

# Unifying the concepts of scattering and structure factor in ordered and disordered samples

Dingning Li and Kai Zhang\*

Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, Jiangsu, 215300, People's Republic of China. \*Correspondence e-mail: kai.zhang@dukekunshan.edu.cn

Received 3 July 2020

Accepted 19 February 2021

Edited by G. J. McIntyre, Australian Nuclear Science and Technology Organisation, Lucas Heights, Australia

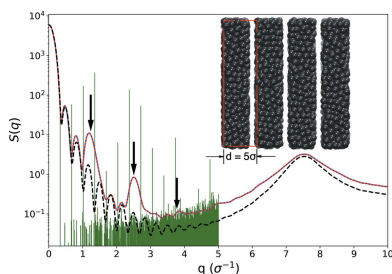
**Keywords:** scattering; structure factors; computation; GPU code.

Scattering methods are widely used in many research areas to analyze and resolve material structures. Given its importance, a large number of textbooks are devoted to this topic. However, technical details in experiments and disconnection between explanations from different perspectives often confuse and frustrate beginner students and researchers. To create an effective learning path, the core concepts of scattering and structure factor are reviewed in this article in a self-contained way. Classical examples of scattering photography and intensity scanning are calculated. Sample CPU and GPU codes are provided to facilitate the understanding and application of these methods.

## 1. Introduction

Scattering methods, using a source of photons, electrons, X-rays, neutrons *etc.*, are powerful tools to examine microscopic structural (Powles, 1973) and dynamical (Goldburg, 1999) properties of matter; they have been successfully applied to study subatomic particles (Xiong *et al.*, 2019), crystals (Azaroff, 1968), liquids (Head-Gordon & Hura, 2002), glasses (Sette *et al.*, 1998), surfactants (Hayter & Penfold, 1983), biomolecules (Kendrew, 1961; Ashkar *et al.*, 2018) and polymers (Roe, 2000). The rule of thumb here is that the wavelength  $\lambda$  of the radiation should be comparable to the length scale of the structure to be observed. To detect ordering over a range much longer than  $\lambda$ , methods like small-angle scattering are needed (Chu & Hsiao, 2001). Another important consideration is the contrast between scattering signals from different elements due to underlying physical mechanisms. Therefore, neutron scattering is often preferred for soft-matter systems, despite having lower accessibility than X-rays. In addition, techniques like resonant soft-X-ray scattering can be used to provide enhanced resolution (Fink *et al.*, 2013; Liu *et al.*, 2016). Compared with real-space microscopy techniques, reciprocal-space probes like scattering methods are good at picking up periodic patterns and revealing three-dimensional (3D) structures as a whole by penetrating deeply into the sample (Mukherjee *et al.*, 2017).

Given the richness of material structures, a variety of experimental methods have been developed during the past century, with the scattering being hard (high energy) or soft (low energy), monochromatic or polychromatic, and elastic or inelastic. Despite the diversity of experimental setups, they can largely be grouped into two categories based on how signals are collected and interpreted. The first category is photography of ordered samples, where they are recorded as spotted scattering signals on a two-dimensional (2D) film (McIntyre, 2015). The second category is intensity scanning of scattering signals from disordered or partially ordered



samples, whose one-dimensional (1D) profile is plotted against one variable (a scalar) that characterizes the existence of periodicities in the system (Hura *et al.*, 2009). In both types, the quantitative measurement of the signal is the scattering intensity  $I(\mathbf{q})$ , or its normalized version, the structure factor  $S(\mathbf{q})$ , which is often expressed as a function of the scattering vector  $\mathbf{q}$ . There are then two central tasks of structural analysis with scattering methods:

(i) The forward problem  $\rho(\mathbf{r}) \rightarrow I(\mathbf{q})$ : given the electron-density distribution  $\rho(\mathbf{r})$  or particle positions  $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ , to predict the scattering pattern  $I(\mathbf{q})$ .

(ii) The inverse problem  $I(\mathbf{q}) \rightarrow \rho(\mathbf{r})$ : given the scattering pattern  $I(\mathbf{q})$ , to resolve the electron-density distribution  $\rho(\mathbf{r})$  or particle positions  $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ .

In this article, we only focus on the forward problem, which could still shed light on some basic structural information. Sometimes, the forward method may also be used to solve  $\rho(\mathbf{r})$  iteratively, through a trial-and-error process. That is, one keeps modifying a proposed structure  $\rho(\mathbf{r})$  until the theoretically computed  $I(\mathbf{q})$  matches the experimentally observed one. The full solution to the inverse problem is, however, challenged by the notorious ‘phase problem’ (Hauptman, 1991).

The concepts of scattering and structure factor are often discussed across different disciplines including condensed-matter physics, materials science, polymer physics, structural biology *etc.* The same idea can take different forms in different areas, causing confusion and misconceptions. Graduate or advanced undergraduate students in need of applying these concepts to their research problems can be frustrated by the convoluted experimental details covered in traditional textbooks. It is thus the purpose of this article to unify the concepts of scattering and structure factor, giving junior researchers an effective pathway to quickly grasp the key ideas in this field without taking a whole course or reading an entire textbook.

To fulfill this task, we first elaborate the fundamentals about scattering (Section 2), crystallography (Section 3) and liquid-state theory (Section 4) based on the Fourier transform and reciprocal lattice. Using concrete examples, we then discuss the photography of ordered samples in Sections 5 and 6 and intensity scanning of isotropic samples in Sections 7 and 8. Relevant CPU and GPU source codes are provided online at <https://github.com/statisticalmechanics/scatter>. Finally, a brief introduction to the 2D structure factor is given in Section 9, before the conclusion in Section 10.

## 2. Scattering

### 2.1. Scattering vector

In a scattering experiment, the incident beam of wavevector  $\mathbf{k}_0$ , after hitting the sample, is deflected from its straight path by a scattering angle  $2\theta$  and becomes the diffracted beam of wavevector  $\mathbf{k}_1$  (Fig. 1). In the case of elastic<sup>1</sup> and monochro-

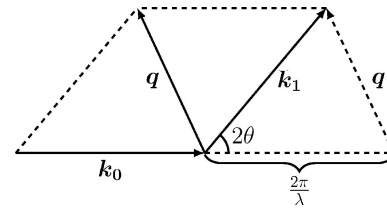


Figure 1

Scattering vector  $\mathbf{q}$  defined as the difference between the diffracted wavevector  $\mathbf{k}_1$  and the incident wavevector  $\mathbf{k}_0$ , both with magnitude  $2\pi/\lambda$  during elastic scattering.

matic scattering (of a fixed wavelength  $\lambda$ ),  $|\mathbf{k}_0| = |\mathbf{k}_1| = 2\pi/\lambda$ . The change of wavevector, called the scattering vector, is

$$\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_0 \quad (1)$$

with a magnitude

$$q = 2|\mathbf{k}_0| \sin \theta = \frac{4\pi}{\lambda} \sin \theta. \quad (2)$$

Let  $\mathbf{s}_0 = \mathbf{k}_0/|\mathbf{k}_0| = \mathbf{k}_0\lambda/(2\pi)$  and  $\mathbf{s}_1 = \mathbf{k}_1/|\mathbf{k}_1| = \mathbf{k}_1\lambda/(2\pi)$  be the unit vectors of the incident and diffracted beam, respectively; then the scattering vector can also be written as

$$\mathbf{q} = \frac{2\pi}{\lambda}(\mathbf{s}_1 - \mathbf{s}_0). \quad (3)$$

### 2.2. Scattering intensity

When a detection screen is placed behind the sample in the path of  $\mathbf{k}_1$ , the diffracted beam may be detected. The strength of such signals is quantified by the scattering intensity  $I(\mathbf{q})$  of the ray, which changes with  $\mathbf{k}_1$  or, equivalently, with  $\mathbf{q}$ . The scattering pattern, or the distribution of  $I(\mathbf{q})$  on the screen, is determined by the structural features of the sample, for instance, the electron-density distribution  $\rho(\mathbf{r})$  in the case of X-ray scattering by atoms.

Both the incident and the diffracted rays can be viewed as plane waves of the form  $\psi_{\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r} | \mathbf{k} \rangle \propto \exp(i\mathbf{k} \cdot \mathbf{r})$ . According to Fermi’s golden rule, the scattering intensity  $I(\mathbf{q})$  is proportional to the square of the transition probability amplitude from state  $\psi_{\mathbf{k}_0}(\mathbf{r})$  to state  $\psi_{\mathbf{k}_1}(\mathbf{r})$ , after interacting with the overall scattering potential  $\rho(\mathbf{r})$ . That is,

$$\begin{aligned} I(\mathbf{q}) &\propto |\langle \mathbf{k}_0 | \rho(\mathbf{r}) | \mathbf{k}_1 \rangle|^2 = \left| \int d\mathbf{r} \psi_{\mathbf{k}_0}^*(\mathbf{r}) \rho(\mathbf{r}) \psi_{\mathbf{k}_1}(\mathbf{r}) \right|^2 \\ &\propto \left| \int d\mathbf{r} \exp(-i\mathbf{k}_0 \cdot \mathbf{r}) \rho(\mathbf{r}) \exp(i\mathbf{k}_1 \cdot \mathbf{r}) \right|^2 \\ &= \left| \int d\mathbf{r} \rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) \right|^2. \end{aligned} \quad (4)$$

Neglecting the coefficient of proportionality, one can write

$$I(\mathbf{q}) = \hat{\rho}_{\mathbf{q}} \hat{\rho}_{-\mathbf{q}}, \quad (5a)$$

where

$$\hat{\rho}_{\mathbf{q}} = \int d\mathbf{r} \rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) \quad (5b)$$

is the Fourier transform of the density distribution and  $\hat{\rho}_{-\mathbf{q}}$  is its complex conjugate (Appendix A).

Unless  $\rho(\mathbf{r})$  has a symmetry center,  $\hat{\rho}_{\mathbf{q}}$  is generally a complex number, *i.e.*  $\hat{\rho}_{\mathbf{q}} = |\hat{\rho}_{\mathbf{q}}| \exp(i\phi_{\mathbf{q}})$ . If  $\hat{\rho}_{\mathbf{q}}$  is known exactly,  $\rho(\mathbf{r})$  can

in principle be reconstructed through the inverse Fourier transform [equation (48)] (Argos, 1977). However, in an experiment, only the scattering intensity  $I(\mathbf{q}) = |\hat{\rho}_{\mathbf{q}}|^2 \times \exp(i\phi_{\mathbf{q}}) \exp(-i\phi_{\mathbf{q}}) = |\hat{\rho}_{\mathbf{q}}|^2$  is directly measurable. This allows us to compute the magnitude of  $\hat{\rho}_{\mathbf{q}}$  by  $|\hat{\rho}_{\mathbf{q}}| = [I(\mathbf{q})]^{1/2}$ . Unfortunately, information about the phase angle  $\phi_{\mathbf{q}}$  is lost during this process, which gives rise to the ‘phase problem’ in crystallography. Special techniques (Hauptman, 1991; Harrison, 1993; Taylor, 2003) have been developed to determine  $\phi_{\mathbf{q}}$ , which are beyond the scope of this article.

In a system of  $N$  atoms or particles at positions  $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  inside a region of volume  $V$ , the density distribution consists of the contributions from each particle  $i$  with a scattering potential  $f_i(\mathbf{r} - \mathbf{r}_i)$  ( $i = 1, 2, \dots, N$ ), *i.e.*

$$\rho(\mathbf{r}) = \sum_{i=1}^N f_i(\mathbf{r} - \mathbf{r}_i) = \sum_{i=1}^N f_i(\mathbf{R}_i), \quad (\mathbf{R}_i \equiv \mathbf{r} - \mathbf{r}_i). \quad (6)$$

In this case

$$\begin{aligned} \hat{\rho}_{\mathbf{q}} &= \int_V d\mathbf{r} \sum_{i=1}^N f_i(\mathbf{r} - \mathbf{r}_i) \exp(i\mathbf{q} \cdot \mathbf{r}) \\ &= \sum_{i=1}^N \int_V d\mathbf{R}_i f_i(\mathbf{R}_i) \exp(i\mathbf{q} \cdot \mathbf{R}_i) \exp(i\mathbf{q} \cdot \mathbf{r}_i) \\ &= \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_i), \end{aligned} \quad (7)$$

where

$$\hat{f}_i(\mathbf{q}) = \int_V d\mathbf{r} f_i(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) \quad (8)$$

is the atomic form factor, or scattering factor, of particle  $i$ .

If the scattering potential of each particle  $f_i(\mathbf{r} - \mathbf{r}_i)$  is symmetric about  $\mathbf{r}_i$ , which should be true for atoms and most particles,  $\hat{f}_i(\mathbf{q})$  is real and even, *i.e.* its complex conjugate  $\hat{f}_i^*(\mathbf{q}) = \hat{f}_i(-\mathbf{q}) = \hat{f}_i(\mathbf{q})$  (Appendix A). Under this circumstance, the scattering intensity

$$\begin{aligned} I(\mathbf{q}) &= \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_i) \sum_{j=1}^N \hat{f}_j(-\mathbf{q}) \exp(-i\mathbf{q} \cdot \mathbf{r}_j) \\ &= \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_i) \sum_{j=1}^N \hat{f}_j(\mathbf{q}) \exp(-i\mathbf{q} \cdot \mathbf{r}_j) \\ &= \left| \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \cos(\mathbf{q} \cdot \mathbf{r}_i) \right|^2 + \left| \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \sin(\mathbf{q} \cdot \mathbf{r}_i) \right|^2 \end{aligned} \quad (9a)$$

or, equivalently,

$$\begin{aligned} I(\mathbf{q}) &= \sum_{i=1}^N \sum_{j=1}^N \hat{f}_i(\mathbf{q}) \hat{f}_j(\mathbf{q}) \exp[i\mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_j)] \\ &= \sum_{i=1}^N \sum_{j=1}^N \hat{f}_i(\mathbf{q}) \hat{f}_j(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_{ij}) \\ &= \sum_{i=1}^N \sum_{j=1}^N \hat{f}_i(\mathbf{q}) \hat{f}_j(\mathbf{q}) \cos(\mathbf{q} \cdot \mathbf{r}_{ij}) \\ &= \sum_{i=1}^N \hat{f}_i^2(\mathbf{q}) + \sum_{i=1}^N \sum_{j \neq i}^N \hat{f}_i(\mathbf{q}) \hat{f}_j(\mathbf{q}) \cos(\mathbf{q} \cdot \mathbf{r}_{ij}). \end{aligned} \quad (9b)$$

Equations (9a) and (9b) are mathematically equivalent because  $\cos(\mathbf{q} \cdot \mathbf{r}_{ij}) = \cos(\mathbf{q} \cdot \mathbf{r}_i - \mathbf{q} \cdot \mathbf{r}_j) = \cos(\mathbf{q} \cdot \mathbf{r}_i) \cos(\mathbf{q} \cdot \mathbf{r}_j) +$

$\sin(\mathbf{q} \cdot \mathbf{r}_i) \sin(\mathbf{q} \cdot \mathbf{r}_j)$ . However, in numerical computation of  $I(\mathbf{q})$  at a given  $\mathbf{q}$ , equation (9a) has a lower cost with a computational complexity  $\mathcal{O}(N)$ , while equation (9b) is of complexity  $\mathcal{O}(N^2)$ . Nevertheless, when there is an appropriate symmetry in the system, the expression  $\mathbf{r}_i$  in equation (9b) allows it to be further simplified and thus to become more computationally efficient, as will be discussed in later sections.

