Study of non-Bragg scattering for crystals with substitution disorder: a theoretical approach

2007 J. Phys.: Condens. Matter 19 236232

(http://iopscience.iop.org/0953-8984/19/23/236232)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **19** (2007) 236232 (28pp) doi:10.1088/0953-8984/19/23/236232

Study of non-Bragg scattering for crystals with substitution disorder: a theoretical approach

N Lei

Center for Advanced Radiation Sources, The University of Chicago, 5640 S Ellis Avenue, Chicago, IL 60637, USA

Received 25 March 2007, in final form 26 April 2007
Published 17 May 2007
Online at stacks.iop.org/JPhysCM/19/236232

Abstract

We derive the mathematical formalism for the ensemble averages of the non-Bragg scattering differential cross sections for crystals of statistically uniformly randomly distributed substitution disorder. In addition, we study the intensity fluctuations from their ensemble averages. The fluctuations are shown to be extremely large and identical to the averaged intensities themselves. However, we demonstrate mathematically that when spectrometer resolutions are not sharp enough, which is typical, the fluctuations are suppressed greatly by the resolutions and are proportional to the square root of the total number of scattering molecules. Computer Monte Carlo simulations are carried out to study the non-Bragg scattering intensity fluctuations, and the simulation results are consistent with our theoretical predictions. Furthermore, we obtain the noise levels in determining the electron density distribution from both the non-Bragg and Bragg scattering. A sharpened Patterson map is proposed to be used to solve the overlap of the non-Bragg scattering for the case that a unit cell contains multiple molecules which are of different angular orientations.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

It is of deep interest to many scientists to know the instantaneous structure of a molecule when it is in a transient state due to reasons such as temperature changes, photo excitations, chemical reactions, and biological processes. In the scope of using crystallographic techniques, the interest is focused on those molecules which can be photo excited. Traditionally, the electron density profile for a transient molecule is obtained through the difference Fourier map [1, 2], from the often prominent x-ray Bragg peak intensities. The difference Fourier map method is an approximation, and the systematic error created by using it [3] may subsequently be corrected by model building refinement [4]. However, if the systematic error is large, it can be imagined that the model building process can be difficult. It is important to supply an accurate electron density profile before any model refinement. Do we have other means to provide the electron

density profile which may have a smaller systematic error? It is known that, for an excited crystal, in addition to the Bragg peaks, there exists diffuse scattering which we denote as non-Bragg scattering [5]. Therefore, it is natural to ask the question of whether the non-Bragg scattering can be used to solve the structure of the excited molecules, and if the answer is yes, at what level of *signal-to-noise ratio* (SNR). In this paper, we shall address these questions.

For a photo excited crystal, the non-Bragg scattering arises from the substitution disorder developed by the excitation. An example of a crystal with substitution disorder is a photo-active yellow protein crystal illuminated by a laser pulse [6]. In this case, the majority of the molecules are in ground state and a portion of the molecules are excited by the photons. For not-too-small-sized crystals, the Bragg scattering is typically negligible at off-Bragg peak positions where the non-Bragg scattering may be large enough to be detected. The Bragg scattering is described by the average electron density across all unit cells (unit cell in this paper always means a primitive unit cell) and the non-Bragg scattering is described by the deviation from the average. For a unit cell containing only one molecule, a mathematical expression for the non-Bragg scattering intensity distribution has been given in classic x-ray scattering books of Guinier [5] and Warren [7].

Retrieval of an electron density profile, from the x-ray scattering intensity distribution of a non-periodic object, has been carried out very successfully through over-sampling [8–13]. The non-Bragg scattering from a photo excited crystal is similar to that from a single non-periodic object. For the case of one molecule per unit cell, the ensemble averaged non-Bragg scattering is simply the scattering of a pseudo-molecule which is the difference, in electron density, of that of the excited molecule and that of the molecule in the ground state, amplified by $Np(1-p)$, where N is the total number of molecules in the crystal and p is the probability of the molecule in the excited state [5, 7]. Therefore, if the non-Bragg scattering is strong enough to be detectable, the electron density profile for the excited molecule may be obtained by applying iterative algorithms [14].

To obtain the pseudo-molecule electron density profile, a very challenging complication arises when there are multiple molecules per unit cell. This complication is due to the overlap of the non-Bragg scattering intensities from the excited molecules which are in different orientations. Mathematically, if not considering the atomicity of the atoms, due to the overlap, there is an infinite number of solutions. Using a vector verification method with a sharpened Patterson function (section 2.5), the corresponding pseudo-molecular structure may be identified from the overlapped total non-Bragg scattering.

In this paper, we are going to obtain mathematical expressions for the x-ray non-Bragg scattering intensity distributions, in the sense of ensemble averages, for a statistically uniformly randomly distributed substitution disordered molecular crystal for the case of one molecule per unit cell, and the general case of multiple molecules per unit cell with multiple excited states per molecule. In addition, we shall obtain mathematical expressions for the non-Bragg scattering intensity fluctuations. Knowing the fluctuations is important since we want to know how many systems one needs to measure in order to be statistically within a certain percentage of the ensemble averages. What we have found, amazingly, is that the fluctuations are large, identical to the averages themselves, indicating that in order to be statistically within 10% of the ensemble averages, one needs to carry out 100 complete measurements, which is extremely time consuming. The large fluctuations prompt us to consider the effect of spectrometer resolutions. We shall show that when the spectrometer resolution widths are relatively large, the fluctuations are suppressed greatly, proportional only to the square root of the total number of x-ray illuminated unit cells.

We shall first develop, in section 2, the mathematical expressions of the x-ray non-Bragg scattering intensity distributions, for a molecular crystal with statistically uniform random

substitution disorder. In section 3, we shall formulate the non-Bragg scattering intensity fluctuations, and shall show the results of computer Monte Carlo simulations to support our theoretical predications. We shall give theories, in section 4, for the noise levels in determining the electron density profiles for the cases of using the non-Bragg scattering and the Bragg scattering difference Fourier map. Finally in section 5, we shall summarize our results.

