uncorrelated atoms larger than this length in addition to shot noise. Essentially, the interference from uncorrelated atoms produces fringes much finer than a pixel width, which have close to zero contribution to the integrated pixel intensity.

### Acknowledgements

We gratefully acknowledge the feedback on the manuscript from Harry Quiney, Les Allen and Duane Loh. AVM was supported by the Australian Research Council's Discovery Early Career Researcher Award scheme (DE140100624) and Centre of Excellence programmes.

### References

- Altarelli, M., Kurta, R. P. & Vartanyants, I. A. (2010). *Phys. Rev. B*, **82**, 104207.
- Anstis, G. R., Liu, Z. & Lake, M. (1988). *Ultramicroscopy*, **26**, 65–69.
- Aquila, A., Barty, A., Bostedt, C., Boutet, S., Carini, G., de Ponte, D., Drell, P., Doniach, S., Downing, K. H., Earnest, T., Elmlund, H., Elser, V., Guhr, M., Hajdu, J., Hastings, J., Hau-Riege, S. P., Huang, Z., Lattman, E. E., Maia, F. R. N. C., Marchesini, S., Ourmazd, A., Pellegrini, C., Santra, R., Schlichting, I., Schroer, C., Spence, J. C. H., Vartanyants, I. A., Wakatsuki, S., Weis, W. I. & Williams, G. J. (2015). *Struct. Dyn.* **2**, 041701.
- Car, R. & Parrinello, M. (1985). *Phys. Rev. Lett.* **55**, 2471–2474.
- Donatelli, J. J., Zwart, P. H. & Sethian, J. A. (2015). *Proc. Natl Acad. Sci.* **112**, 10286–10291.
- Elliott, S. R. (1983). *Physics of Amorphous Materials*. London: Longman.
- Elliott, S. R. (1991). *Nature*, **354**, 445–452.
- Emma, P., Akre, R., Arthur, J., Bionta, R. M., Bostedt, C., Bozek, J., Brachmann, A., Bucksbaum, P. H., Coffee, R., Decker, F. J., Ding, Y., Dowell, D., Edstrom, S., Fisher, A., Frisch, J., Gilevich, S., Hastings, J., Hays, G., Hering, P., Huang, Z., Iverson, R., Loos, H., Messerschmidt, M., Miahnahri, A., Moeller, S., Nuhn, H.-D., Pile, D., Ratner, D., Rzepiela, J., Schultz, D., Smith, T., Stefan, P., Tompkins, H., Turner, J., Welch, J., White, W., Wu, J., Yocky, G. & Galayda, J. N. (2010). *Nat. Photon.* **4**, 641–648.
- Fan, L., McNulty, I., Paterson, D., Treacy, M. & Gibson, J. (2005). *Nucl. Instrum. Methods Phys. Res. Sect. B*, **238**, 196–199.
- Fischer, H., Barnes, A. & Salmon, P. (2006). *Rep. Prog. Phys.* **69**, 233–299.
- Gibson, J. M., Treacy, M. M. J., Sun, T. & Zaluzec, N. J. (2010). *Phys. Rev. Lett.* **105**, 125504.
- Gibson, J. M., Treacy, M. M. J. & Voyles, P. M. (2000). *Ultramicroscopy*, **83**, 169–178.
- Kam, Z. (1977). *Macromolecules*, **10**, 927–934.