### 2.3. Atomic form factor

For realistic scattering potentials, the atomic form factor  $\hat{f}_i(\mathbf{q})$  changes with the direction and magnitude of the scattering vector  $\mathbf{q}$ , and thus often drops as the scattering angle  $\theta$  increases (Fig. 2). If, however, the scattering potential is spherically symmetric, *i.e.*  $f_i(\mathbf{r}) = f_i(r)$ , we can write

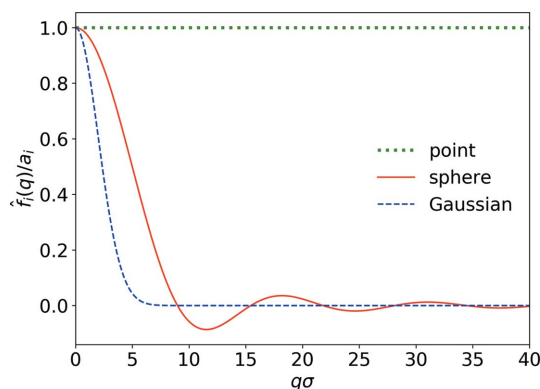
$$\begin{aligned} \hat{f}_i(\mathbf{q}) &= \hat{f}_i(q) = 2\pi \int dr r^2 f_i(r) \int_0^\pi d\theta \sin \theta \exp(iqr \cos \theta) \\ &= 2\pi \int dr r^2 f_i(r) \frac{2 \sin(qr)}{qr} \\ &= 4\pi \int dr r^2 f_i(r) \frac{\sin(qr)}{qr} \quad (q \neq 0). \end{aligned} \quad (10)$$

It is useful to consider the three simple spherically symmetric scattering potentials listed below (Fig. 2).

(i)  $f_i(\mathbf{r} - \mathbf{r}_i) = a_i \delta(\mathbf{r} - \mathbf{r}_i)$ , the scattering by each atom is idealized as from a sizeless point at the atomic center. This model can be mapped onto the physical scenario of nuclear scattering or the abstract scenario of point-mass scattering. The scattering strength  $a_i$  of atom  $i$  generally has different values for different elements, and has also been called the atomic scattering factor, because here

$$\hat{f}_i(\mathbf{q}) = a_i. \quad (11)$$

The electron-density distribution is then  $\rho(\mathbf{r}) = \sum_{i=1}^N a_i \delta(\mathbf{r} - \mathbf{r}_i)$ , which, in the case of  $a_i = 1$ , becomes the particle density distribution  $\rho(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)$ .



**Figure 2** Atomic form factor  $\hat{f}_i(q)$  of a sizeless point [green dotted line, equation (11)], a uniform sphere [red solid line, equation (13)] and a Gaussian scattering center [blue dashed line, equation (15)] as a function of  $q$ .

(ii)  $f_i(\mathbf{r} - \mathbf{r}_i)$  is homogeneous and bounded within a sphere of radius  $\sigma/2$ ,

$$f_i(\mathbf{r} - \mathbf{r}_i) = \begin{cases} a_i/(\pi\sigma^3/6) & \text{for } |\mathbf{r} - \mathbf{r}_i| \leq \sigma/2 \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

and

$$\begin{aligned} \hat{f}_i(q) &= \frac{4\pi a_i/(\pi\sigma^3/6)}{q^3} [\sin(q\sigma/2) - q\sigma/2 \cos(q\sigma/2)] \\ &= \frac{3a_i}{(q\sigma/2)^3} [\sin(q\sigma/2) - q\sigma/2 \cos(q\sigma/2)]. \end{aligned} \quad (13)$$

(iii)  $f_i(\mathbf{r} - \mathbf{r}_i)$  is Gaussian-like with standard deviation  $\sigma/2$ ,

$$\begin{aligned} f_i(\mathbf{r} - \mathbf{r}_i) &= a_i \left\{ \frac{1}{[2\pi(\sigma/2)^2]^{1/2}} \right\}^3 \exp\left[-\frac{|\mathbf{r} - \mathbf{r}_i|^2}{2(\sigma/2)^2}\right] \\ &= \frac{a_i}{\sigma^3(\pi/2)^{3/2}} \exp\left(-\frac{2R_i^2}{\sigma^2}\right) \quad (R_i = |\mathbf{r} - \mathbf{r}_i|) \end{aligned} \quad (14)$$

and

$$\hat{f}_i(q) = a_i \exp\left(-\frac{\sigma^2 q^2}{8}\right). \quad (15)$$

In all numerical results shown below, we will assume  $\hat{f}_i(\mathbf{q}) = 1$ , *i.e.* point scattering, for all particles.

### 3. Crystallography

We now review concepts and theories about scattering methods used for crystal samples. The earlier theory of von Laue (McQuarrie & Simon, 1997) that considers diffraction of parallel beams by arrays of atoms is skipped here. Instead, we apply the more intuitive Bragg's law which envisages crystallographic planes as reflective mirrors to understand the principle, although there is no such reflection in the physical sense.

#### 3.1. Bragg's law

For an incident ray of wavelength  $\lambda$  to generate a strong constructive scattering signal by a family of crystallographic planes ( $hkl$ ) of interplanar spacing  $d_{hkl}$  (Appendix B), the scattering angle  $2\theta$  needs to obey Bragg's law (Bragg, 1968) (Fig. 3):

$$n\lambda = 2d_{hkl} \sin \theta, \quad n = 1, 2, 3, \dots \quad (16)$$

This is because the path difference of the two scattering rays 'reflected' by two neighboring lattice planes is

$$\begin{aligned} l(\cos \alpha + \cos \beta) &= l[\cos \alpha + \cos(\pi - 2\theta - \alpha)] \\ &= l[\cos \alpha - \cos(2\theta + \alpha)] \\ &= l \sin(\theta + \alpha) \sin \theta \\ &= d_{hkl} \sin \theta. \end{aligned} \quad (17)$$

<sup>2</sup> Throughout the paper, we use  $\sigma$  as the unit of length and  $1/\sigma$  as the unit of wavevector. We choose  $\sigma$  to be the particle diameter, which can be mapped onto the length scale of Å for atomic systems, nm for nano-systems and  $\mu\text{m}$  for colloidal systems.

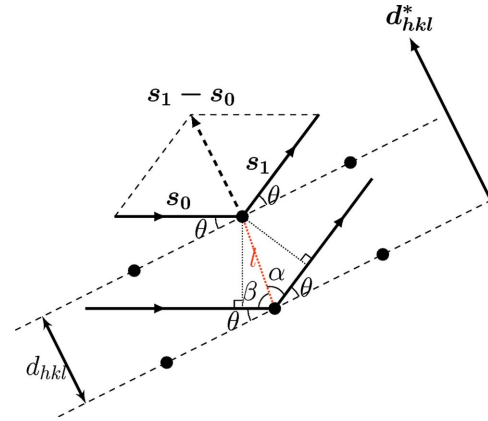


Figure 3

The scattering paths of two rays diffracted by two layers of ordered particles (black dots) with interplanar distance  $d_{hkl}$  and scattering angle  $2\theta$ .  $\mathbf{s}_0$  and  $\mathbf{s}_1$  are unit vectors of the incident and diffracted rays, respectively. When Bragg's law is satisfied, the scattering vector is parallel to the normal vector  $\mathbf{d}_{hkl}^*$  of the lattice planes.

The rescaled scattering vector  $\mathbf{s}_1 - \mathbf{s}_0 = (\lambda/2\pi)\mathbf{q}$  (of length  $2 \sin \theta$ ) is parallel to the normal vector, or reciprocal vector  $\mathbf{d}_{hkl}^*$  (of length  $1/d_{hkl}$ ), of the lattice planes ( $hkl$ ). Thus, it is sometimes convenient to express Bragg's law in a vector form, for instance, for the primary  $n = 1$  scattering, as

$$\frac{\mathbf{s}_1 - \mathbf{s}_0}{\lambda} = \mathbf{d}_{hkl}^*. \quad (18)$$

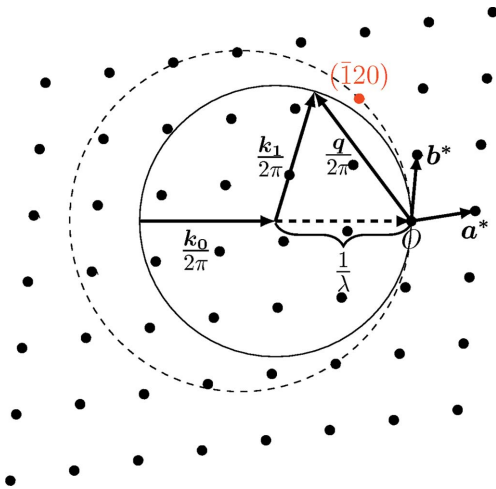
Using equation (3), the necessary condition to receive a strong signal for scattering vector  $\mathbf{q}$  in crystals is thus

$$\mathbf{q} = 2\pi\mathbf{d}_{hkl}^*. \quad (19)$$

#### 3.2. The Ewald construction

Bragg's law needs to be satisfied to have a strong scattering signal in the direction of  $\mathbf{s}_1$ . However, this does not mean that, given an arbitrary experimental setup, Bragg's law is guaranteed to be satisfied somewhere. In particular, if a monochromatic incident beam (fixed  $\lambda$ ) is directed onto a single crystal at an arbitrarily fixed position (fixed  $\theta$  and  $d_{hkl}$ ), it is possible that none of the lattice planes will be able to produce a strong scattering signal. If this happens, either  $\lambda$  (polychromatic) or  $\theta$  (rotate the sample or use polycrystals) has to be tuned to satisfy equations (16), (18) and (19).

An alternative view to check that Bragg's law is satisfied is to use Ewald's sphere in the reciprocal space (Hammond, 2001; Barbour, 2018). Here, each point at vector  $\mathbf{d}_{hkl}^*$  represents a family of parallel planes ( $hkl$ ) in the direct space. When the orientation of the crystal sample is fixed, the relative positions of the incident beam and reciprocal-lattice points are also fixed. One can align the endpoint of the incident wavevector  $\mathbf{k}_0$  (in practice  $\mathbf{k}_0/2\pi$ ) with the origin  $O$  of the reciprocal lattice and then draw a sphere of radius  $1/\lambda$ . The center of the sphere is found by moving from point  $O$  by a vector displacement  $-\mathbf{k}_0/2\pi$  (Fig. 4). It can be seen that the endpoint of the scattering vector  $\mathbf{q}$ , normalized by  $2\pi$ , falls on the



**Figure 4**  
The Ewald construction: Ewald’s sphere of radius  $1/\lambda$  (solid circle) depicts all possible scattering wavevectors  $\mathbf{q}$  under the current setup. Lattice planes with Miller indices  $(hkl)$  are represented by points on the reciprocal lattice (black dots). For wavelength  $\lambda$ , no reciprocal-lattice point is on Ewald’s sphere implying that no scattering signal will be generated at any scattering angle. If the wavelength is appropriately tuned, some reciprocal-lattice points can fall on the new Ewald sphere (dashed circle) to satisfy Bragg’s law, for instance,  $(\bar{1}20)$ .

surface of this Ewald sphere. According to the vector form of Bragg’s law [equation (19)], scattering from certain lattice planes  $(hkl)$  is possible only when the corresponding reciprocal-vector point  $\mathbf{d}_{hkl}^*$  falls on the surface of the Ewald sphere. If the wavelength and crystal orientation are not appropriately chosen, this condition may not be met at all and no scattering signal is generated by the sample.

### 3.3. Crystal structure factor $F_{hkl}$

Bragg’s law is actually the necessary (but not sufficient) condition to have a strong scattering signal. Even if Bragg’s law is obeyed by lattice planes  $(hkl)$ , it is still possible that the scattering signal cancels due to special lattice symmetries. In fact, when Bragg’s law is presented as in Fig. 3, a simple square or oblique lattice structure is often used, which misses the complexity in other 3D lattices. Generally, not every family of lattice planes  $(hkl)$  can produce a constructive scattering.

Because the density distribution  $\rho(\mathbf{r})$  is periodic in crystals, one only needs to consider the particle distribution within one unit cell. If each unit cell has a volume  $V_{\text{cell}}$  and  $m$  atoms, then

$$\hat{\rho}_{\mathbf{q}} = \frac{N}{m} \int_{V_{\text{cell}}} \mathbf{dr} \rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}), \quad (20)$$

where  $N/m$  is the number of unit cells in the  $N$ -particle system. In crystallography, it is customary to define  $\hat{\rho}_{\mathbf{q}}$  per unit cell as the structure factor,

$$F_{\mathbf{q}} = \int_{V_{\text{cell}}} \mathbf{dr} \rho(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}). \quad (21)$$

For crystals, only  $\mathbf{qs}$  satisfying Bragg’s law [equation (19)] can possibly generate a large  $\hat{\rho}_{\mathbf{q}}$  or  $F_{\mathbf{q}}$ . Therefore, we only need to

consider  $\mathbf{qs}$  of the form  $\mathbf{q} = 2\pi\mathbf{d}_{hkl}^* = 2\pi(h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)$ , where  $\mathbf{d}_{hkl}^*$  represents a family of lattice planes  $(hkl)$  of spacing  $d_{hkl} = 1/|\mathbf{d}_{hkl}^*|$  (Appendix B). The associated structure factor can thus be denoted as  $F_{hkl}$ :

$$F_{hkl} = \int_{V_{\text{cell}}} \mathbf{dr} \rho(x, y, z) \exp[2\pi i(hx + ky + lz)]. \quad (22)$$

Inversely, the density distribution within each unit cell is

$$\rho(\mathbf{r}) = \frac{1}{V_{\text{cell}}} \sum_{hkl} F_{hkl} \exp[-2\pi i(hx + ky + lz)]. \quad (23)$$

For point-like scattering centers,  $\rho(\mathbf{r}) = \sum_{i=1}^m a_i \delta(\mathbf{r} - \mathbf{r}_i)$  and equation (22) reduces to

$$F_{hkl} = \sum_{i=1}^m a_i \exp[2\pi i(hx_i + ky_i + lz_i)], \quad (24)$$

after substituting equation (11) and following the steps in equation (7), where  $(x_i, y_i, z_i)$  are the coordinates of the  $m$  atoms inside one unit cell and are expressed as fractions of lattice vectors. The strength of  $F_{hkl}$  by planes  $(hkl)$  is the vector sum of each term  $a_i \exp[2\pi i(hx_i + ky_i + lz_i)]$  in equation (24), where the phase angle  $hx_i + ky_i + lz_i$  defines the direction of each vector. For typical crystal lattices of point-like atoms of the same type ( $a_i = a$ ),  $F_{hkl}$  can be easily computed.