2. Non-Bragg scattering intensity distributions

2.1. One excited state and one molecule per unit cell

The x-ray scattering structure factor for an excited crystal in the far field may be expressed as

$$F(\mathbf{q}) = \int_{\text{sample}} \left[\sum_{j=1}^{N} \rho(\mathbf{r} - \mathbf{R}_j) + \sum_{k=1}^{K} \Delta\rho(\mathbf{r} - \mathbf{R}_{n(k)}) \right] \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}} \mathrm{d}\mathbf{r}, \tag{2.1.1}$$

where $\mathbf{q}$ is the x-ray wavevector transfer in the scattering process, $\rho(\mathbf{r})$ is the electron density distribution of the molecule in its ground state, $\mathbf{R}_j$ is the location of the jth unit cell and N is the total number of unit cells. In (2.1.1), $\Delta\rho(\mathbf{r})$ is the electron density distribution deviation from $\rho(\mathbf{r})$ when the molecule is in its excited state, K is the number of excited molecules in the crystal, and $n(k)$ is the unit cell index for the kth excited molecule. The Bragg scattering structure factor for the crystal can be defined as

$$F_{\text{B}}(\mathbf{q}) = \int_{\text{sample}} \sum_{j=1}^{N} \left(\rho(\mathbf{r} - \mathbf{R}_j) + \frac{K}{N} \Delta\rho(\mathbf{r} - \mathbf{R}_j) \right) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}} \mathrm{d}\mathbf{r}, \tag{2.1.2}$$

and the non-Bragg scattering structure factor for the crystal is defined as the total scattering structure factor subtracting the Bragg scattering structure factor, namely,

$$F_{\text{NB}}(\mathbf{q}) = \int_{\text{sample}} \left[\sum_{k=1}^{K} \Delta\rho(\mathbf{r} - \mathbf{R}_{n(k)}) - \sum_{j=1}^{N} \frac{K}{N} \Delta\rho(\mathbf{r} - \mathbf{R}_j) \right] \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}} \mathrm{d}\mathbf{r}. \tag{2.1.3}$$

The Bragg scattering structure factor contains a factor whose absolute value squared is the Bragg scattering lattice factor, defined as

$$L_{\text{B}}(\mathbf{q}) = \left(\frac{\sin(N_a q_a a/2)}{\sin(q_a a/2)} \right)^2 \left(\frac{\sin(N_b q_b b/2)}{\sin(q_b b/2)} \right)^2 \left(\frac{\sin(N_c q_c c/2)}{\sin(q_c c/2)} \right)^2, \tag{2.1.4}$$

where $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ are the unit cell base vectors, and $(N_a\mathbf{a}, N_b\mathbf{b}, N_c\mathbf{c})$ are the sizes of the crystal. Due to this lattice factor, the Bragg scattering is relatively very small at off-Bragg peak positions.

From (2.1.3), it is easy to see that at Bragg peak positions, the non-Bragg scattering structure factor is zero. At off-Bragg peak positions, the second term on the right-hand side of (2.1.3) can be ignored since the term is comparatively very small, and the non-Bragg scattering structure factor is simply

$$F_{\text{NB}}(\mathbf{q}) = \int_{\text{molecule}} \Delta\rho(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}} \, \mathrm{d}\mathbf{r} \sum_{k=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)}}. \tag{2.1.5}$$

Therefore, ignoring x-ray polarization and absorption, the non-Bragg scattering differential cross section at off-Bragg peak positions is

$$\frac{\mathrm{d}\sigma_{\text{NB}}}{\mathrm{d}\Omega}(\mathbf{q}) = r_0^2 \left| \int_{\text{molecule}} \Delta\rho(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}} \mathrm{d}\mathbf{r} \right|^2 \left| \sum_{k=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)}} \right|^2. \tag{2.1.6}$$

In (2.1.6), r_0 is the classical electron radius with $r_0^2 = 7.94 \times 10^{-10}$ Å^2.

From (2.1.6), we can see that the non-Bragg scattering is a product of two terms, with one depending only on the change of the electron density profile and the other depending only on which molecules are excited. For simplicity, we define the difference molecular scattering factor as

$$\Delta M(\mathbf{q}) = \left| \int_{\text{molecule}} \Delta\rho(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}} \mathrm{d}\mathbf{r} \right|^2 \tag{2.1.7}$$

and the non-Bragg scattering lattice factor for the case of one molecule per unit cell as

$$L_{\text{NB}}(\mathbf{q}) = \left| \sum_{k=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)}} \right|^2 . \tag{2.1.8}$$

Therefore, (2.1.6) becomes

$$\frac{\mathrm{d}\sigma_{\text{NB}}}{\mathrm{d}\Omega}(\mathbf{q}) = r_0^2 \Delta M(\mathbf{q}) L_{\text{NB}}(\mathbf{q}). \tag{2.1.9}$$

To simplify the right-hand side of (2.1.8), we realize that in (2.1.8), each $\exp(-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)})$ in the series of $\{\exp(-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)})\}$ represents a random walk in a complex plane with step size one and a random direction if $\mathbf{q}$ is away from the Bragg peaks and $K/N \gg 1$. Therefore it is expected that $\langle|\sum_{k=1}^{K} \exp(-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)})|^2\rangle = \langle K\rangle$. $\langle\cdots\rangle$ indicates an ensemble average. However, the random walk is not completely random since the same site cannot be walked on again, namely $n(k)$ cannot be the same for different k. To handle this case, we use

$$\left| \sum_{k=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k)}} \right|^2 = K + \sum_{k\neq k'}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})}, \tag{2.1.10}$$

and

$$\left\langle \sum_{k\neq k'}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})} \right\rangle_{\text{suben}} = K(K-1) \left\langle \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})} \right\rangle_{k\neq k',\text{suben}} . \tag{2.1.11}$$

In (2.1.11), the subscript suben indicates the subensemble in which every sample has K molecules excited. Since, in the subensemble, $\mathbf{R}_{n(k)}$ can go through each unit cell in the crystal and $\mathbf{R}_{n(k')}$ can go through the remaining $N-1$ unit cells if $k \neq k'$, we have (the more vigorous mathematical formalism developed in appendix B can be used here).

$$\left\langle \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})} \right\rangle_{k\neq k',\ \text{suben}} = \frac{1}{N(N-1)} \left(\sum_{k,k'}^{N} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_k - \mathbf{R}_{k'})} - N \right). \tag{2.1.12}$$

The first term in the bracket of the right-hand side of (2.1.12) is simply the Bragg scattering lattice factor, and it can be ignored when $\mathbf{q}$ is at off-Bragg peak positions. Therefore, at off-Bragg peak positions, the right-hand side is $-1/(N-1)$, and (2.1.11) becomes

$$\left\langle \sum_{k\neq k'}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})} \right\rangle_{\text{suben}} = -K(K-1)/(N-1). \tag{2.1.13}$$

Therefore the ensemble average of the non-Bragg scattering lattice factor is $\langle K\rangle - \langle K(K-1)/(N-1)\rangle$, namely

$$\langle L_{\text{NB}}(\mathbf{q})\rangle = Np(1-p), \tag{2.1.14}$$

where p is the molecular excitation probability, and the non-Bragg scattering differential cross section is

$$\left\langle \frac{\mathrm{d}\sigma_{\text{NB}}}{\mathrm{d}\Omega}(\mathbf{q}) \right\rangle = r_0^2 Np(1-p)\Delta M(\mathbf{q}). \tag{2.1.15}$$

Equation (2.1.15) is the well-known Laue formula [5, 7].

It is evident from (2.1.7) and (2.1.15) that $\Delta\rho(\mathbf{r})$ and $\pm\Delta\rho(\mathbf{r}-\mathbf{r}_1)$, for any $\mathbf{r}_1$, generate the same non-Bragg scattering. This may create problems for uniqueness in determining $\Delta\rho(\mathbf{r})$, and the knowledge of $\rho(\mathbf{r})$ may be used to break the ambiguity. In addition, $\Delta\rho(\mathbf{r})$ and $\pm\Delta\rho(-\mathbf{r})$ also create the same non-Bragg scattering [13].