- Kam, Z., Koch, M. H. & Bordas, J. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 3559–3562.
- Kirian, R. (2012). *J. Phys. B*, **45**, 223001.
- Kirian, R. A., Schmidt, K. E., Wang, X., Doak, R. B. & Spence, J. C. H. (2011). *Phys. Rev. E*, **84**, 011921.
- Lanusse, F., Rassat, A. & Starck, J.-L. (2012). *Astron. Astrophys.* **540**, A92.
- Lipfert, J. & Doniach, S. (2007). *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307–327.
- Loh, N., Hampton, C., Martin, A., Starodub, D., Sierra, R., Barty, A., Aquila, A., Schulz, J., Lomb, L., Steinbrener, J., Shoeman, R., Kassemeyer, S., Bostedt, C., Bozek, J., Epp, S. W., Erk, B., Hartmann, R., Rolles, D., Rudenko, A., Rudek, B., Foucar, L., Kimmel, N., Weidenspointner, G., Hauser, G., Holl, P., Pedersoli, E., Liang, M., Hunter, M., Hunter, M. M., Gumprecht, L., Coppola, N., Wunderer, C., Graafsma, H., Maia, F. R., Ekeberg, T., Hantke, M., Fleckenstein, H., Hirsemann, H., Nass, K., White, T. A., Tobias, H. J., Farquar, G. R., Benner, W. H., Hau-Riege, S. P., Reich, C., Hartmann, A., Soltau, H., Marchesini, S., Bajt, S., Barthelmess, M., Bucksbaum, P., Hodgson, K. O., Strüder, L., Ullrich, J., Frank, M., Schlichting, I., Chapman, H. N. & Bogan, M. J. (2012). *Nature*, **486**, 513–517.
- Lu, J. & Szpunar, J. A. (1997). *Philos. Mag. A*, **75**, 1057–1066.
- Martin, J. D., Goettler, S. J., Fossé, N. & Iton, L. (2002). *Nature*, **419**, 381–384.
- Mauro, J. C. (2014). *Front. Mater.* **1**, 1–4.
- Plimpton, S. J. (1995). *J. Comput. Phys.* **117**, 1–19.
- Rezikyan, A., Jibben, Z. J., Rock, B. A., Zhao, G., Koeck, F. A., Nemanich, R. F. & Treacy, M. M. (2015). *Microsc. Microanal.* **21**, 1455–1474.
- Sakdinawat, A. & Attwood, D. (2010). *Nature Photon.* **4**, 840–848.
- Saldin, D. K., Poon, H.-C., Schwander, P., Uddin, M. & Schmidt, M. (2011). *Opt. Express*, **19**, 17318–17335.
- Saldin, D., Shneerson, V., Fung, R. & Ourmazd, A. (2009). *J. Phys. Condens. Matter*, **21**, 134014.
- Salmon, P. S., Martin, R. A., Mason, P. E. & Cuello, G. J. (2005). *Nature*, **435**, 75–78.
- Sellberg, J. A., Huang, C., McQueen, T. A., Loh, N. D., Laksmono, H., Schlesinger, D., Sierra, R. G., Nordlund, D., Hampton, C. Y., Starodub, D., DePonte, D. P., Beye, M., Chen, C., Martin, A. V., Barty, A., Wikfeldt, K. T., Weiss, T. M., Caronna, C., Feldkamp, J., Skinner, L. B., Seibert, M. M., Messerschmidt, M., Williams, G. J., Boutet, S., Pettersson, L. G. M., Bogan, M. J. & Nilsson, A. (2014). *Nature*, **510**, 381–384.
- Sheng, H. W., Luo, W. K., Alamgir, F. M., Bai, J. M. & Ma, E. (2006). *Nature*, **439**, 419–425.
- Starodub, D., Aquila, A., Bajt, S., Barthelmess, M., Barty, A., Bostedt, C., Bozek, J., Coppola, N., Doak, R., Epp, S., Erk, B., Foucar, L., Gumprecht, L., Hampton, C., Hartmann, A., Hartmann, R., Holl, P., Kassemeyer, S., Kimmel, N., Laksmono, H., Liang, M., Loh, N., Lomb, L., Martin, A., Nass, K., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schulz, J., Shoeman, R., Sierra, R., Soltau, H., Steinbrener, J., Stellato, F., Stern, S., Weidenspointner, G., Frank, M., Ullrich, J., Strüder, L., Schlichting, I., Chapman, H., Spence, J. & Bogan, M. (2012). *Nat. Commun.* **3**, 1276.
- Steinhardt, P. J., Nelson, D. R. & Ronchetti, M. (1983). *Phys. Rev. B*, **28**, 784–805.
- Treacy, M. M. J. & Borisenko, K. B. (2012). *Science*, **335**, 950–953.
- Treacy, M. M. J., Gibson, J. M., Fan, L., Paterson, D. J. & McNulty, I. (2005). *Rep. Prog. Phys.* **68**, 2899–2944.
- Waasmaier, D. & Kirfel, A. (1995). *Acta Cryst.* **A51**, 416–431.
- Wochner, P., Gutt, C., Autenrieth, T., Demmer, T., Bugaev, V., Ortiz, A. D., Duri, A., Zontone, F., Grübel, G. & Dosch, H. (2009). *Proc. Natl Acad. Sci.* **106**, 11511–11514.
- Wu, Z. W., Li, M. Z., Wang, W. H. & Liu, K. X. (2015). *Nat. Commun.* **6**, 6035.