(i) Simple cubic (SC)

$m = 1$  and  $(x_1, y_1, z_1) = (0, 0, 0)$ :

$$F_{hkl}^{\text{SC}} = a \exp[2\pi i(h0 + k0 + l0)] = a \quad (25)$$

for any  $h, k, l$ .

(ii) Body-centered cubic (b.c.c.)

$m = 2$ ,  $(x_1, y_1, z_1) = (0, 0, 0)$  and  $(x_2, y_2, z_2) = (1/2, 1/2, 1/2)$ :

$$\begin{aligned} F_{hkl}^{\text{b.c.c.}} &= a \exp[2\pi i(h0 + k0 + l0)] \\ &\quad + a \exp[2\pi i(h\frac{1}{2} + k\frac{1}{2} + l\frac{1}{2})] \\ &= a + a \exp[\pi i(h + k + l)]. \end{aligned} \quad (26)$$

(iii) Face-centered cubic (f.c.c.)

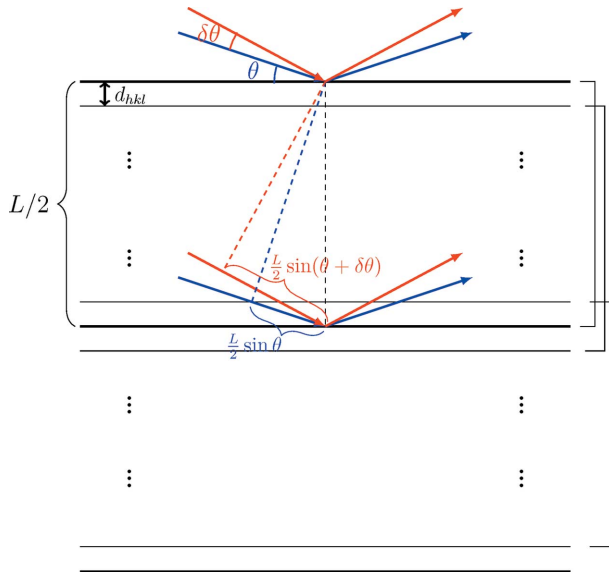
$m = 4$ ,  $(x_1, y_1, z_1) = (0, 0, 0)$ ,  $(x_2, y_2, z_2) = (1/2, 1/2, 0)$ ,  $(x_3, y_3, z_3) = (0, 1/2, 1/2)$  and  $(x_4, y_4, z_4) = (1/2, 0, 1/2)$ :

$$\begin{aligned} F_{hkl}^{\text{f.c.c.}} &= a \exp[2\pi i(h0 + k0 + l0)] \\ &\quad + a \exp[2\pi i(h\frac{1}{2} + k\frac{1}{2} + l0)] \\ &\quad + a \exp[2\pi i(h0 + k\frac{1}{2} + l\frac{1}{2})] \\ &\quad + a \exp[2\pi i(h\frac{1}{2} + k0 + l\frac{1}{2})] \\ &= a + a \exp[\pi i(h + k)] + a \exp[\pi i(k + l)] \\ &\quad + a \exp[\pi i(h + l)]. \end{aligned} \quad (27)$$

The  $F_{hkl}$  of b.c.c. and f.c.c. lattices completely vanishes for certain  $h, k, l$ . The resulting reflection Miller indices should successively be (110), (200), (211), (220), (310), (222) ... for b.c.c. and (111), (200), (220), (311), (222), (400) ... for f.c.c. crystals.

### 3.4. Finite-size crystals and Bragg peak broadening

When Bragg’s law is satisfied by wavelength  $\lambda$  at an incident angle  $\theta$ , a small deviation  $\delta\theta$  from  $\theta$  only slightly changes the path difference between two rays reflected by a pair of



**Figure 5**  
Illustration of the Scherrer equation. In a finite-size crystal of thickness  $L$ , the path difference between two rays reflected by a pair of planes that are  $L/2$  apart sets the limit of the signal broadening  $2\delta\theta$ .

neighboring planes (of spacing  $d_{hkl}$ ), which still add constructively. If we consider two reflection planes that are  $2d_{hkl}, 3d_{hkl}, \dots$  apart, the change in path difference due to  $\delta\theta$  accumulates and, at large enough spacing, becomes  $\lambda/2$  such that the two rays completely cancel. For a beam reflected by a crystallographic plane in a large crystal sample, it is always possible to find another remote plane whose reflected beam interferes destructively, even for very small  $\delta\theta$ . Therefore, when other broadening effects are excluded, diffraction signals in large samples at fixed  $\lambda$ , if there are any, should in principle be of infinitely small size (in terms of the range of  $\theta$ ).

For small crystal samples, it is possible that the change in path difference due to  $\delta\theta$  is much less than  $\lambda/2$  such that Bragg's law is still approximately satisfied at  $\theta + \delta\theta$  and the diffraction signal is broadened by an amount  $\sim\delta\theta$ . The quantitative relationship between the broadening  $2\delta\theta$  of the signal and the linear dimension  $L$  of a finite-size crystal can be found by considering all pairs of planes that are  $L/2$  apart. When  $\theta$  changes to  $\theta + \delta\theta$ , the path difference for such a pair of planes increases by  $2(L/2)[\sin(\theta + \delta\theta) - \sin\theta] = L \cos\theta\delta\theta$  (Fig. 5). The diffraction signal broadens until destructive interference occurs at  $\lambda/2 = L \cos\theta\delta\theta$ , which gives the Scherrer equation:

$$2\delta\theta = \frac{\lambda}{L \cos\theta} = \frac{2 \tan\theta}{L/d_{hkl}}. \quad (28)$$

Thus diffraction signals tend to be larger in smaller systems.

#### 4. Liquid-state theory

According to liquid-state theory, a static structure factor  $S(\mathbf{q})$  can be used to address short-range order (Thomas & Gingrich, 1941; Rahman *et al.*, 1962) and the glass transition (Janssen, 2018) in amorphous/liquid samples (Fischer *et al.*, 2006) and

more generally in nano-structured or other structurally disordered systems (Billinge, 2019). In an  $N$ -particle system, it is defined as

$$S(\mathbf{q}) = \frac{1}{\sum_{i=1}^N \hat{f}_i^2(\mathbf{q})} \langle \hat{\rho}_{\mathbf{q}} \hat{\rho}_{-\mathbf{q}} \rangle = \frac{1}{\sum_{i=1}^N \hat{f}_i^2(\mathbf{q})} I(\mathbf{q}), \quad (29)$$

where the ensemble average  $\langle \dots \rangle$  is usually taken over configurations at thermal equilibrium (Hansen & McDonald, 2013). Practically, this ensemble average results from a sum over all the different coherence volumes in the sample, after being Fourier transformed, giving a real-space representation of the sample's ensemble-averaged instantaneous local structure.

If scattering centers are point like, *i.e.*  $\rho(\mathbf{r}) = \sum_{i=1}^N a_i \delta(\mathbf{r} - \mathbf{r}_i)$ , then  $\hat{\rho}_{\mathbf{q}} = \sum_{i=1}^N a_i \exp(i\mathbf{q} \cdot \mathbf{r}_i)$  and

$$S(\mathbf{q}) = \frac{1}{\sum_{i=1}^N a_i^2} \left\langle \left| \sum_{i=1}^N a_i \cos(\mathbf{q} \cdot \mathbf{r}_i) \right|^2 + \left| \sum_{i=1}^N a_i \sin(\mathbf{q} \cdot \mathbf{r}_i) \right|^2 \right\rangle. \quad (30)$$

For monodisperse systems ( $a_i$  is the same for all particles),  $S(0) = N$ .

In the case of  $a_i = 1$ ,  $S(\mathbf{q})$  is related to the radial distribution function  $g(\mathbf{r})$  or the pair correlation function  $h(\mathbf{r}) = g(\mathbf{r}) - 1$  by

$$\begin{aligned} S(\mathbf{q}) &= 1 + \rho_0 \int d\mathbf{r} [g(\mathbf{r}) - 1] \exp(i\mathbf{q} \cdot \mathbf{r}_i) + \rho_0 \int d\mathbf{r} \exp(i\mathbf{q} \cdot \mathbf{r}_i) \\ &= 1 + \rho_0 \int d\mathbf{r} [g(\mathbf{r}) - 1] \exp(i\mathbf{q} \cdot \mathbf{r}_i) + \rho_0 (2\pi)^3 \delta_{\mathbf{D}}(\mathbf{q}) \\ &= 1 + \rho_0 \int_V d\mathbf{r} [g(\mathbf{r}) - 1] \exp(i\mathbf{q} \cdot \mathbf{r}_i) + \rho_0 V \delta_{\mathbf{q},0} \quad (\text{finite } V) \\ &= 1 + \rho_0 \int_V d\mathbf{r} h(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}_i) + N \delta_{\mathbf{q},0} \\ &= 1 + \rho_0 \hat{h}_{\mathbf{q}} + N \delta_{\mathbf{q},0}, \end{aligned} \quad (31)$$

where the global number density  $\rho_0 = N/V$  and the Fourier transform  $\hat{h}_{\mathbf{q}} = \int_V d\mathbf{r} h(\mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}_i)$ . Note that  $S(\mathbf{q})$  is singular or discontinuous at  $\mathbf{q} = 0$ , *i.e.*  $\lim_{\mathbf{q} \rightarrow 0} S(\mathbf{q}) \neq S(0) = N$ . Correspondingly,  $\lim_{\mathbf{q} \rightarrow 0} \hat{h}_{\mathbf{q}} \neq \hat{h}_0 = -1/\rho_0$ .

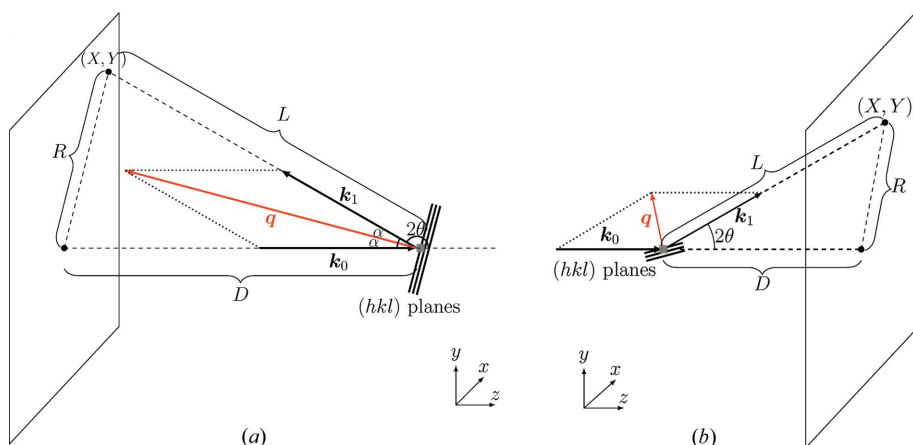
The radial distribution function can be obtained from the structure factor by the inverse Fourier transform:

$$g(\mathbf{r}) = 1 + \frac{1}{(2\pi)^3} \int_{\mathbf{q} \rightarrow 0} d\mathbf{q} \frac{S(\mathbf{q}) - 1}{\rho_0} \exp(-i\mathbf{q} \cdot \mathbf{r}_i), \quad (32)$$

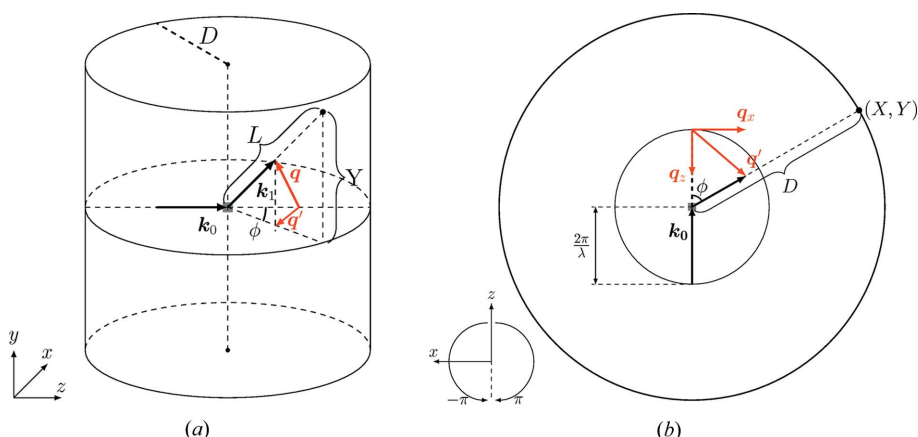
where the value  $\lim_{\mathbf{q} \rightarrow 0} S(\mathbf{q})$  should be used at  $\mathbf{q} = 0$  in the integration. When the system's structure is isotropic over the sample volume, *i.e.*  $g(\mathbf{r}) = g(r)$ , more convenient relationships can be derived (Keen, 2001):

$$S(q) = 1 + 4\pi\rho_0 \int_{0^+}^{\infty} dr [g(r) - 1] r^2 \frac{\sin(qr)}{qr}, \quad (33a)$$

$$g(r) = 1 + \frac{1}{2\pi^2} \int_{0^+}^{\infty} dq \frac{S(q) - 1}{\rho_0} q^2 \frac{\sin(qr)}{qr}, \quad (33b)$$



**Figure 6** Illustration of (a) back-reflection and (b) transmission methods. The scattering vector  $\mathbf{q}$  resulting from crystallographic planes  $(hkl)$  maps onto 2D coordinates  $(X, Y)$  on the film, which is placed at a distance  $D$  from the sample. The incident beam is along the  $z$  axis.



**Figure 7** Illustration of the cylindrical method from (a) the side view and (b) the top view. A cylindrical film is placed at a radius  $D$  around the sample.

where  $\lim_{x \rightarrow 0} (\sin x/x) = 1$  should be used in the integration. The limit value of  $S(q)$  as  $q$  approaches zero is related to the isothermal compressibility  $\kappa$  by (Barrat & Hansen, 2003)

$$\lim_{q \rightarrow 0} S(q) = \rho_0 k_B T \kappa. \quad (34)$$

## 5. Experimental setups in photography

In this section, we discuss some technical details about photography methods, which collect signals of  $I(\mathbf{q}) = I(X, Y)$  on a 2D film with coordinates  $(X, Y)$ . Three popular experimental setups are often used as described below, which map  $\mathbf{q}$  onto  $(X, Y)$  differently.

### 5.1. Back-reflection and transmission methods

In back-reflection and transmission methods, the recording film is a rectangular plane, which is placed either before (back-reflection) or after (transmission) the sample as shown in

Fig. 6. In both methods, it can be seen that the ratio  $q_x/q_y$  equals  $X/Y$ . If the incident wavenumber is  $|\mathbf{k}_0| = 2\pi/\lambda$ , then

$$(q_x, q_y) = \frac{2\pi}{\lambda} \left( \frac{X}{L}, \frac{Y}{L} \right), \quad (35)$$

where  $L^2 = R^2 + D^2$  and  $R^2 = X^2 + Y^2$ . The difference lies in the  $z$  component  $q_z$ .