Note that since $\int \Delta\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} = 0$, when q is small enough that $\mathbf{qr} \ll 1$, (2.1.15) becomes

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 Np(1-p)q^2 \left| \int_{\text{molecule}} \Delta\rho(\mathbf{r})\hat{q}\mathbf{r}\,\mathrm{d}\mathbf{r} \right|^2, \tag{2.1.16}$$

where $\hat{q}$ is a unit vector of $\mathbf{q}$, indicating that $\langle \mathrm{d}\sigma_{\mathrm{NB}}/\mathrm{d}\Omega\rangle/q^2$ is independent of q at small q. This behaviour is very different from that of a liquid whose scattering intensity approaches a constant when q becomes small [15]. Moreover, (2.1.16) shows that there is a significant difference in the non-Bragg scattering intensity distribution at a small q for $\Delta\rho(\mathbf{r})$ and $\Delta\rho(\mathbf{r}) + \rho_0$, both of which create the same non-Bragg scattering at a large q.

The non-Bragg scattering's non-linearity dependence on the excitation probability p, as shown in (2.1.15), offers a way to obtain the probability. If the probability is varied from p to np, achieved, for example, through increasing the trigger laser power by a factor of n, (2.1.15) becomes

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega}(\mathbf{q}) \right\rangle = r_0^2 Nnp\Delta M(\mathbf{q}) - r_0^2 Nn^2p^2\Delta M(\mathbf{q}). \tag{2.1.17}$$

The coefficients of n and n^2 can be obtained through fitting the measured intensity (integrated over reciprocal space to improve the SNR) versus n to (2.1.17), and the ratio of these two coefficients gives the value of the excitation probability.

2.2. Any number of excited states, with each unit cell containing any number of molecules

For the case of more than one excited state per molecule, the substitution disorder created non-Bragg scattering differential cross section contains an interference term. Detailed derivations are given in appendix A. Here we just list the results.

(a) Any number of excited states, with each unit cell having one molecule.

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 \sum_{n=1}^{N_{\mathrm{e}}} Np_n|\Delta F_n(\mathbf{q})|^2 - \frac{r_0^2}{N}\left| \sum_{n=1}^{N_{\mathrm{e}}} Np_n\Delta F_n(\mathbf{q}) \right|^2, \tag{2.2.1}$$

where $\Delta F_n(\mathbf{q})$ is the Fourier transform of $\Delta\rho_n(\mathbf{r})$, which is the electron density difference between the molecule in its nth excited state and the ground state.

(b) One excited state per molecule, with each unit cell having multiple molecules.

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 Np(1-p) \sum_{m=1}^{N_{\mathrm{m}}} \Delta M_m(\mathbf{q}), \tag{2.2.2}$$

where $\Delta M_m(\mathbf{q})$ is the difference molecule scattering factor for the mth molecule in a unit cell.

(c) Any number of excited states per molecule, with each unit cell having any number of molecules.

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 \sum_{m=1}^{N_{\mathrm{m}}} \left[\left(\sum_{n=1}^{N_{\mathrm{e}}} Np_n\Delta M_{m,n}(\mathbf{q}) \right) - \frac{1}{N}\left| \sum_{n=1}^{N_{\mathrm{e}}} Np_n\Delta F_{m,n}(\mathbf{q}) \right|^2 \right]. \tag{2.2.3}$$

In the expressions above, N_{e} is the number of excited states per molecule, N_{m} is the number of molecules per unit cell, p_n is the probability for a molecule being at its nth excited state, $\Delta F_{m,n}$ is the Fourier transform of $\rho_{m,n}(\mathbf{r}) - \rho_m(\mathbf{r})$, with $\rho_{m,n}(\mathbf{r})$ and $\rho_m(\mathbf{r})$ the electron densities

for the mth molecule (in a unit cell) in its nth excited and ground states, respectively, and $\Delta M_{m,n} = |\Delta F_{m,n}|^2$. For case (b) (only one excited state), to determine each $\Delta M_m(\mathbf{q})$ from the experimental data, assuming that the excitation probabilities are small, we may use a Patterson vector verification method (see section 2.5).

2.3. Temperature effect on substitution disorder created non-Bragg scattering

At a non-zero temperature, the molecules may move around their equilibrium positions as well as the atoms within the molecules. The correlation among the motions of the atoms in a single unit cell creates diffuse scattering [16]. Additional motion correlation of unit cells creates the typical thermal diffuse scattering [7] for crystals. It is reasonable to assume that the diffuse scattering unrelated to photo exitations does not change much before and after a photo excitation. Beyond the thermal diffuse scattering, what happens to the substitution disorder created non-Bragg scattering due to motions of the unit cells?

Assuming that the molecules do not rotate and are rigid, the unit cell motion (no rotation) affected substitution disorder created non-Bragg scattering structure factor can be described by

$$F_{\mathrm{NB}}(\mathbf{q}) = \int_{\text{molecule}} \Delta\rho(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}}\, \mathrm{d}\mathbf{r} \sum_{k=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)}+\boldsymbol{\delta}_{n(k)})}, \tag{2.3.1}$$

where $\boldsymbol{\delta}_{n(k)}$ represents the deviation from its equilibrium position at a particular time for the molecule of index $n(k)$. Therefore, the non-Bragg scattering lattice factor is

$$L_{\mathrm{NB}}(\mathbf{q}) = \left| \sum_{k=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)}+\boldsymbol{\delta}_{n(k)})} \right|^2. \tag{2.3.2}$$

We ignore the correlation among and assume a Gaussian probability for $\boldsymbol{\delta}_{n(k)}$, and in addition we assume that $\langle(\mathbf{q}\boldsymbol{\delta}_{n(k)})^2\rangle$ is the same for all molecules. Taking an ensemble average on (2.3.2), we have

$$\langle L_{\mathrm{NB}}(\mathbf{q})\rangle = \left\langle \sum_{k\neq k'}^{K} \mathrm{e}^{\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)}-\mathbf{R}_{n(k')})} \right\rangle \mathrm{e}^{-\langle(\mathbf{q}\boldsymbol{\delta})^2\rangle} + \langle K\rangle. \tag{2.3.3}$$

From (2.1.16), we know that (2.3.3) is just

$$\langle L_{\mathrm{NB}}(\mathbf{q})\rangle = Np - Np^2 \mathrm{e}^{-\langle(\mathbf{q}\boldsymbol{\delta})^2\rangle}. \tag{2.3.4}$$

For the general case, it is easy to show that the substitution disorder created non-Bragg scattering differential cross section with temperature effect is, assuming that each molecule's displacement from its equilibrium position is uncorrelated to others,

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 \sum_{m=1}^{N_{\mathrm{m}}} \left[\left(\sum_{n=1}^{N_{\mathrm{e}}} N p_n \Delta M_{m,n}(\mathbf{q}) \right) - \frac{1}{N} \left| \sum_{n=1}^{N_{\mathrm{e}}} N p_n \Delta F_{m,n}(\mathbf{q}) \right|^2 \mathrm{e}^{-\langle(\mathbf{q}\boldsymbol{\delta}_m)^2\rangle} \right], \tag{2.3.5}$$

where $\boldsymbol{\delta}_m$ is the displacement from its equilibrium position for the mth molecule in a unit cell. Equation (2.3.5) shows that the thermal motion does not affect the main non-Bragg scattering term $\sum_{m=1}^{N_{\mathrm{m}}} \sum_{n=1}^{N_{\mathrm{e}}} N p_n \Delta M_{m,n}(\mathbf{q})$, unlike the case for the Bragg peak intensities which decrease dramatically once q is over a threshold due to the Debye–Waller thermal motion factors.

2.4. Is the non-Bragg scattering experimentally measurable?

To use the substitution disorder created non-Bragg scattering, it is crucial to know theoretically whether, amid the noise created by the solvent in the case of a biological sample and air scattering, the non-Bragg scattering is detectable. In this subsection, we shall answer this question with a simplified example.

We assume that a single carbon atom moves from $(-1, 0, 0)$ to $(1, 0, 0)$ during an excitation. The unit cell is a cube with a size of 50^3 Å^3, containing only one molecule. We assume that the entire crystal has a size of 8×10^{18} Å^3, indicating that the sample contains 6.4×10^{13} unit cells. The excitation probability is 0.1. Ignoring x-ray polarization and absorption, the substitution disorder created non-Bragg scattering intensity is close to

$$\langle I_{\mathrm{NB}}(\mathbf{q})\rangle = 1.8 \times 10^4 I_0 \sin^2 q_x \, |F_{\mathrm{C}}(q)|^2 \, \Delta\Omega\Delta t, \tag{2.4.1}$$

where $F_{\mathrm{C}}(q)$ is the carbon atom x-ray scattering factor, $\Delta\Omega$ is the solid angle that the detection area (typically the area of a pixel or binned pixels) makes with the sample, Δt is the time duration of measurement, and I_0 is the incoming x-ray intensity per unit time per Å^2 (we assume that the sample is completely covered by the incoming x-rays).

We only consider the noise from solvent scattering and treat the solvent as water. Assuming that half of the sample volume is water, there are nearly 2.4×10^{17} water molecules. Therefore, the scattering intensity from the water is

$$\langle I_{\mathrm{water}}\rangle = 1.9 \times 10^8 I_0 I_{\mathrm{water\ molecule}}(q)\Delta\Omega\Delta t. \tag{2.4.2}$$

In (2.4.2), $I_{\mathrm{water\ molecule}}(q)$ is the water molecular x-ray scattering differential cross-section, whose value ranges from about 100 electron units at $q = 0$ to about 600 electron units at its peak location of $q = 2.0$ Å^{-1} [17]. The average noise from the water scattering is simply $\sqrt{\langle I_{\mathrm{water}}\rangle}$. To be able to measure $\langle I_{\mathrm{NB}}(\mathbf{q})\rangle$ accurately, $\langle I_{\mathrm{NB}}(\mathbf{q})\rangle$ has to be much larger than $\sqrt{\langle I_{\mathrm{water}}\rangle}$. For a beamline of incoming x-ray photon intensity of 10^{13} s^{-1} over an area of 0.01 mm^2, $I_0 = 10$ Å^{-2} s^{-1}. For a solid angle $\Delta\Omega = 1.0 \times 10^{-3}$, $\langle I_{\mathrm{NB}}(\mathbf{q})\rangle = 180 \times \Delta t |F_{\mathrm{C}}(q)|^2 \sin^2 q_x$ and $\langle I_{\mathrm{water}}\rangle = 1.9 \times 10^6 I_{\mathrm{watermolecule}}(q)\Delta t$. If we set, at $q = 2.0$ Å^{-1}, $\langle I_{\mathrm{NB}}(\mathbf{q})\rangle > \sqrt{\langle I_{\mathrm{water}}\rangle}$, and take $q_x = 2$ Å^{-1}, meaning $F_{\mathrm{C}}(q_x) = 4.1$, we have $\Delta t > 182$ s, which is much longer than the time scales associated with a typical biological time-dependent process (further analysis in section 4.2 seems to suggest that even though the non-Bragg scattering data are very noisy, the reconstructed electron density profile may be good enough). It is obvious that the solvent scattering is detrimental to using the non-Bragg scattering to extract time-dependent structural changes. Therefore, the usage of substitution disorder created non-Bragg scattering is likely limited to (a) crystals without solvent, such as those formed by chemical compounds, and (b) shining the trigger laser continuously on to the sample so that the excitation probabilities are not time variants and the data collection time duration can be as long as desired.

2.5. Patterson vector verification method to solve non-Bragg scattering overlap

We give an example to illustrate how to use a Patterson map to solve the overlap problem in (2.2.2) for the case of multiple molecules in a unit cell. We draw in figure 1(a) two atoms (in a molecule) which move after an excitation from their original positions denoted by circles to new locations denoted by filled circles. We assume that there are four molecules in a unit cell which has a four-fold symmetry. The Patterson peaks from the total non-Bragg scattering are shown in figure 1(b), with filled circles indicating positive peaks and non-filled circles indicating negative peaks. In generating the Patterson peaks, we can use a sharpened Patterson map scheme described in appendix C.

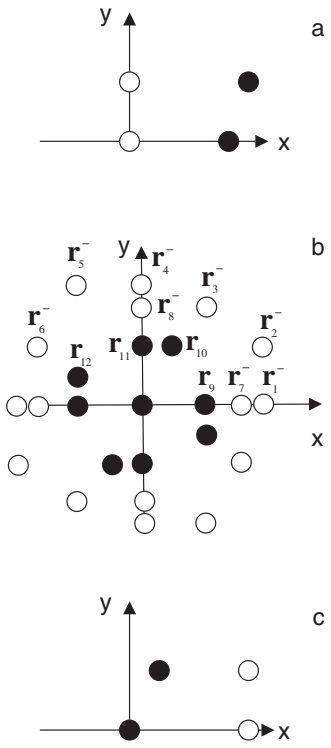

Figure 1. (a) Atom locations. The original locations are represented by circles and the locations after excitation are represented by filled circles. (b) Patterson map for atoms in (a). (c) Obtained atom locations from the Patterson map through the procedure described in section 2.5.