In the back-reflection method, because  $\alpha = (\pi/2) - \theta$  satisfies  $\cos(2\alpha) = D/L$ , it follows that

$$\begin{aligned} q_z &= -2 \frac{2\pi}{\lambda} \cos \alpha \cos \alpha \\ &= -\frac{2\pi}{\lambda} [1 + \cos(2\alpha)] \\ &= -\frac{2\pi}{\lambda} \left( 1 + \frac{D}{L} \right). \end{aligned} \quad (36)$$

Therefore,

$$(q_x, q_y, q_z) = \frac{2\pi}{\lambda} \left( \frac{X}{L}, \frac{Y}{L}, -\frac{D+L}{L} \right) \quad (\text{back-reflection}). \quad (37)$$

In contrast, in the transmission method,

$$(q_x, q_y, q_z) = \frac{2\pi}{\lambda} \left( \frac{X}{L}, \frac{Y}{L}, -\frac{L-D}{L} \right) \quad (\text{transmission}). \quad (38)$$

### 5.2. Cylindrical method

Compared with the above two setups, the cylindrical method is more informative as it collects signals from all azimuthal angles  $\phi$  (Fig. 7). In fact, a cylindrical film, which better preserves the natural shape of scattering spots, can be considered as the sum of an infinitely wide back-reflection film and an infinitely wide transmission film, on which scattering patterns farther away from the film center are more distorted.

To map  $\mathbf{q}$  onto the film, one can unfold the cylinder into a plane with coordinates  $(X, Y) = (D \sin \phi, Y)$  with the azimuthal angle  $\phi \in (-\pi, \pi)$ . The relationship is

$$(q_x, q_y, q_z) = \frac{2\pi}{\lambda} \left( \frac{D \sin \phi}{L}, \frac{Y}{L}, \frac{D \cos \phi - D}{L} \right). \quad (39)$$

## 6. Photography of single-crystalline samples

The illustration of Bragg's law using Ewald's sphere suggests two ways to make reciprocal-lattice points fall on the sphere to generate constructive scattering signals from specific crystallographic planes. One is to tune the wavelength and the other is to change the orientation of the sample. These correspond to two experimental strategies in designing photography

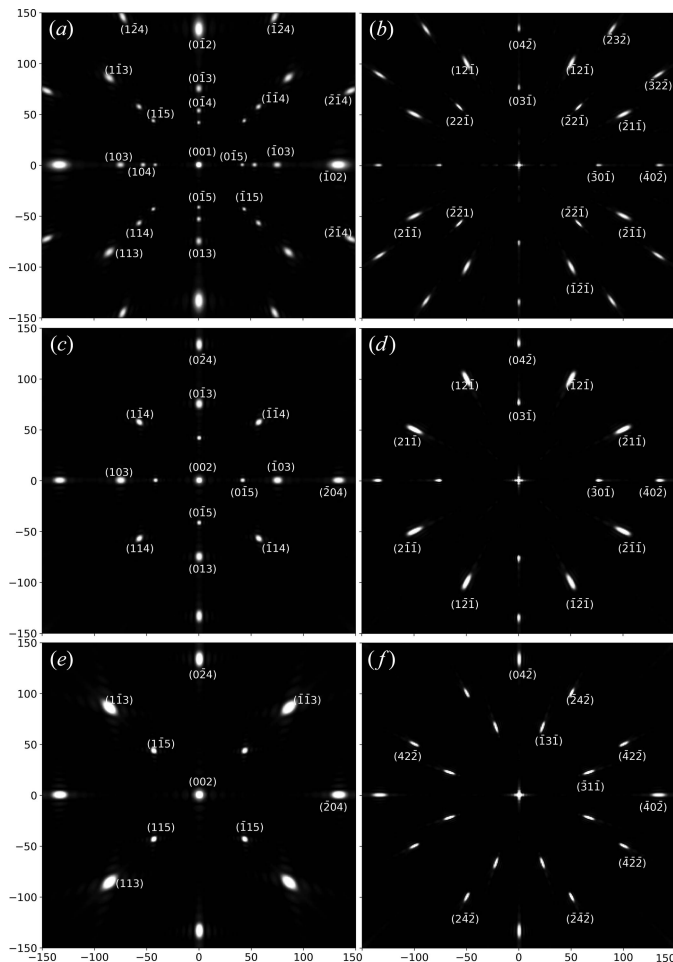
methods for single crystals – the Laue method and the (monochromatic) rotation method.

### 6.1. Varying wavelength at fixed angle: Laue method

In the Laue method, one fixes the orientation of the sample (thus the angle  $\theta$  in Bragg's law) and changes the wavelength of the incident beam over a certain range  $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ , which is thus called 'white color'.

For each pixel  $(X, Y)$  on the film, the scattering intensity is then the sum of contributions from all wavelengths, or equivalently all parallel scattering vectors  $\mathbf{q}$ , which can be formally written as

$$\bar{I}(X, Y) = \sum_{\mathbf{q}} I(\mathbf{q}) = \sum_{\mathbf{q}} \left[ \left| \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \cos(\mathbf{q} \cdot \mathbf{r}_i) \right|^2 + \left| \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \sin(\mathbf{q} \cdot \mathbf{r}_i) \right|^2 \right]. \quad (40)$$



**Figure 8**

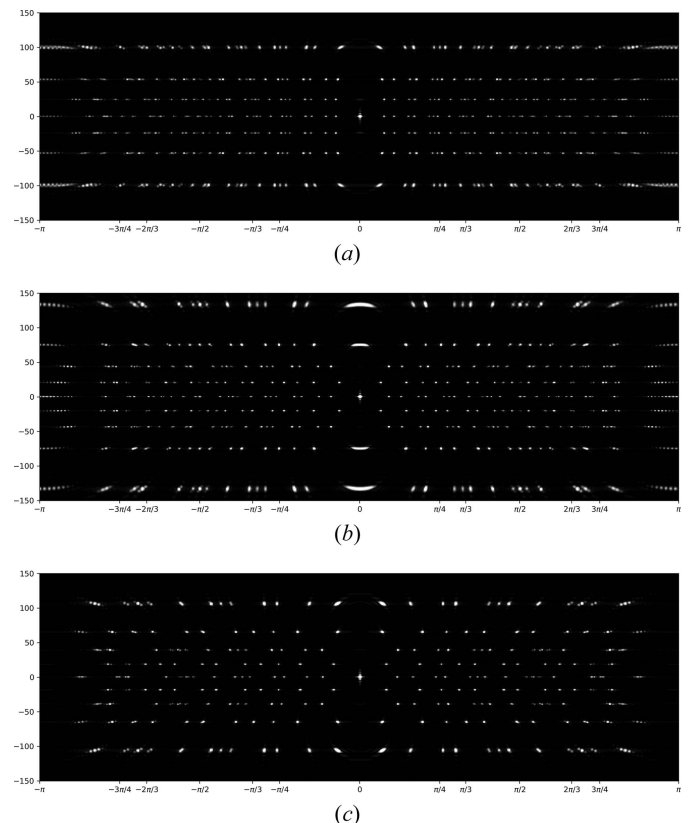
Back-reflection (left column) and transmission (right column) photography of SC (a), (b) ( $N = 3375$ ), b.c.c. (c), (d) ( $N = 4394$ ) and f.c.c. (e), (f) ( $N = 5324$ ) crystals. The range of wavelength  $\lambda$  is  $0.35\text{--}1.0\sigma$  for SC back-reflection,  $0.199\text{--}0.35\sigma$  for SC transmission,  $0.4\text{--}1.2\sigma$  for b.c.c. back-reflection,  $0.23\text{--}0.4\sigma$  for b.c.c. transmission,  $0.5\text{--}1.42\sigma$  for f.c.c. back-reflection and  $0.23\text{--}0.49\sigma$  for f.c.c. transmission. The  $(X, Y)$  coordinate range  $[-150, 150]$  is now set by the grid resolution of the computer code, which can be mapped onto the real length unit on a physical film.

This type of general equation, which computes scattering intensity from all atoms in the sample, reduces to a simple summation over atoms in the unit cell for ideal crystals, as explained in Section 3.3.

We demonstrate the photography results using perfect SC, b.c.c. and f.c.c. samples (Fig. 8). The incident beam is along the  $[001]$  direction and the nearest-neighbor distance  $\sigma$  is set as the unit of length. The code to compute  $\bar{I}(X, Y)$  numerically implementing equation (40) is provided online. The value of  $D$  can be chosen arbitrarily, with all other lengths calculated accordingly, because it only leads to a scaling of the photograph. Here, we set  $D = 100\sigma$  for numerical convenience. If the total number of pixels on the film is  $N_{XY}$  and the number of wavelengths scanned is  $N_\lambda$ , then the computational complexity using equation (40) is  $\mathcal{O}(N_{XY}N_\lambda N)$ .

### 6.2. Varying angle using fixed wavelength: rotation method

We use the conventional setup – cylindrical film – to explain the rotation method for the same SC, b.c.c. and f.c.c. crystalline samples as above (Fig. 9). The wavelength  $\lambda$  of the incident beam is fixed in this method, and the sample placed at the central axis of the cylinder is rotated by a certain angle to probe possible orientations and scattering angles  $2\theta$  for given crystallographic planes. A full circle of  $2\pi$  rotation is only necessary for noncentrosymmetric crystals containing elements that exhibit anomalous dispersion; a rotation of  $2\pi$  is sufficient for centrosymmetric crystals.



**Figure 9**

Rotation photography of SC (a) ( $N = 3375$  and  $\lambda = 0.235\sigma$ ), b.c.c. (b) ( $N = 4394$  and  $\lambda = 0.231\sigma$ ) and f.c.c. (c) ( $N = 5324$  and  $\lambda = 0.257\sigma$ ) crystals.



The scattering intensity at coordinates  $(X, Y)$  is then

$$\bar{I}(X, Y) = \sum_{\Omega} \left\{ \left| \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \cos[\mathbf{q} \cdot \mathbf{r}_i(\Omega)] \right|^2 + \left| \sum_{i=1}^N \hat{f}_i(\mathbf{q}) \sin[\mathbf{q} \cdot \mathbf{r}_i(\Omega)] \right|^2 \right\}, \quad (41)$$

where  $\Omega$  represents the orientation of the sample due to rotation. For a given sample, we apply a rotation matrix about its  $y$  axis to transform particle coordinates into new values. The accumulated signal  $\bar{I}(X, Y)$  on the cylinder is then unfolded onto a rectangle. If a total number  $N_{\Omega}$  of rotation angles within  $(0, 2\pi)$  are scanned, the computational complexity to implement equation (41) to produce results on  $N_{XY}$  pixels is then  $\mathcal{O}(N_{XY}N_{\Omega}N)$ .

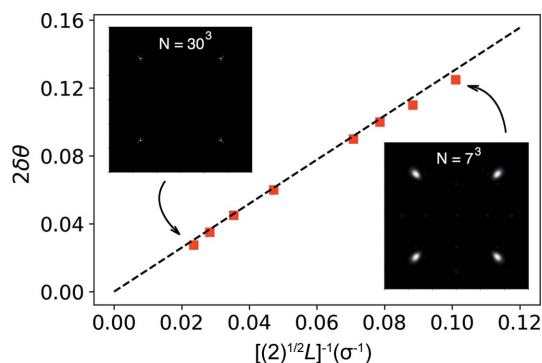
### 6.3. Broadening due to the finite-size effect

So far we have assumed that either varying wavelength or varying sample orientation is needed to satisfy Bragg's law and produce nonvanishing scattering signals on the photograph. However, this is only true for infinitely large systems. In our small samples with  $N \sim 10^3$  particles, signal broadening allows us to observe certain scattering patterns, even when the wavelength  $\lambda$  is fixed at one appropriate value.

For example, in the previously mentioned SC crystals, we can see four scattering spots in the back-reflection method at fixed wavelength  $\lambda = 0.55\sigma$ , which correspond to the (113) planes and equivalents (Fig. 10). When the system size is varied from  $N = 7^3$  to  $30^3$ , the size of each spot decreases. It can be confirmed that the relationship between box size  $L = N^{1/3}$  and spot size  $2\delta\theta$  roughly satisfies the Scherrer equation  $2\delta\theta \propto 1/L$ . An empirical scaling factor  $2^{1/2}$  is needed on  $L$  to estimate the actual dimension of the sample perpendicular to the (113) planes and to agree with the theoretical slope  $\lambda/\cos\theta$ .

### 6.4. DNA double helix

One of the most successful and famous applications of scattering methods is the determination of the DNA structure,



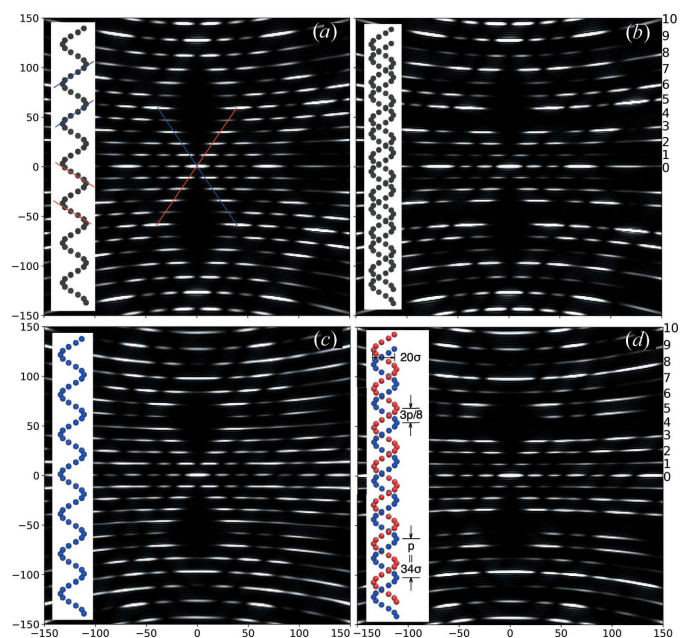
**Figure 10** System size effect on scattering size in back-reflection of SC samples with fixed wavelength  $\lambda = 0.55\sigma$ . The scattering angle  $2\theta$  for these four spots can be computed from  $\tan(\pi - 2\theta) = 85(2)^{1/2}/100$ . After  $L$  is scaled by a factor of  $2^{1/2}$ , the data (red squares) agree with the theoretical slope (dashed line) from the Scherrer equation  $\lambda/\cos\theta$ .

whose X-ray photography shows a characteristic 'X'-shape pattern with horizontal stripes (Franklin, 1953). The form of the pattern can be understood analytically by diffraction from the 2D projected sinusoidal waves of the single or double helix (Kittel, 1968; Thompson *et al.*, 2018). Here we produce the transmission photography of a single model DNA fiber with only backbone particles. Each helix has  $N = 70$  particles with 10 particles per turn (pitch). The parameters of the right-handed B-DNA,  $34\sigma$  for pitch and  $20\sigma$  for helix diameter, are used (Fig. 11). The unit of length  $\sigma$  can be mapped onto the real length unit  $\text{\AA}$ .