We put the Patterson peak vectors in the upper plane into a set denoted by P (for peaks fall on the x axis, we select only those with $x > 0$), namely $P = \{\mathbf{r}_1^-, \mathbf{r}_2^-, \mathbf{r}_3^-, \mathbf{r}_4^-, \mathbf{r}_5^-, \mathbf{r}_6^-, \mathbf{r}_7^-, \mathbf{r}_8^-, \mathbf{r}_9, \mathbf{r}_{10}, \mathbf{r}_{11}, \mathbf{r}_{12}\}$. The superscript '$-$' indicates that the Patterson peak is negative. Because a peak in a Patterson map corresponds to a vector starting from one atom to the next in a molecule, we can assign (0, 0) as the location of one atom and one of the vectors in P as the location of another atom. We first take away the largest vector in P, say $\mathbf{r}_2^-$, and put $\mathbf{r}_2^-$ in a solution set called S (the vectors in S are the locations of atoms). Now we have $S = \{(0, 0), \mathbf{r}_2^-\}$. We look for symmetry elements of $\mathbf{r}_2^-$ among the vectors in set P. We find that $\mathbf{r}_5^-$ is a symmetry vector of $\mathbf{r}_2^-$. Therefore, we remove $\mathbf{r}_5^-$ from P and have $P = \{\mathbf{r}_1^-, \mathbf{r}_3^-, \mathbf{r}_4^-, \mathbf{r}_6^-, \mathbf{r}_7^-, \mathbf{r}_8^-, \mathbf{r}_9, \mathbf{r}_{10}, \mathbf{r}_{11}, \mathbf{r}_{12}\}$.

Next, we pick the largest vector $\mathbf{r}_1^-$ in P. To determine whether $\mathbf{r}_1^-$ is an atom location, we need to check whether P contains all the vectors which are from $\mathbf{r}_1^-$ to known atom locations specified in set S or from a vector in set S to $\mathbf{r}_1^-$. At this moment, except for the origin, there is only one atom located at $\mathbf{r}_2^-$ in the solution set S. Therefore we only need to check $\mathbf{r}_2^- - \mathbf{r}_1^-$. We find that $\mathbf{r}_2^- - \mathbf{r}_1^- = \mathbf{r}_{10}$ is in set P. Therefore $\mathbf{r}_1^-$ is an atom location, and we remove it from P and put it into S. $\mathbf{r}_4^-$ is removed from set P since $\mathbf{r}_4^-$ is a symmetry element of $\mathbf{r}_1^-$. $\mathbf{r}_{11}$ is removed from set P since $\mathbf{r}_{11} = \mathbf{r}_2^- - \mathbf{r}_1^-$. Since $\mathbf{r}_9$ is a symmetry element of $\mathbf{r}_{11}$, we remove $\mathbf{r}_9$ from P as well. At this point, $P = \{\mathbf{r}_3^-, \mathbf{r}_6^-, \mathbf{r}_7^-, \mathbf{r}_8^-, \mathbf{r}_{10}, \mathbf{r}_{12}\}$ and $S = \{0, \mathbf{r}_1^-, \mathbf{r}_2^-\}$.

We now examine $\mathbf{r}_3^-$ (we try first not to pick those vectors in P which may be the symmetry elements of the vectors in the first quadrant). $\mathbf{r}_3^-$ is rejected as an atom location since P does

not contain vector $\mathbf{r}_3^- - \mathbf{r}_1^-$. Similarly, $\mathbf{r}_7^-$ is also rejected. Now we examine $\mathbf{r}_{10}$. We examine the vectors in the solution set S to $\mathbf{r}_{10}$ or from $\mathbf{r}_{10}$ to those vectors in S. We find that both $\mathbf{r}_2^- - \mathbf{r}_{10} = \mathbf{r}_7^-$ and $\mathbf{r}_{10} - \mathbf{r}_1^- = \mathbf{r}_6^-$ are in set P. Therefore $\mathbf{r}_{10}$ is an atom location and is moved from P to S. $\mathbf{r}_{12}$ is a symmetry element of $\mathbf{r}_{10}$ and is removed from P. Therefore, at this point we have $P = \{\mathbf{r}_3^-, \mathbf{r}_6^-, \mathbf{r}_7^-, \mathbf{r}_8^-\}$ and $S = \{0, \mathbf{r}_2^-, \mathbf{r}_1^-, \mathbf{r}_{10}\}$. Since $\mathbf{r}_8^-$ is a symmetry element of the non-atom location vector $\mathbf{r}_7^-$, and $\mathbf{r}_6^-$ is a symmetry element of the non-atom location vector $\mathbf{r}_3^-$, we remove $\mathbf{r}_3^-, \mathbf{r}_6^-, \mathbf{r}_7^-$, and $\mathbf{r}_8^-$ from set P. At this point set P is empty and the solution $S = \{0, \mathbf{r}_2^-, \mathbf{r}_1^-, \mathbf{r}_{10}\}$. Figure 1(c) shows the solution atom locations which are the origin inverted plus a shift along the x axis from the original one shown in figure 1(a).

3. Differential cross section fluctuation of the non-Bragg scattering

Using the approach developed in section 2, it is possible to calculate the non-Bragg scattering differential cross section deviation for a particular system from the ensemble average. We shall consider both cases of infinitely sharp and the realistic not-sharp-enough spectrometer resolutions.

3.1. Infinitely sharp spectrometer resolutions

In this subsection, we consider the fluctuation of the scattering differential cross section due to the uncertainties in the distribution of the excited molecules $n(k)$ and the total number of excited molecules K under the condition of constant trigger laser power.

We need to know $\langle (L_{\mathrm{NB}}(\mathbf{q}))^2 \rangle$. It is easier to first consider the subensemble average, with the subensemble as the one for systems having the same number of excited molecules. The ensemble average is simply the average of the subensemble averages. Intuitively $\langle (L_{\mathrm{NB}}(\mathbf{q}))^2 \rangle_{\mathrm{suben}}$ can be very large, as can be seen from (2.1.10). On the right-hand side of (2.1.10), K is about the average of the non-Bragg scattering lattice factor, and $\sum_{k \neq k'}^{K} \exp\left[\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})\right]$ is basically the deviation from the average. Using a random walk argument, this deviation has a magnitude of $\sqrt{K(K-1)} \approx K$ equaling the subensemble average. The detailed mathematical derivation in appendix B.1 shows that

$$\langle L_{\mathrm{NB}}^2(\mathbf{q}) \rangle_{\mathrm{suben}} = 2K^2 \left(1 - \frac{K}{N}\right)^2. \tag{3.1.1}$$

Since K follows a binomial distribution, we have

$$\langle (L_{\mathrm{NB}}(\mathbf{q}))^2 \rangle - (\langle L_{\mathrm{NB}}(\mathbf{q}) \rangle)^2 = (Np)^2(1-p)^2, \tag{3.1.2}$$

namely

$$\frac{\sqrt{\langle (L_{\mathrm{NB}}(\mathbf{q}) - \langle L_{\mathrm{NB}}(\mathbf{q}) \rangle)^2 \rangle}}{\langle L_{\mathrm{NB}}(\mathbf{q}) \rangle} = 1. \tag{3.1.3}$$

For the more general case of multiple molecules per unit cell, coupled with multiple excited states per molecule, (3.1.3) also holds true (see appendix B.3). Equation (3.1.3) suggests that the non-Bragg scattering intensity fluctuation average is large without spectrometer resolution considerations. Equation (3.1.3) is the result that the excited molecules are uniformly randomly distributed. For general disordered materials, the x-ray diffuse scattering differential cross section fluctuation is related to how the atoms are correlated spatially statistically [18].