All four photographs, with the fiber being single- or double-stranded, 2D projected, or 3D stereoscopic, have an 'X'-shape pattern at the center and are made of horizontal broken stripes (Fig. 11). The two branches of the 'X' pattern can be viewed as scattering signals from the two series of parallel particles on the sinusoidal wave [Fig. 11(a)]. Because each pitch of the helix has ten particles, the pattern has a vertical period of ten stripes (Kittel, 1968). The brightness and darkness along each horizontal stripe depend sensitively on the relative position between different particles (Lucas & Lambin, 2005). For example, the level 4 stripe disappears when two double strands with a phase difference of  $3/8$  pitch are present. The bright level 8 signal of 3D samples at  $X = 0$  is missing for 2D structures.

## 7. Scattering vector $\mathbf{q}$ in intensity scanning

In this section, we discuss the choice of scattering vector  $\mathbf{q}$  in the case of disordered or partially ordered samples that are



**Figure 11** Transmission photography of single-strand (a), (c) and double-strand (b), (d) 2D sinusoidal waves (a), (b) and 3D DNA helices (c), (d) using  $\lambda = 1.54\sigma$ . Each helix is made of a backbone of  $N = 70$  particles with a pitch of  $p = 34\sigma$  and a diameter of  $20\sigma$ . There are 10 particles per pitch. The two helices in the double-strand structure are offset by  $3/8$  pitch. The particle size in the insets is set as  $5\sigma$  to enhance visibility.

spatially isotropic or approximately isotropic. When samples are isotropic, the scattering intensity  $I(\mathbf{q})$  or its normalized version, the structure factor  $S(\mathbf{q})$ , only depends on the magnitude  $q$  of the scattering vector, and thus does not generate isolated spotty signals as in photography of ordered samples. The photography  $I(X, Y)$ , often of less interest in this context, should ideally exhibit concentric circular patterns. The intensity scanning  $I(q)$  or  $S(q)$  as a function of  $q$  is the primary method used for isotropic samples.

### 7.1. Vector $\mathbf{q}$ along a single direction to represent magnitude $q$ in isotropic systems

In an experiment, one can vary  $q$  by observing signals at continuously changing scattering angle  $2\theta$  using a fixed incident wavelength  $\lambda$ . Because experimental samples are generally large enough, a well averaged scattering signal can be detected along one particular direction at  $2\theta$ , as in the powder method with a diffractometer.

For example, consider a polycrystal with  $M$  randomly oriented crystalline grains (domains), each of  $N$  particles. The scattering intensity at  $\mathbf{q}$  computed from equation (9a), assuming  $\hat{f}_i(\mathbf{q}) = 1$ , is

$$I(\mathbf{q}) = \left| \sum_{n=1}^M \sum_{i=1}^N \cos(\mathbf{q} \cdot \mathbf{r}_{n,i}) \right|^2 + \left| \sum_{n=1}^M \sum_{i=1}^N \sin(\mathbf{q} \cdot \mathbf{r}_{n,i}) \right|^2, \quad (42)$$

where  $\mathbf{r}_{n,i}$  is the position vector of particle  $i$  in grain  $n$ . If  $M$  is large and the crystalline grains are uniformly oriented in all directions,  $I(\mathbf{q})$  at the particular vector  $\mathbf{q}$  can be accurate enough to represent  $I(q)$  at the magnitude  $q$ , without averaging over all directions of  $\mathbf{q}$ . A similar argument applies to bulk liquids or glasses, in which  $I(\mathbf{q})$  is also well self-averaged.

### 7.2. Random rotation of a small anisotropic sample

The above method of using a scattering vector  $\mathbf{q}$  in one direction to represent the magnitude  $q$  does not work well for simulation samples, which are usually small and anisotropic (single crystal instead of polycrystal). To simulate experimental results, we can fix the direction of the incident ray but randomly rotate the small sample to many orientations. This is done by applying a 3D rotation matrix to the original particle coordinates  $\mathbf{r}_i$ , whose rotation axis is uniformly distributed on a sphere and rotation angle is uniformly chosen from  $[0, 2\pi]$ . Then the signal  $I(\mathbf{q})$  in equation (42) can be approximated by accumulating intensities from all those orientations  $\Omega$ :

$$I(\mathbf{q}) = \sum_{\Omega} \left\{ \left| \sum_{i=1}^N \cos[\mathbf{q} \cdot \mathbf{r}_i(\Omega)] \right|^2 + \left| \sum_{i=1}^N \sin[\mathbf{q} \cdot \mathbf{r}_i(\Omega)] \right|^2 \right\}. \quad (43)$$

Here,  $\mathbf{r}_i(\Omega)$  represents the new coordinates of particle  $i$  after rotation to orientation  $\Omega$ .

Note that the sum  $\sum_{\Omega}$  is applied to the intensity rather than within the square like  $|\sum_{\Omega} \sum_{i=1}^N \dots|^2$ . The latter choice would imply a virtual system of many randomly orientated overlapping grains, each of size  $N$ . The positions of these virtual grains generated by rotation do not reflect the absolute posi-

tions of grains in the real polycrystal. According to the equivalency of equations (9a) and (9b), equation (43) is an approximation to equation (42) by only considering relative positions of particles within each grain  $\mathbf{r}_i(\Omega) - \mathbf{r}_j(\Omega)$ . Therefore, the difference between the coordinates of particles  $i$  and  $j$  at two different orientations,  $\mathbf{r}_i(\Omega) - \mathbf{r}_j(\Omega')$ , does not affect the result of equation (43), but will lead to different and wrong results if the sum is taken as  $|\sum_{\Omega} \sum_{i=1}^N \dots|^2$ .

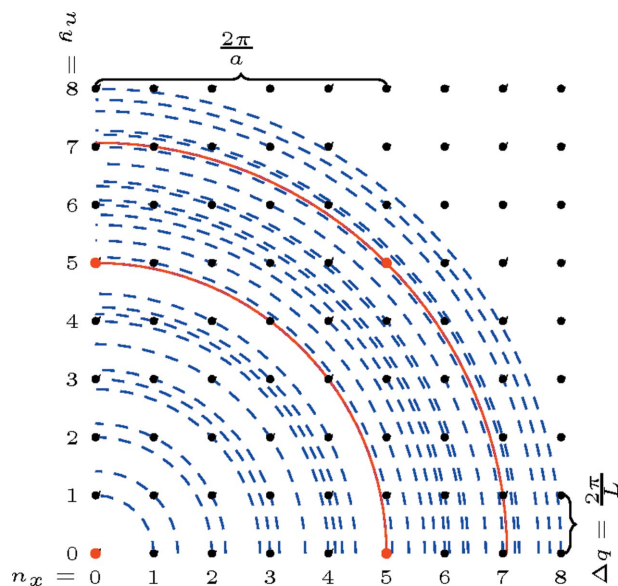
For small and nearly isotropic liquids or glasses, one can replace random rotations of the sample by averaging over many thermally equilibrated configurations. For anisotropic systems, however, rotations are needed to sample different directions.

### 7.3. Scattering vector $\mathbf{q}$ on a lattice

An alternative and more convenient way to simulate experiments is to fix the sample coordinates and choose  $\mathbf{q}$  of a given  $q$  from all directions. It is often suggested to select  $\mathbf{q}$  from a 3D orthorhombic lattice,  $\mathbf{q} = \Delta q(n_x, n_y, n_z)$ , with integers  $n_x, n_y, n_z$  and increment  $\Delta q = 2\pi/L$ , where  $L$  is the linear dimension of the cubic simulation box (Allen & Tildesley, 1987). The motivation here is that  $L$  sets the maximum periodicity of the simulation sample that is still physically meaningful, and thus the resolution of  $q$ . The integers  $n_x, n_y, n_z$  may be chosen to run from negative to positive values to sample spherically symmetric  $\mathbf{q}$ s, or to start from zero to sample only  $\mathbf{q}$ s on 1/8 of the sphere. At the expense of symmetry and averaging, the latter choice can reach a higher-magnitude  $q$  with the same number of lattice points.

There are multiple  $\mathbf{q}$ s on this lattice that correspond to the same magnitude  $q$ , from which we can compute an average  $S(q)$ . The number of  $\mathbf{q}$ s for a given magnitude  $q$  tends to, but does not necessarily, increase with  $q$ . For example, in a 2D system with  $\mathbf{q}$ s on a square lattice, there are 1, 2, 1, 2, 2, 1, 2, ...  $\mathbf{q}$  points on the lattice at magnitude  $q/\Delta q = 0, 1, 2^{1/2}, 2, 5^{1/2}, 8^{1/2}, 9, \dots$ , respectively (Fig. 12). When reporting the result of  $S(q)$ , one can assign  $q$  values into bins of equal size or just use the original  $q$  values visited by the lattice points. In both cases,  $S(q)$  should be the mean value averaged over all the  $\mathbf{q}$ s at that  $q$ .

If the sample is crystalline and  $L$  is an integer multiple of the crystallographic lattice constant  $a$ , then the  $\mathbf{q}$  lattice contains the reciprocal-lattice points of the crystal (subject to a  $2\pi$  factor difference) (Allen & Tildesley, 1987). If  $L = 5a$  in the above 2D example, then  $\mathbf{q}$  points  $\Delta q(0, 0)$ ,  $\Delta q(5, 0)$ ,  $\Delta q(0, 5)$ ,  $\Delta q(5, 5)$  correspond to reciprocal-lattice points  $(0, 0)$ ,  $(1/a, 0)$ ,  $(0, 1/a)$ ,  $(1/a, 1/a)$ , respectively (Fig. 12). These lattice points are where Bragg's law [equation (19)] is obeyed. Therefore, according to the discussion in Section 3.3, if all atomic form factors are unity,  $I(q) = |F_{hkl}|^2 = N^2$  and  $S(q) = N$  at each of these reciprocal-lattice points. The intensity scanning result  $S(q)$  needs to be an average over all  $\mathbf{q}$  points at that  $q$ , some of which are not reciprocal-lattice points and thus have  $S(q) = 0$ . For example, at  $q = 5\Delta q$  of the 2D system, two points have  $S(q) = N$  and two have  $S(q) = 0$ . The average  $S(q = 5\Delta q)$  is thus  $(N + N + 0 + 0)/4 = N/2$  (Fig. 12).



**Figure 12** Scattering vector  $\mathbf{q} = \Delta q(n_x, n_y)$  on a square lattice used for a 2D system of box size  $L$ .  $\mathbf{q}$  points at the same magnitude  $q$  are connected by concentric quarter circles up to  $q = 8\Delta q$ . If the system is a crystal of lattice constant  $a = L/5$ , then four points shown in red correspond to reciprocal-lattice points.

Using lattice points to approximate  $\mathbf{q}$ s from all directions is problematic when  $L$  is small and thus the increment  $\Delta q$  is large, such that only a few  $\mathbf{q}$ s are available at each  $q$ . The issue is more severe at small  $q$  or towards corners of the cubic lattice at high  $q$ . The calculated signal  $S(q)$  can then be quite noisy because  $\mathbf{q}$  is not averaged enough over all directions.

### 7.4. Scattering vector $\mathbf{q}$ on a sphere and Debye’s scattering equation

In order to obtain a smooth curve of  $S(q)$  that better matches experimental results, we need to use enough spherically distributed  $\mathbf{q}$ s. To guarantee a uniform distribution of points on a sphere, we apply the Fibonacci grid approach to randomly choose  $N_q$  scattering vectors  $\mathbf{q}$  from a sphere of radius  $q$  (Saff & Kuijlaars, 1997). Increasing  $N_q$  improves the effect of averaging. The complexity to compute  $I(q)$  or  $S(q)$  at each  $q$  is then  $\mathcal{O}(N_q N)$ .

In the limit of  $N_q \rightarrow \infty$ , using equation (9b), we can integrate over all  $\mathbf{q}$  directions and then normalize by the full solid angle of  $4\pi$  to compute the average  $I(q)$ :

$$\begin{aligned}
 I(q) &= \frac{1}{4\pi} \int_{|\mathbf{q}|=q} d\mathbf{q} \sum_{i=1}^N \sum_{j=1}^N \hat{f}_i(\mathbf{q}) \hat{f}_j(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}_{ij}) \\
 &= \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \sum_{i=1}^N \sum_{j=1}^N \hat{f}_i(q) \hat{f}_j(q) \exp(iqr_{ij} \cos \theta) \\
 &= \sum_{i=1}^N \sum_{j=1}^N \hat{f}_i(q) \hat{f}_j(q) \frac{\sin(qr_{ij})}{qr_{ij}}. \tag{44}
 \end{aligned}$$

This is known as Debye’s scattering equation (Thomas, 2010; Gelisio & Scardi, 2016), which can also be viewed as the

discrete version of the Fourier transform of the radial distribution function  $g(r)$  in equation (33a). The computational complexity of Debye’s method is  $\mathcal{O}(N^2)$  and it becomes more efficient than numerically sampling  $N_q$   $\mathbf{q}$  vectors on a sphere when  $N < N_q$ .

## 8. Photography and intensity scanning of disordered or partially ordered samples

Although intensity scanning  $I(q)$  or  $S(q)$  as a function of  $q$  generally gives more useful structural information about isotropic samples, it is sometimes interesting to show the corresponding photography  $I(X, Y)$ . In fact, intensity scanning can be obtained from photography by moving along a specific radial direction on the  $(X, Y)$  film, as in the early days of the powder method (Cohen, 1935).

To generate scattering photography of isotropic samples, we use the rotation or thermal averaging method of Section 7.2. Intensity scanning profiles are calculated using the three methods mentioned in Sections 7.3 and 7.4.

### 8.1. Liquids and glasses

If a scattering photograph is taken for disordered samples like liquids or glasses using a fixed wavelength, a characteristic ring signal is expected at peak value  $q^* \sim 2\pi/\sigma$  which corresponds to the molecular size  $\sigma$ . This ring is regular and clear, when the sample, like most experimental bulk samples, is large enough such that a good average is taken within the system in the calculation of  $I(\mathbf{q})$ . However, in a small simulation system ( $N = 10^3$ – $10^4$ ), photography of one static disordered sample gives spotty and noisy signals with certain traces of ring features [Fig. 13(a)]. To enhance the sharpness of the ring, one can either increase the size  $N$  of the sample or take the ensemble average of  $I(\mathbf{q})$  over many configurations [Fig. 13(c)].

For homogeneous liquids and glasses, the static structure factor  $S(\mathbf{q}) = S(q)$  varies only with the magnitude  $q$  of the scattering vector and exhibits a major peak at  $q^* \sim 2\pi/\sigma$ . Using  $\mathbf{q}$ s on a sphere numerically or Debye’s equation can generate well averaged smooth  $S(q)$  curves for liquids or glasses [Figs. 13(b) and 13(d)]. If only one disordered configuration is analyzed, the  $S(q)$  curve is much noisier using  $\mathbf{q}$ s on a cubic lattice [Fig. 13(b)].

### 8.2. Polycrystalline samples – powder method

The powder method is often used to analyze polycrystals, in which a crystalline sample is ground into powder to produce many small randomly oriented crystalline grains. Then, at any scattering angle  $2\theta$  where a strong signal is expected, at least one of the grains has the correct orientation by chance to satisfy Bragg’s law. The measured intensity scanning  $S(q)$  can be used to calculate interplanar spacings in the crystal and, with some limitations, even to determine the crystal structure.