3.2. Spectrometer resolution widths $>1/$sample *linear sizes*

The actual x-ray scattering differential cross section is the convolution of the ideal scattering differential cross section and the corresponding spectrometer resolution function. We first consider the simple case of one molecule per unit cell and one excited state per molecule. With spectrometer resolution considerations, the real non-Bragg scattering differential cross section can be described by

$$\frac{\mathrm{d}\sigma_{\mathrm{NB,Re}}}{\mathrm{d}\Omega} = r_0^2 \int_\infty \Delta M(\mathbf{q}+\mathbf{q}') L_{\mathrm{NB}}(\mathbf{q}+\mathbf{q}') R(\mathbf{q}')\, \mathrm{d}\mathbf{q}', \tag{3.2.1}$$

where R represents the spectrometer resolution function.

$\Delta M(\mathbf{q})$ is nearly a constant over the range of the resolution widths, which are often 0.001 Å^{-1} or smaller. Therefore (3.2.1) can be written as

$$\frac{\mathrm{d}\sigma_{\mathrm{NB,Re}}}{\mathrm{d}\Omega} = r_0^2 \Delta M(\mathbf{q}) \int_\infty L_{\mathrm{NB}}(\mathbf{q}+\mathbf{q}') R(\mathbf{q}')\, \mathrm{d}\mathbf{q}'. \tag{3.2.2}$$

Since $\langle L_{\mathrm{NB}}(\mathbf{q})\rangle$ is independent of $\mathbf{q}$, (3.2.2) indicates that the non-Bragg scattering ensemble average is irrelevant to the spectrometer resolution, namely

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB,Re}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 \Delta M(\mathbf{q}) \langle L_{\mathrm{NB}} \rangle. \tag{3.2.3}$$

However, the ensemble average of the square of the non-Bragg scattering is spectrometer resolution dependent.

Denoting the resolution convoluted non-Bragg scattering lattice factor as $L_{\mathrm{NB,Re}}(\mathbf{q})$, we have

$$\langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2 \rangle = \int_\infty \int_\infty \langle L_{\mathrm{NB}}(\mathbf{q}+\mathbf{q}_1') L_{\mathrm{NB}}(\mathbf{q}+\mathbf{q}_2') \rangle R(\mathbf{q}_1') R(\mathbf{q}_2')\, \mathrm{d}\mathbf{q}_1'\, \mathrm{d}\mathbf{q}_2'. \tag{3.2.4}$$

Since when $\mathbf{q}_1' \neq \mathbf{q}_2'$, $\langle L_{\mathrm{NB}}(\mathbf{q}+\mathbf{q}_1') L_{\mathrm{NB}}(\mathbf{q}+\mathbf{q}_2') \rangle$ is not the same as $\langle L_{\mathrm{NB}}(\mathbf{q}) L_{\mathrm{NB}}(\mathbf{q}) \rangle$, $\langle (L_{\mathrm{NB}}(\mathbf{q}))^2 \rangle \neq \langle L_{\mathrm{NB}}^2(\mathbf{q}) \rangle$ and therefore the non-Bragg scattering intensity fluctuation ensemble average differs between infinitely sharp and non-infinitely sharp spectrometer resolutions.

The detailed mathematical derivation in appendix B.4 shows that

$$\langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2 \rangle_{\mathrm{suben}} = \left(1 + \frac{(2\pi)^{d/2}}{V \sigma_x \sigma_y \sigma_z}\right) (\langle L_{\mathrm{NB}}(\mathbf{q}) \rangle_{\mathrm{suben}})^2, \tag{3.2.5}$$

where d is the dimension of the sample and V is the x-ray illuminated sample volume. σ_x, σ_y, and σ_z are the Gaussian half widths of the spectrometer resolution function in reciprocal space along q_x, q_y, and q_z, respectively. From (3.2.5), we have

$$\frac{\sqrt{\langle (L_{\mathrm{NB,Re}}(\mathbf{q}) - \langle L_{\mathrm{NB,Re}}(\mathbf{q}) \rangle)^2 \rangle}}{\langle L_{\mathrm{NB,Re}}(\mathbf{q}) \rangle} = \sqrt{\frac{(2\pi)^{d/2}}{N V_{\mathrm{cell}} \sigma_x \sigma_y \sigma_z}}. \tag{3.2.6}$$

Using the method developed in appendix B.3, we can show that (3.2.6) is also valid for the more general case of multiple molecules in a unit cell coupled with multiple excited states per molecule. Note that (3.2.6) is valid under the condition of $1 \ll (\pi)^{d/2}/(V_{\mathrm{cell}} \sigma_x \sigma_y \sigma_z) \ll N$, where V_{cell} is the unit cell volume.

Equation (3.2.6) shows that the relative intensity fluctuation can be negligibly small in the case of a non-sharp resolution function since $N V_{\mathrm{cell}} \sigma_x \sigma_y \sigma_z \gg 1$. For example, at the Advanced Photon Source 14 ID-B, the angular divergences (HWHM) are about 11 μrads (horizontal) by 3 μrads (vertical), and the energy spread is about 1 eV (HWHM). At an x-ray wavelength of 1 Å, we have $\sigma_x \approx 6.9 \times 10^{-5}$ Å^{-1}, $\sigma_y \approx 1.9 \times 10^{-5}$ Å^{-1}, $\sigma_z \approx 2.3 \times 10^{-4}$ Å^{-1}

at $\sqrt{q_x^2+q_y^2} = 6$ Å^{-1}. Therefore, for a sample size of $(0.1\text{ mm})^3$, the relative non-Bragg scattering intensity fluctuation is $\sqrt{(2\pi)^{d/2}/(V\sigma_x\sigma_y\sigma_z)} \approx 0.0072$, which is much smaller than what is expected for the case of an infinitely sharp resolution. Detector pixel and illuminated sample sizes typically increase greatly the effective spectrometer resolution widths, making the non-Bragg scattering fluctuation even smaller.

The quench effect on the non-Bragg scattering differential cross section fluctuation due to spectrometer resolutions can be easily understood by examining (2.1.10). With a finite-width spectrometer resolution function, the fluctuation term in (2.1.10) $\sum_{k\neq k'}^{K} \exp[-i\mathbf{q}(\mathbf{R}_{n(k)} - \mathbf{R}_{n(k')})]$ is suppressed by the resolution and has contributions from only those pairs which are relatively close to each other, leading to a much smaller fluctuation.

3.3. Computer Monte Carlo simulations

To support our claims in the last two subsections for the scattering fluctuations, we carry out computer Monte Carlo simulations on a one-dimensional crystal. We use the random number generator provided by JAVA 1.4.2. In the crystal, each unit cell contains only one molecule. What we want to test are (2.1.15) and (3.1.3), which, respectively, describe the ensemble averages of the non-Bragg scattering lattice factor and its fluctuation in the case of an infinitely sharp spectrometer resolution, and (3.2.6), which describes the ensemble average of the non-Bragg scattering differential cross section relative fluctuation with a non-infinitely sharp resolution. The unit cell size is 100 Å. We set $q = 0.03$ Å^{-1} and $p = 0.2$ for the excitation probability. We consider an ensemble consisting of 100 such systems.