It is difficult to produce a well randomized polycrystalline sample in simulation, given the limit of system size. Nevertheless, we can start from a small single-crystal sample and use random rotation or  $\mathbf{q}$ s from different directions to simulate

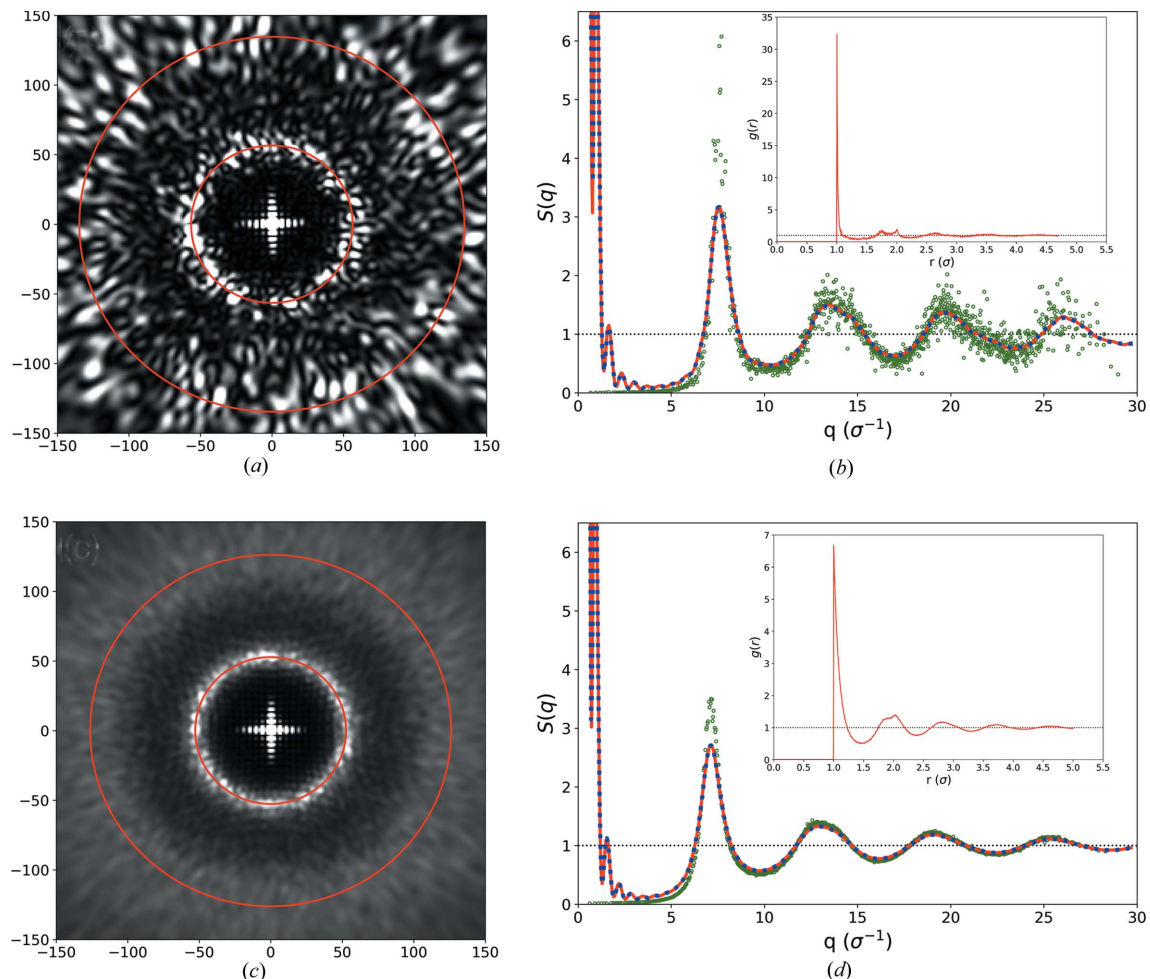
scattering signals of a polycrystal. In particular, we use the same SC, b.c.c. and f.c.c. crystals used above to generate photographs and intensity scanning results of corresponding polycrystals (Fig. 14).

The sharp concentric rings in the photograph  $I(X, Y)$  and the narrow peaks in  $S(q)$  correspond to scattering from different crystallographic planes ( $hkl$ ) of the three crystals (SC, b.c.c. and f.c.c.).  $S(q)$  peaks computed from spherically distributed  $\mathbf{q}$ s are lower and broader than those from cubic lattice  $\mathbf{q}$ s. The peak height using cubic lattice  $\mathbf{q}$ s often scales with system size  $N$ . For example, the SC crystal has  $L = 15\sigma$  and  $N = L^3 = 3375$  particles. Given  $\Delta q = 2\pi/15\sigma$ , the  $S(q)$  peak from (100) planes is expected to occur at six  $\mathbf{q}$  points,  $\Delta q(15, 0, 0)$ ,  $\Delta q(0, 15, 0)$ ,  $\Delta q(0, 0, 15)$ ,  $\Delta q(-15, 0, 0)$ ,  $\Delta q(0, -15, 0)$  and  $\Delta q(0, 0, -15)$ , each having a value  $S(q) = 3375$ . However, there are other  $\mathbf{q}$  points with magnitude  $q = 15\Delta q$ , which correspond to integer solutions to  $n_x^2 + n_y^2 + n_z^2 = 15^2$ . In total, at  $q = 15$ , there are six (15,0,0)-like (considering its permutation and  $\pm$ ), 24 (12,9,0)-like, 24

(10,10,5)-like, 48 (11,10,2)-like and 48 (14,5,2)-like  $\mathbf{q}$  points. Out of these 150 points, only six have  $S(q) = N$  while the others have  $S(q) = 0$ . So the peak height  $S(q = 15\Delta q) = 3375 \times 6/150 = 135$ .

### 8.3. Mesophases: small-angle method

Mesophases are states of matter intermediate between liquids and solids found in block copolymers (Sakurai *et al.*, 1991), liquid crystals (Mitchell *et al.*, 1983), structural DNAs (Tian *et al.*, 2020) *etc.*, which present mesoscopic ordering of length scales larger than molecular size  $\sigma$ . To detect these long-wavelength structures, small-angle X-ray scattering (Chu & Hsiao, 2001) or small-angle neutron scattering (Richards & Thomason, 1983) methods are needed because scattering signals are expected at small  $q$  (before the first major diffraction peak  $\sim 2\pi/\sigma$ ) and thus small  $\theta$  as seen from equation (2). A logarithmic scale axis is often set for  $S(q)$  in the structure-factor plot because at  $q \rightarrow 0$  the signal scales with system size  $N$  (Schneidman-Duhovny *et al.*, 2010).



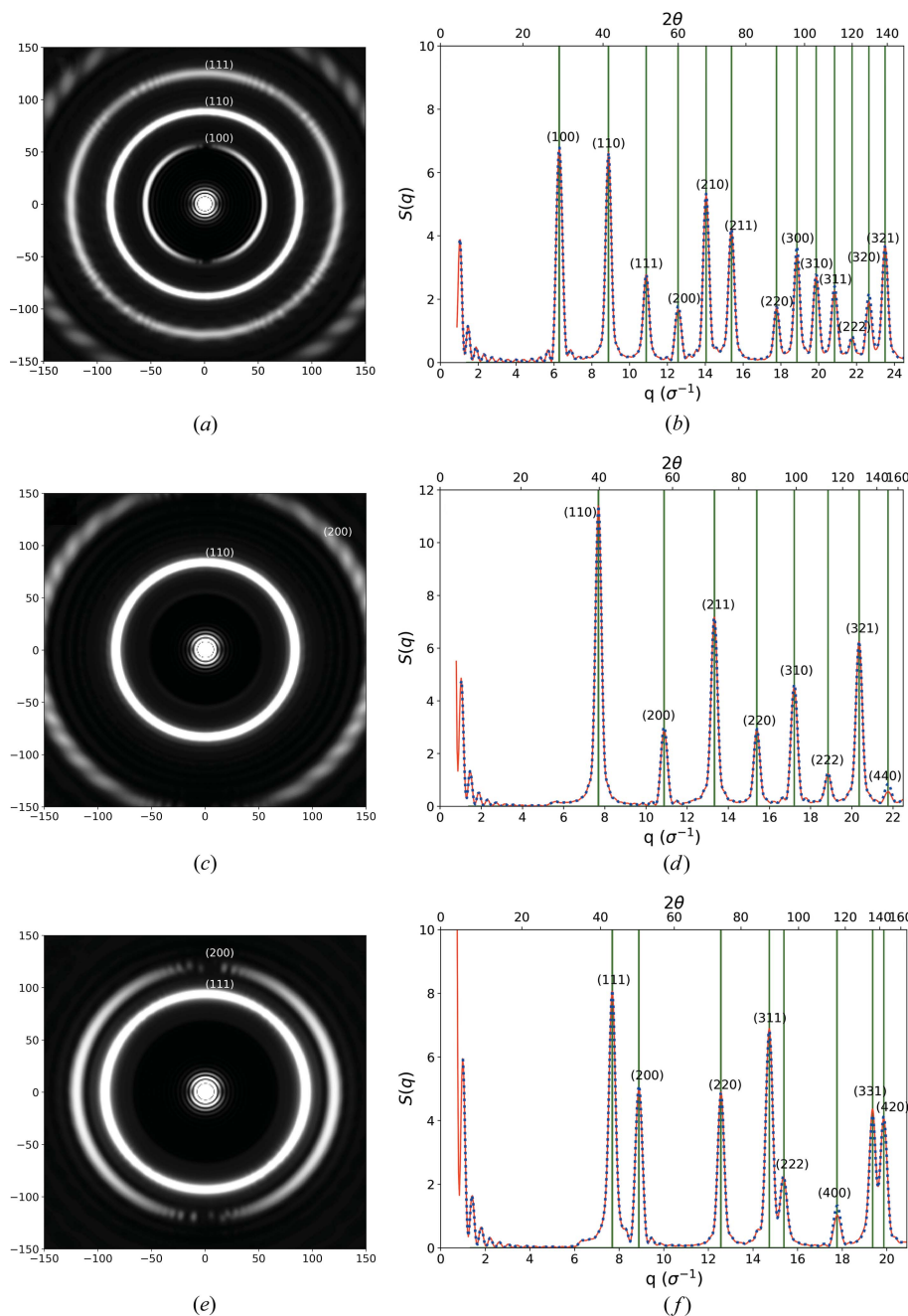
**Figure 13**

Simulated transmission photography using fixed wavelength  $\lambda = 0.4\sigma$  (a), (c) and structure factor  $S(q)$  (b), (d) of a static glass sample (a), (b) and a thermally averaged liquid sample (c), (d). Red solid rings in the photographs correspond to the first and second peaks in  $S(q)$ . The cross pattern at the center of each photograph is due to Fraunhofer diffraction from the small simulation box, effectively a cubic obstacle. Three methods are used to compute  $S(q)$ : with  $\mathbf{q}$ s on a cubic lattice (green circles), with  $\mathbf{q}$ s on spheres (blue dotted line) and Debye's scattering equation (red solid line). Insets show the radial distribution function  $g(r)$ . Both samples are  $N = 1000$  hard spheres of diameter  $\sigma$ . The glass sample has one configuration at packing fraction 0.64. The liquid sample has 1000 thermally equilibrated configurations at packing fraction  $\pi/6 = 0.5236$ .

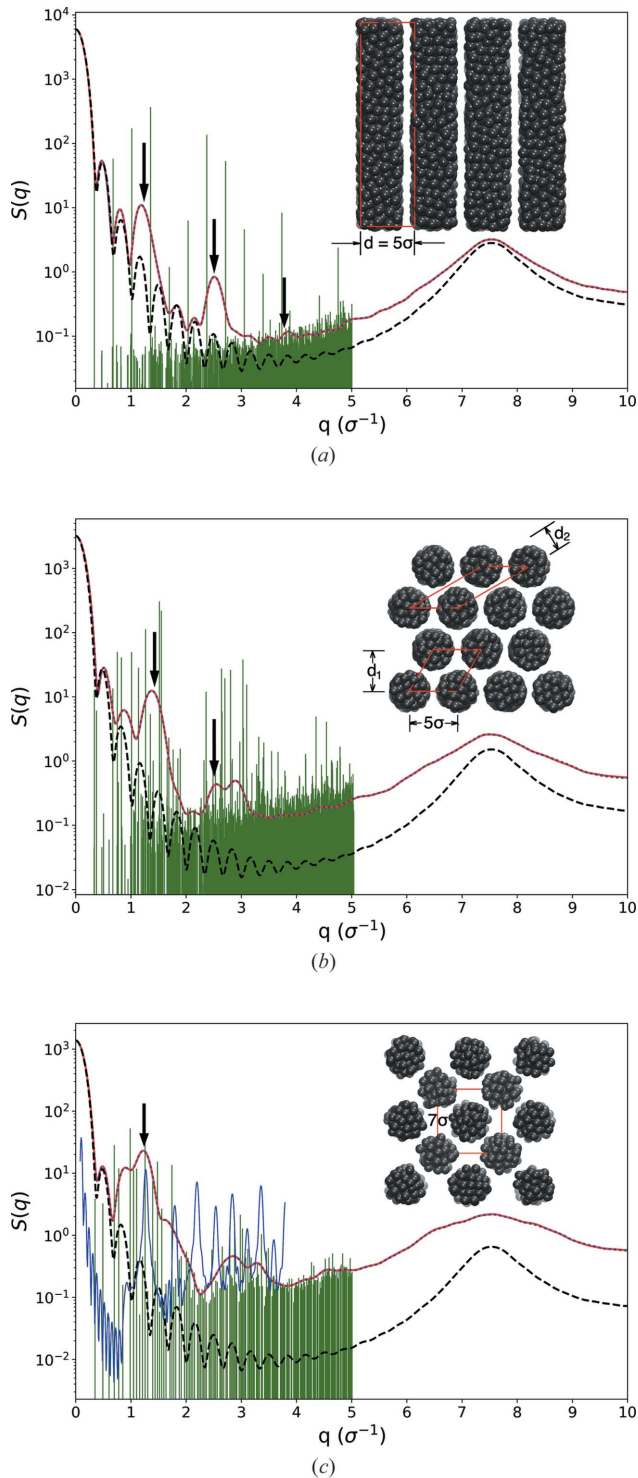
Experimental mesophases are often synthesized as polycrystals or many small crystalline domains randomly embedded in an amorphous matrix, for which intensity scanning  $S(q)$  at small  $q$  is used to reveal the ordering. We illustrate the concepts of the mesophase structure factor using a lamellar, a cylindrical and a b.c.c. spherical configuration of domains, cut from a disordered glass sample of hard spheres of diameter  $\sigma$ . These structures are thus amorphous within each

domain, but the domains form a 1D, 2D or 3D superlattice for the lamellar, cylindrical or spherical configuration, respectively.