For an infinitely sharp spectrometer resolution, we show in figure 2(a) our computer simulation results. In the figure, we set the number of unit cells $N = 20\,000$. The figure clearly indicates large intensity fluctuation from system to system. The average value of the non-Bragg scattering lattice factors in the simulation is 3611, whereas our theoretical prediction is $N \times 0.2 \times (1-0.2) = 3200$. The relative fluctuation of the lattice factors $\sqrt{\langle(L_{\mathrm{NB}}(\mathbf{q}) - \langle L_{\mathrm{NB}}(\mathbf{q})\rangle)^2\rangle}/\langle L_{\mathrm{NB}}(\mathbf{q})\rangle$ is 1.07, while our theoretical prediction is 1. In figure 2(b), we plot $\sqrt{\langle(L_{\mathrm{NB}}(\mathbf{q}) - \langle L_{\mathrm{NB}}(\mathbf{q})\rangle)^2\rangle}/\langle L_{\mathrm{NB}}(\mathbf{q})\rangle$ versus N for N varying from 500 to 40 000. It can be easily seen that the relative fluctuation is about one, agreeing with our prediction.

To show the spectrometer resolution effect, we use a Gaussian shaped functional form for the resolution function with a half width of $\sigma_x = 0.0002$ Å^{-1}. In figure 3(a), we plot the computer simulation result with $N = 20\,000$. The figure shows that the large fluctuation presented in figure 2(a) is greatly reduced. The simulation gives an average non-Bragg scattering lattice factor of 3202 and a relative fluctuation of 0.083, agreeing very well with our theoretical predictions of 3200 and 0.080, respectively. In figure 3(b), we plot, in logarithmic scale for both axes, $\sqrt{\langle(L_{\mathrm{NB,Re}}(\mathbf{q}) - \langle L_{\mathrm{NB,Re}}(\mathbf{q})\rangle)^2\rangle}/\langle L_{\mathrm{NB,Re}}(\mathbf{q})\rangle$ versus N with N varying from 400 to 40 000. The figure clearly demonstrates that the relative non-Bragg scattering lattice factor behaves as $1/\sqrt{N}$, consistent with our theoretical prediction.

4. Electron density noise calculations

Since experimental data contain noise, it is important to understand what determines the noise for the electron density profile so that an experiment can be better designed. The signal is the computed $\Delta\rho(\mathbf{r})$ from experimental data, and the noise is the deviation of the computed $\Delta\rho(\mathbf{r})$ from its ensemble average. In this section, we shall investigate the noise in $\Delta\rho(\mathbf{r})$ for both cases of using Bragg and non-Bragg scattering.

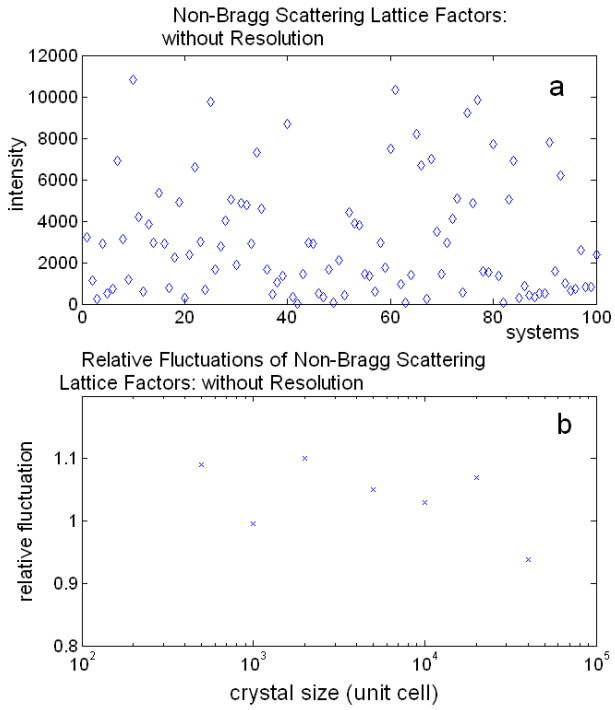

Figure 2. (a) Computer simulated substitution disorder created non-Bragg scattering lattice factors for 100 one-dimensional (1D) crystals with an infinitely sharp spectrometer resolution. There is only one excitation state per molecule. Each crystal contains exactly the same 20 000 unit cells and each unit cell has a size of 10^2 Å. $q = 0.03$ Å^{-1}. The excitation probability p is 0.2. The average scattering lattice factor is 3611 and the relative fluctuation $\sqrt{\langle(L_{\rm NB}(\mathbf{q}) - \langle L_{\rm NB}(\mathbf{q})\rangle)^2\rangle}/\langle L_{\rm NB}(\mathbf{q})\rangle$ is 1.07. (b) Plot of computer simulated $\sqrt{\langle(L_{\rm NB}(\mathbf{q}) - \langle L_{\rm NB}(\mathbf{q})\rangle)^2\rangle}/\langle L_{Nb}(\mathbf{q})\rangle$ for different total numbers of unit cells in the crystal with an infinitely sharp spectrometer resolution. The other parameters take the same values as those in (a).

Often, to use Bragg peaks to extract $\Delta\rho(\mathbf{r})$, the corresponding difference Fourier map is used [19]. In this scenario, the contribution to the noise in $\Delta\rho(\mathbf{r})$ can come from a few sources which include the quantum fluctuations of the scattering process and the excitation of the molecules. The scattered x-ray intensity fluctuates as a Poisson random variable, and the excitation number fluctuations have been considered in section 3. We want to know what the uncertainties are in determining $\Delta\rho(\mathbf{r})$ due to these fluctuations. We shall consider the simple case of only one molecule per unit cell with only one excited state per molecule. We ignore the non-Bragg scattering intensity fluctuation created by the excitation process since it is greatly suppressed by the spectrometer resolution. Factors which we ignore but do contribute to the $\Delta\rho(\mathbf{r})$ noise are the uncertainties in the incoming x-ray intensity and the intensity of the laser light used to excite the molecules, and the solvent (in biological samples) and air scattering fluctuations.