In the lamellar phase, each period is of length  $d = 5\sigma$ , consisting of a layer with thickness  $3.5\sigma$  and a gap with thickness  $1.5\sigma$ . We find three peaks of  $S(q)$  at one, two and three times  $2\pi/d \simeq 1.257\sigma^{-1}$ , corresponding to the first, second and third order of Bragg diffraction of the superlattice



**Figure 14** Simulated powder method. Transmission photography using fixed wavelength  $\lambda = 0.4\sigma$  (a), (c), (d) and structure factor  $S(q)$  (b), (d), (f) of polycrystalline SC (a), (b), b.c.c. (c), (d) and f.c.c. (e), (f) samples. Miller indices ( $hkl$ ) are labeled next to each signal peak. The photograph is produced by randomly rotating a single-crystalline sample in three dimensions and taking the average of  $I(\mathbf{q})$  over 5000 orientations. The small concentric circular pattern at the center of each photograph is due to Fraunhofer diffraction from the small simulation box, effectively a circular obstacle, after random rotation. Three methods are used to compute  $S(q)$ : with  $\mathbf{q}$ s on a cubic lattice (green vertical lines), with  $\mathbf{q}$ s on spheres (blue dotted line) and Debye's scattering equation (red solid line).



**Figure 15** Simulated small-angle structure factor  $S(q)$  for (a) lamellar, (b) cylindrical and (c) b.c.c. spherical mesophases. Three methods are used to compute  $S(q)$ : with  $\mathbf{q}$ s on a cubic lattice (green vertical lines), with  $\mathbf{q}$ s on spheres (blue dotted line) and Debye's scattering equation (red solid line). The Debye result for  $S(q)$  of a homogeneous glass, after being vertically rescaled to align at  $q \rightarrow 0$ , is shown for comparison (black dashed line). Black downward-pointing arrows mark signature peaks for each structure. The broad peak at  $7.5\sigma^{-1}$  corresponds to particle size  $\sigma$ . Insets show top/side views of the configurations under consideration. The blue solid line in (c) is the Debye result for  $S(q)$  of a b.c.c. sphere mesophase with one particle per domain, obtained by rescaling the curve of a b.c.c. crystal.

[Fig. 15(a)]. The peak height drops as  $q$  increases, and when the layer thickness equals the gap thickness, peaks at even multiples of  $2\pi/d$  disappear.

The cylindrical phase with a disc radius  $1.8\sigma$  resides on a 2D triangular superlattice with lattice constant  $5\sigma$ . By assigning unit cells in two different ways with interplanar spacing  $d_1 = \{[5(3)^{1/2}]/2\}\sigma$  and  $d_2 = 2.5\sigma$ , we can identify two peaks at  $q_1 = 2\pi/d_1 \simeq 1.451\sigma^{-1}$  and  $q_2 = 2\pi/d_2 \simeq 2.513\sigma^{-1}$  [Fig. 15(b)]. The second-order peak around  $2q_1 \simeq 2.9\sigma^{-1}$  is also visible (not marked).

The spherical phase has spheres of radius  $2\sigma$  that pack on a b.c.c. superlattice with a lattice constant  $7\sigma$ . If each sphere domain had only one particle, the structure factor would be the same as for a normal b.c.c. crystal apart from a change of unit for  $q$ . We can obtain  $S(q)$  of this one-particle spherical phase by rescaling the  $q$  axis of  $S(q)$  of the b.c.c. crystal, which has a lattice constant  $a = (2/3^{1/2})\sigma$ , by a factor of  $\{[7(3)^{1/2}]/2\} \simeq 6.062$ . This moves the (110) peak from  $7.7\sigma^{-1}$  to  $1.27\sigma^{-1}$  [Fig. 15(c)]. This helps us to identify that only the peak from the (110) planes of the b.c.c. superlattice is sharply distinguishable from the background signals.

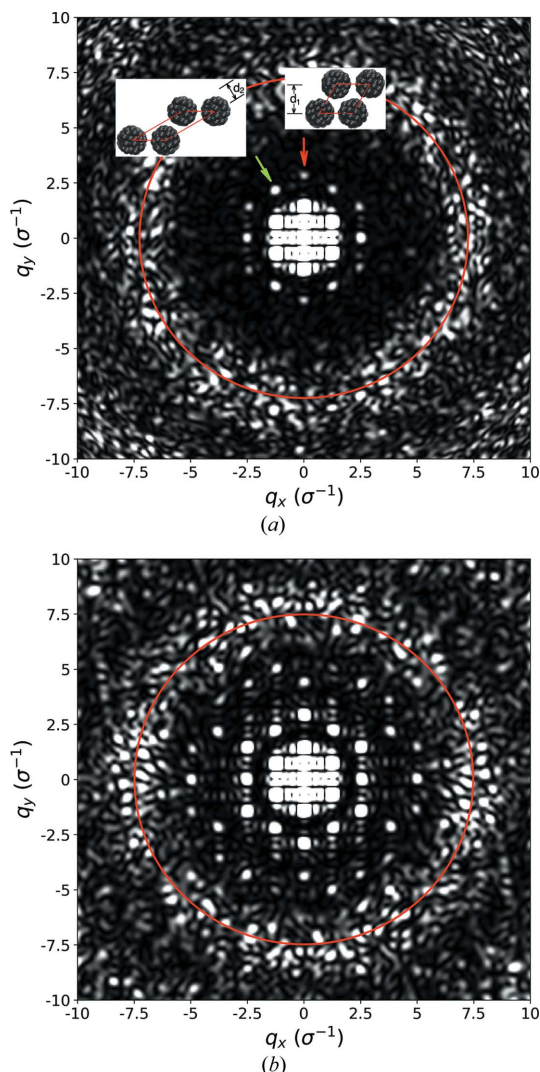
### 9. 2D structure factor

For 2D samples or a 2D projection of 3D samples, it is sometimes useful to express  $S(\mathbf{q})$  as a 2D function of  $(q_x, q_y)$  (Tutsch *et al.*, 2014) or scattering angles  $(\theta_x, \theta_y)$  (Lee *et al.*, 2005). The 2D structure factor  $S(q_x, q_y)$  is related to the photography  $I(X, Y)$  by converting coordinates  $(X, Y)$  on the film into components  $(q_x, q_y)$  of the scattering vector using equation (35). For 3D structures, the component  $q_z$  can be expressed as a function of  $q_x$  and  $q_y$ , for example, in the case of the transmission method [equation (38)],

$$\begin{aligned} q_z &= -\frac{2\pi}{\lambda}(1 - D/L) \\ &= -\frac{2\pi}{\lambda} \left\{ 1 - \left[ 1 - \frac{\lambda^2}{(2\pi)^2} q_x^2 - \frac{\lambda^2}{(2\pi)^2} q_y^2 \right]^{1/2} \right\} \\ &= -\frac{2\pi}{\lambda} + \left[ \left( \frac{2\pi}{\lambda} \right)^2 - q_x^2 - q_y^2 \right]^{1/2}. \end{aligned} \quad (45)$$

Note that knowing  $(q_x, q_y)$  does not uniquely determine  $q_z$ . The constant  $2\pi/\lambda$  still needs to be specified. For small-angle scattering with  $q_x, q_y \rightarrow 0$ , an approximation to set  $q_z = 0$  is valid if there is no long-range periodicity along the  $z$  direction.

We compute  $S(q_x, q_y)$  for the cylindrical mesophase in Fig. 15(b). The cylinder axis is aligned with the incident ray in the  $z$  direction. We first use  $q_z$  calculated from equation (45) with  $\lambda = 0.4\sigma$ . Besides the isotropic circular signal corresponding to particle size  $\sigma$ , a characteristic hexagonal pattern with sixfold symmetry is observed at small  $q$ , which results from the cylinders packed on a 2D triangular lattice. We can identify two sets of spots on the vertices of hexagons – one corresponds to the unit cell with spacing  $d_1$  and the other corresponds to the unit cell with spacing  $d_2$  [Fig. 16(a)]. The second-order peak related to  $d_1$  and first-order peak related to



**Figure 16**  
 2D structure factor  $S(q_x, q_y)$  of the cylindrical mesophase using  $\lambda = 0.4\sigma$  with (a)  $q_z = -(2\pi/\lambda) + [(2\pi/\lambda)^2 - q_x^2 - q_y^2]^{1/2}$  and (b)  $q_z = 0$ . The red solid ring marks the broad peak corresponding to particle diameter  $\sigma$  at  $\sim 2\pi/\sigma$ . The red arrow points to the second-order peak from the unit cell with spacing  $d_1$ . The green arrow points to the first-order peak from the unit cell with spacing  $d_2$ .

$d_2$  form a hexagon together, while the first-order peak related to  $d_1$  is mixed with the Fraunhofer diffraction pattern at smaller  $q$ .

If we set  $q_z = 0$ , the  $S(q_x, q_y)$  pattern is approximately the same at small  $q$ , with a certain degree of enhancement [Fig. 16(b)]. Some higher-order peaks become visible at larger  $q$ .

### 10. Conclusion

In this article, we give a comprehensive and coherent review of the core concepts of scattering methods used to determine the structures of ordered and disordered samples. Typical examples of scattering photography and intensity scanning are calculated, which can be used as benchmarks. Sample CPU codes are provided on GitHub at <https://github.com/>

statisticalmechanics/scatter to illustrate the mathematics and algorithms. Accelerating GPU codes that can reduce hours of computation to seconds are also provided for efficient simulation of scattering signals.

Note that, for simplicity, the intensity calculation discussed in this paper has omitted several important wavelength and/or angle-dependent factors due to, for example, absorption, extinction, multiple scattering, polarization and the Lorentz factor, as well as the issue of normalization of the measured intensity to an absolute scale. We have also omitted the effect of temperature, which generally adds a Gaussian co-factor to each atomic scattering factor.

### APPENDIX A

#### Fourier transform: continuous and discrete

The Fourier transform  $\hat{F}_{\mathbf{k}}$  of a function  $F(\mathbf{r})$  defined continuously in 3D real space of infinite volume is

$$\hat{F}_{\mathbf{k}} = \int d\mathbf{r} F(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}), \quad (46)$$

where  $\mathbf{k}$  is a wavevector used to extract the spatial periodicity of  $F(\mathbf{r})$  (Lighthill, 1958). For instance, if  $F(\mathbf{r})$  has a periodic pattern of wavelength  $\lambda$  along the  $x$  axis, *i.e.*  $F(x, y, z) = F(x + \lambda, y, z)$ , then the value of  $\hat{F}_{\mathbf{k}}$  is large for a  $\mathbf{k}$  of magnitude  $|\mathbf{k}| = 2\pi/\lambda$  pointing in the  $\mathbf{x}$  direction, *i.e.*  $\mathbf{k} = (2\pi/\lambda, 0, 0)$ . Physically, if  $\exp(i\mathbf{k} \cdot \mathbf{r})$  is viewed as a plane wave traveling in the  $\mathbf{k}$  direction, then  $\hat{F}_{\mathbf{k}}$  would exhibit a peak value when  $F(\mathbf{r})$  has wavelike properties coherent with  $\exp(i\mathbf{k} \cdot \mathbf{r})$  such that they add constructively in the integral. In this sense, the Fourier transform [equation (46)] quantifies the existence and the extent of periodicity corresponding to  $\mathbf{k}$  in  $F(\mathbf{r})$ .

In general, even if  $F(\mathbf{r})$  is a real function,  $\hat{F}_{\mathbf{k}}$  can be complex. However, if  $F(\mathbf{r})$  is real [ $F^*(\mathbf{r}) = F(\mathbf{r})$ ] and even [ $F(-\mathbf{r}) = F(\mathbf{r})$ , *i.e.* with a symmetry center], its Fourier transform  $\hat{F}_{\mathbf{k}}$  is also real and even, because the conjugate of  $\hat{F}_{\mathbf{k}}$  is

$$\begin{aligned} \hat{F}_{\mathbf{k}}^* &= \int d\mathbf{r} F^*(\mathbf{r}) \exp(-i\mathbf{k} \cdot \mathbf{r}) = \int d\mathbf{r} F(\mathbf{r}) \exp(-i\mathbf{k} \cdot \mathbf{r}) = \hat{F}_{-\mathbf{k}} \\ &= \int_{-\infty}^{\infty} d\mathbf{r} F(-\mathbf{r}) \exp(-i\mathbf{k} \cdot \mathbf{r}) \stackrel{\mathbf{r}' = -\mathbf{r}}{=} \int_{\infty}^{-\infty} d\mathbf{r}' F(\mathbf{r}') \exp(i\mathbf{k} \cdot \mathbf{r}') \\ &= \int_{-\infty}^{\infty} d\mathbf{r}' F(\mathbf{r}') \exp(i\mathbf{k} \cdot \mathbf{r}') = \hat{F}_{\mathbf{k}}. \end{aligned} \quad (47)$$

Here the integration limits for the vector variable  $\mathbf{r}$ , formally denoted as  $\pm\infty$ , are to be understood as for each of its components.

The inverse Fourier transform of  $\hat{F}_{\mathbf{k}}$  is an integral in the wavevector space which gives the original real-space function:

$$F(\mathbf{r}) = \frac{1}{(2\pi)^3} \int d\mathbf{k} \hat{F}_{\mathbf{k}} \exp(-i\mathbf{k} \cdot \mathbf{r}). \quad (48)$$

This expands  $F(\mathbf{r})$  in terms of an infinite number of periodic basis functions  $\exp(-i\mathbf{k} \cdot \mathbf{r})$  characterized by different  $\mathbf{k}$ s. The coefficient or contribution of each  $\mathbf{k}$  is just the Fourier transform  $\hat{F}_{\mathbf{k}}$ . In principle, the collection of all  $\hat{F}_{\mathbf{k}}$ s contains the

entire information about the original function  $F(\mathbf{r})$  such that knowing  $\hat{F}_{\mathbf{k}}$ s allows us to reconstruct  $F(\mathbf{r})$ .

In physical systems,  $F(\mathbf{r})$  is often defined within a finite volume  $V$  and the Fourier transform should be integrated over the region  $V$ :

$$\hat{F}_{\mathbf{k}} = \int_V d\mathbf{r} F(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}). \quad (49)$$

If such a finite system is of a cubic shape with a linear dimension  $L$ , *i.e.*  $V = L^3$ , then any periodicity or wavelength  $\lambda > L$  is unphysical. This imposes a lower bound,  $2\pi/L$ , on the smallest wavevector to be considered. The inverse Fourier transform [equation (48)] thus should not vary  $\mathbf{k}$  continuously as in an integral, but only take discrete values of  $\mathbf{k}$  with increments  $(\Delta k_x, \Delta k_y, \Delta k_z) = (2\pi/L, 2\pi/L, 2\pi/L)$ . The integral then becomes (Chaikin *et al.*, 1995)

$$\begin{aligned} F(\mathbf{r}) &= \frac{1}{(2\pi)^3} \sum_{\mathbf{k}} \hat{F}_{\mathbf{k}} \exp(-i\mathbf{k} \cdot \mathbf{r}) \left(\frac{2\pi}{L}\right)^3 \\ &= \frac{1}{V} \sum_{\mathbf{k}} \hat{F}_{\mathbf{k}} \exp(-i\mathbf{k} \cdot \mathbf{r}). \end{aligned} \quad (50)$$

Mathematically, for the Fourier transform [equation (46)] to exist, the function  $F(\mathbf{r})$  needs to be absolutely integrable. If  $F(\mathbf{r})$  equals some nonzero constants, or without loss of generality,  $F(\mathbf{r}) = 1$ , in order to reconcile the singularity, the result of the Fourier transform is formally written as  $(2\pi)^3 \delta_{\mathbf{D}}(\mathbf{k}) = \int d\mathbf{r} \exp(i\mathbf{k} \cdot \mathbf{r})$ , or equivalently,

$$\delta_{\mathbf{D}}(\mathbf{k}) = \frac{1}{(2\pi)^3} \int d\mathbf{r} \exp(i\mathbf{k} \cdot \mathbf{r}), \quad (51)$$

where  $\delta_{\mathbf{D}}(\mathbf{x})$  is the (3D) singular Dirac delta function [ $\delta_{\mathbf{D}}(0) \rightarrow \infty$ ]. Usually, the Dirac delta function with the property that  $\int d\mathbf{x} \delta_{\mathbf{D}}(\mathbf{x}) f(\mathbf{x}) = f(0)$  is introduced as the limiting case of a normalized Gaussian function with its standard deviation approaching zero. According to the above notation, the inverse Fourier transform of the Dirac delta function readily reduces to  $[1/(2\pi)^3] \int d\mathbf{k} (2\pi)^3 \delta_{\mathbf{D}}(\mathbf{k}) \exp(-i\mathbf{k} \cdot \mathbf{r}) = \exp(-i\mathbf{0} \cdot \mathbf{r}) = 1$ . For a system of a finite volume  $V$ , it is also customary to write

$$\int_V d\mathbf{r} \exp(i\mathbf{k} \cdot \mathbf{r}) = V \delta_{\mathbf{k},0}, \quad (52)$$

where  $\delta_{i,j} = 1, i = j$  and  $0, i \neq j$  is the Kronecker delta function.

### APPENDIX B Direct and reciprocal lattices

The position vector  $\mathbf{r}$  of particles or atoms residing on a crystal lattice, the direct lattice, can be expressed as a linear combination,

$$\mathbf{r} = x\mathbf{a} + y\mathbf{b} + z\mathbf{c}, \quad (53)$$

of the (direct) lattice vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c}$ , which are basis vectors of the unit cell with volume  $V_{\text{cell}} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ . Generally,  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  may not be orthogonal to each other and thus  $x, y, z$  are not necessarily the projections of  $\mathbf{r}$  in a Cartesian coordinate

system. If particles coincide with lattice points, then  $x, y, z$  are integers; if particles are contained inside the unit cell, their coordinates  $x, y, z$  can be fractions (Sands, 1993).

Particles on regular crystal lattices are situated on different families of parallel crystallographic planes when viewed from different angles. Such parallel planes are denoted by three integers  $(hkl)$ , the Miller indices, whose reciprocals are proportional to the intercepts of the planes with the three axes of the direct lattice. The spacing or distance,  $d_{hkl}$ , between neighboring lattice planes in the family  $(hkl)$  is a function of the Miller indices and lattice parameters [Fig. 17(a)]. In the special case of an orthorhombic lattice,

$$\frac{1}{d_{hkl}^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}. \quad (54)$$

The reciprocal lattice is defined mathematically in a space spanned by the reciprocal-lattice vectors  $\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*$ , which are related to the direct lattice vectors by

$$\mathbf{a}^* = (\mathbf{b} \times \mathbf{c})/V_{\text{cell}}, \quad (55)$$

$$\mathbf{b}^* = (\mathbf{c} \times \mathbf{a})/V_{\text{cell}}, \quad (56)$$

$$\mathbf{c}^* = (\mathbf{a} \times \mathbf{b})/V_{\text{cell}}. \quad (57)$$

Since  $\mathbf{a}^*$  is orthogonal to  $(\mathbf{b}, \mathbf{c})$ ,  $\mathbf{b}^*$  is orthogonal to  $(\mathbf{a}, \mathbf{c})$  and  $\mathbf{c}^*$  is orthogonal to  $(\mathbf{a}, \mathbf{b})$ ,

$$\mathbf{a}^* \cdot \mathbf{a} = 1, \mathbf{a}^* \cdot \mathbf{b} = 0, \mathbf{a}^* \cdot \mathbf{c} = 0 \text{ etc.} \quad (58)$$

Note that, in general,  $\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*$  are not orthogonal to each other. Positions of reciprocal-lattice points can be represented by vectors of the form

$$\mathbf{d}_{hkl}^* = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*, \quad (59)$$

where  $h, k, l$  are integers [Fig. 17(b)].

In crystallography, as the notation here implies, the physical meaning of the reciprocal lattice is related to lattice planes in the direct space as follows (Chen, 1986):

(i) Each point with a vector  $\mathbf{d}_{hkl}^*$  on the reciprocal lattice represents a family of lattice planes with Miller indices  $(hkl)$ .

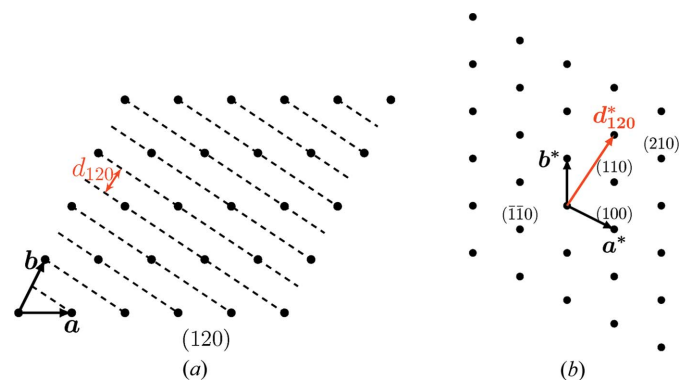


Figure 17 (a) Direct and (b) reciprocal lattices. The  $(120)$  planes (dashed lines) with interplanar distance  $d_{120}$  in direct space correspond to the point denoted by the vector  $\mathbf{d}_{120}^*$  in reciprocal space.  $\mathbf{d}_{120}^*$  is normal to the  $(120)$  planes and  $|\mathbf{d}_{120}^*| = 1/d_{120}$ .



(ii) The direction of  $\mathbf{d}_{hkl}^*$  is perpendicular to (or normal to) the lattice planes ( $hkl$ ).

(iii) The magnitude of  $\mathbf{d}_{hkl}^*$  is equal to the reciprocal of the interplanar spacing  $d_{hkl}$ , i.e.  $|\mathbf{d}_{hkl}^*| = 1/d_{hkl}$ .

## Acknowledgements

We thank Corey O'Hern, Alex Grigas, Robert Hoy and Joseph Dietz for testing some of our codes. We also thank Patrick Charbonneau for helpful discussions.

## Funding information

This work benefits from Duke Kunshan startup funding and resources made available at the Duke Compute Cluster (DCC).

## References

- Allen, M. P. & Tildesley, D. J. (1987). *Computer Simulation of Liquids*. New York: Oxford University Press.
- Argos, P. (1977). *Am. J. Phys.* **45**, 31–37.
- Ashkar, R., Bilheux, H. Z., Bordallo, H., Briber, R., Callaway, D. J. E., Cheng, X., Chu, X.-Q., Curtis, J. E., Dadmun, M., Fenimore, P., Fushman, D., Gabel, F., Gupta, K., Herberle, F., Heinrich, F., Hong, L., Katsaras, J., Kelman, Z., Kharlampieva, E., Kneller, G. R., Kovalevsky, A., Krueger, S., Langan, P., Lieberman, R., Liu, Y., Losche, M., Lyman, E., Mao, Y., Marino, J., Mattos, C., Meilleur, F., Moody, P., Nickels, J. D., O'Dell, W. B., O'Neill, H., Perez-Salas, U., Peters, J., Petridis, L., Sokolov, A. P., Stanley, C., Wagner, N., Weinrich, M., Weiss, K., Wymore, T., Zhang, Y. & Smith, J. C. (2018). *Acta Cryst.* **D74**, 1129–1168.
- Azaroff, L. V. (1968). *Elements of X-ray Crystallography*. New York: McGraw-Hill.
- Barbour, L. J. (2018). *J. Appl. Cryst.* **51**, 1734–1738.
- Barrat, J. & Hansen, J. (2003). *Basic Concepts for Simple and Complex Liquids*. New York: Cambridge University Press.
- Billinge, S. J. L. (2019). *IUCr Newslett.* **27**, [https://www.iucr.org/news/letter/et/articles?issue=145052&result\\_138339\\_result\\_page=7](https://www.iucr.org/news/letter/et/articles?issue=145052&result_138339_result_page=7).
- Bragg, L. (1968). *Sci. Am.* **219**, 58–70.
- Chaikin, P. M., Lubensky, T. C. & Witten, T. A. (1995). *Principles of Condensed Matter Physics*. Cambridge University Press.
- Chen, N.-X. (1986). *Am. J. Phys.* **54**, 1000–1002.
- Chu, B. & Hsiao, B. S. (2001). *Chem. Rev.* **101**, 1727–1762.
- Cohen, M. (1935). *Rev. Sci. Instrum.* **6**, 68–74.
- Fink, J., Schierle, E., Weschke, E. & Geck, J. (2013). *Rep. Prog. Phys.* **76**, 056502.
- Fischer, H. E., Barnes, A. C. & Salmon, P. S. (2006). *Rep. Prog. Phys.* **69**, 233–299.
- Franklin, R. E. & Gosling, R. G. (1953). *Nature*, **171**, 740–741.
- Gelasio, L. & Scardi, P. (2016). *Acta Cryst.* **A72**, 608–620.
- Goldburg, W. (1999). *Am. J. Phys.* **67**, 1152–1160.
- Hammond, C. (2001). *The Basics of Crystallography and Diffraction*. Oxford Science Publications.
- Hansen, J. & McDonald, I. R. (2013). *Theory of Simple Liquids with Applications to Soft Matter*. New York: Academic Press.
- Harrison, R. W. (1993). *J. Opt. Soc. Am. A*, **10**, 1046–1055.
- Hauptman, H. A. (1991). *Rep. Prog. Phys.* **54**, 1427–1454.
- Hayter, J. & Penfold, J. (1983). *Colloid Polym. Sci.* **261**, 1022–1030.
- Head-Gordon, T. & Hura, G. (2002). *Chem. Rev.* **102**, 2651–2670.
- Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L. II, Tsutakawa, S. E., Jenney, F. E. Jr, Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S., Scott, J. W., Dillard, B. D., Adams, M. W. W. & Tainer, J. A. (2009). *Nat. Methods*, **6**, 606–612.
- Janssen, L. (2018). *Front. Phys.* **6**, 97.
- Keen, D. A. (2001). *J. Appl. Cryst.* **34**, 172–177.
- Kendrew, J. C. (1961). *Sci. Am.* **205**, 96–110.
- Kittel, C. (1968). *Am. J. Phys.* **36**, 610–616.
- Lee, B., Park, I., Yoon, J., Park, S., Kim, J., Kim, K.-W., Chang, T. & Ree, M. (2005). *Macromolecules*, **38**, 4311–4323.
- Lighthill, M. (1958). *An Introduction to Fourier Analysis and Generalized Functions*. Cambridge University Press.
- Liu, F., Brady, M. A. & Wang, C. (2016). *Eur. Polym. J.* **81**, 555–568.
- Lucas, A. A. & Lambin, P. (2005). *Rep. Prog. Phys.* **68**, 1181–1249.
- McIntyre, G. J. (2015). *J. Phys. D Appl. Phys.* **48**, 504002.
- McQuarrie, D. A. & Simon, J. D. (1997). *Physical Chemistry: a Molecular Approach*. New York: University Science Books.
- Mitchell, D. J., Tiddy, G. J., Waring, L., Bostock, T. & McDonald, M. P. (1983). *J. Chem. Soc. Faraday Trans. 1*, **79**, 975–1000.
- Mukherjee, S., Herzog, A. A., Zhao, D., Wu, Q., Yu, L., Ade, H., DeLongchamp, D. M. & Richter, L. J. (2017). *J. Mater. Res.* **32**, 1921–1934.
- Powles, J. (1973). *Adv. Phys.* **22**, 1–56.
- Rahman, A., Singwi, K. & Sjölander, A. (1962). *Phys. Rev.* **126**, 986–996.
- Richards, R. W. & Thomason, J. (1983). *Macromolecules*, **16**, 982–992.
- Roe, R.-J. (2000). *Methods of X-ray and Neutron Scattering in Polymer Science*. New York: Oxford University Press.
- Saff, E. B. & Kuijlaars, A. B. (1997). *Math. Intelligencer*, **19**, 5–11.
- Sakurai, S., Okamoto, S., Kawamura, T. & Hashimoto, T. (1991). *J. Appl. Cryst.* **24**, 679–684.
- Sands, D. E. (1993). *Introduction to Crystallography*. New York: Dover Publications.
- Schneidman-Duhovny, D., Hammel, M. & Sali, A. (2010). *Nucleic Acids Res.* **38**, W540–W544.
- Sette, F. (1998). *Science*, **280**, 1550–1555.
- Taylor, G. (2003). *Acta Cryst.* **D59**, 1881–1890.
- Thomas, C. & Gingrich, N. (1941). *Am. J. Phys.* **9**, 10–13.
- Thomas, N. W. (2010). *Acta Cryst.* **A66**, 64–77.
- Thompson, J., Braun, G., Tierney, D., Wessels, L., Schmitzer, H., Rossa, B., Wagner, H. & Dultz, W. (2018). *Am. J. Phys.* **86**, 95–104.
- Tian, Y., Lhermitte, J. R., Bai, L., Vo, T., Xin, H. L., Li, H., Li, R., Fukuto, M., Yager, K. G., Kahn, J. S., Xiong, Y., Minevich, B., Kumar, S. K. & Gang, O. (2020). *Nat. Mater.* **19**, 789–796.
- Tutsch, U., Wolf, B., Wessel, S., Postulka, L., Tsui, Y., Jeschke, H., Opahle, I., Saha-Dasgupta, T., Valentí, R., Brühl, A., Remović-Langer, K., Kretz, T., Lerner, H. W., Wagner, M. & Lang, M. (2014). *Nat. Commun.* **5**, 5169.
- Xiong, W., Gasparian, A., Gao, H., Dutta, D., Khandaker, M., Liyanage, N., Pasyuk, E., Peng, C., Bai, X., Ye, L., Gnanvo, K., Gu, C., Levillain, M., Yan, X., Higinbotham, D. W., Mezziane, M., Ye, Z., Adhikari, K., Aljawrneh, B., Bhatt, H., Bhetuwal, D., Brock, J., Burkert, V., Carlin, C., Deur, A., Di, D., Dunne, J., Ekanayaka, P., El-Fassi, L., Emmich, B., Gan, L., Glamazdin, O., Kabir, M. L., Karki, A., Keith, C., Kowalski, S., Lagerquist, V., Larin, I., Liu, T., Liyanage, A., Maxwell, J., Meekins, D., Nazeer, S. J., Nelyubin, V., Nguyen, H., Pedroni, R., Perdrisat, C., Pierce, J., Punjabi, V., Shabestari, M., Shahinyan, A., Silwal, R., Stepanyan, S., Subedi, A., Tarasov, V. V., Ton, N., Zhang, Y. & Zhao, Z. W. (2019). *Nature*, **575**, 147–150.