4.1. Difference Fourier map

In this subsection, we consider the Bragg scattering difference map created uncertainty in $\Delta\rho(\mathbf{r})$. We denote $F_{\rm B,G}$ and $F_{\rm B,E}$ as the Bragg scattering structure factors of the molecule in the ground and excited states, respectively. We assume that the magnitudes of $F_{\rm B,G}$ and $F_{\rm B,E}$

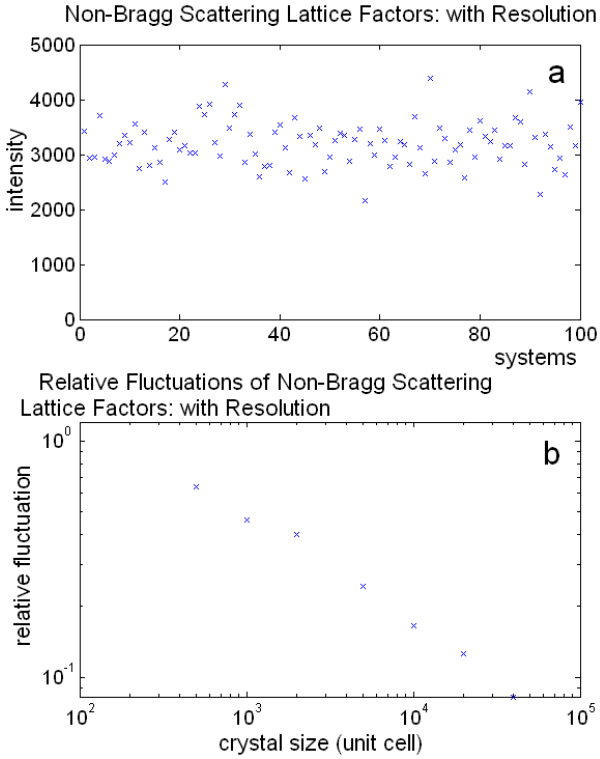

Figure 3. (a) Computer simulated substitution disorder created non-Bragg scattering lattice factors for 100 1D crystals with a Gaussian shaped spectrometer resolution function. The Gaussian half width is 0.0002 Å^{-1}. Each crystal contains exactly the same 20 000 unit cells and each unit cell has a size of 10^2 Å. $q = 0.03$ Å^{-1}. The probability of excitation p is 0.2. The average lattice factor has a value of 3202, and the relative fluctuation $\sqrt{\langle(L_{\mathrm{NB}}(\mathbf{q}) - \langle L_{\mathrm{NB}}(\mathbf{q})\rangle)^2\rangle}/\langle L_{\mathrm{NB}}(\mathbf{q})\rangle$ is 0.083. (b) Plot of computer simulated $\sqrt{\langle(L_{\mathrm{NB}}(\mathbf{q}) - \langle L_{\mathrm{NB}}(\mathbf{q})\rangle)^2\rangle}/\langle L_{Nb}(\mathbf{q})\rangle$ for different total numbers of unit cells in the crystal with a Gaussian shaped spectrometer resolution function of half width 0.0002 Å^{-1}. The other parameters are the same as those in (a).

are close. Under the strong assumption that the summation of $(F_{\mathrm{B,E}} - F_{\mathrm{B,G}})\exp(-2i\alpha_{\mathrm{G}})$ over all the Bragg peaks is zero, where α_{G} is the phase in $F_{\mathrm{B,G}}$, it is easy to show that $\Delta\rho(\mathbf{r})$ can be obtained by [1]

$$\Delta\rho(\mathbf{r}) = \frac{2}{pV_{\mathrm{cell}}}\sum_{\mathbf{Q}_j}(|F_{\mathrm{B,E}}| - |F_{B,G}|)\mathrm{e}^{-\mathrm{i}(\mathbf{Q}_j\mathbf{r}-\alpha_{\mathrm{G}})}. \tag{4.1.1}$$

Therefore, the noise squared in $\Delta\rho(\mathbf{r})$ due to the scattering process is

$$|\delta\Delta\rho(\mathbf{r})|^2 = \left(\frac{2}{pV_{\mathrm{cell}}}\right)^2 \sum_{\mathbf{Q}_j\mathbf{Q}_{j'}}(\delta|F_{\mathrm{B,E}}(\mathbf{Q}_j)| - \delta|F_{\mathrm{B,G}}(\mathbf{Q}_j)|) \times (\delta|F_{\mathrm{B,E}}(\mathbf{Q}_{j'})| - \delta|F_{\mathrm{B,G}}(\mathbf{Q}_{j'})|)\mathrm{e}^{-\mathrm{i}[\mathbf{Q}_j\mathbf{r}-\alpha_{\mathrm{G}}-\mathbf{Q}_{j'}\mathbf{r}+\alpha'_{\mathrm{G}}]}, \tag{4.1.2}$$

where $\delta|F_{\mathrm{B,E}}(\mathbf{Q}_j)|$ and $\delta|F_{\mathrm{B,G}}(\mathbf{Q}_j)|$ are the noises associated with the Bragg scattering structure factor magnitudes for the crystal in excited and ground states, respectively. To simplify (4.1.2), we note that $\delta|F_{\mathrm{B,E}}(\mathbf{Q}_j)|$ and $\delta|F_{\mathrm{B,G}}(\mathbf{Q}_j)|$ are uncorrelated, and $\delta|F_{\mathrm{B,E}}(\mathbf{Q}_j)|$ and $\delta|F_{\mathrm{B,E}}(\mathbf{Q}_{j'})|$ are also uncorrelated unless $j = j'$. We assume that the excitation probability

$p \ll 1$ and $\langle(\delta|F_{\mathrm{B,E}}(\mathbf{Q}_j)|)^2\rangle \approx \langle(\delta|F_{\mathrm{B,G}}(\mathbf{Q}_j)|)^2\rangle$. Taking the ensemble average, (4.1.2) becomes

$$\langle|\delta\Delta\rho(\mathbf{r})|^2\rangle = 2\left(\frac{2}{pV_{\mathrm{cell}}}\right)^2 \sum_{\mathbf{Q}_j} \langle(\delta|F_{\mathrm{B,G}}(\mathbf{Q}_j)|)^2\rangle. \tag{4.1.3}$$

To compute $\langle(\delta|F_{\mathrm{B,G}}(\mathbf{Q}_j)|)^2\rangle$, we note that the x-ray scattering intensity for the crystal in the ground state is

$$I_{\mathrm{B,G}} = r_0^2 N I_0 \Delta t |F_{\mathrm{B,G}}|^2. \tag{4.1.4}$$

Since $\langle(\delta I_{\mathrm{B,G}})^2\rangle = \langle I_{\mathrm{B,G}}\rangle$, (4.1.4) gives

$$\langle(\delta|F_{\mathrm{B,G}}|)^2\rangle = \frac{1}{4r_0^2\Delta t I_0 N}. \tag{4.1.5}$$

Substituting (4.1.5) in (4.1.3), we have

$$\langle|\delta\Delta\rho(\mathbf{r})|^2\rangle = \left(\frac{2}{pV_{\mathrm{cell}}}\right)^2 \sum_{\mathbf{Q}_j} \frac{1}{2r_0^2\Delta t I_0 N}. \tag{4.1.6}$$

Carrying out the summation in $\mathbf{Q}_j$ up to a surface of radius $Q_{\max}$, (4.1.6) becomes

$$\langle(\delta\Delta\rho_{\mathrm{B}}(\mathbf{r}))^2\rangle = \frac{2D(d)Q_{\max}^d}{Vp^2 r_0^2 \Delta t I_0 (2\pi)^d}. \tag{4.1.7}$$

In (4.1.7), $D(d)$ is a coefficient associated with the volume of a sphere in d-dimension. Equation (4.1.7) indicates that the noise in the determined electron density is of the same value everywhere in the unit cell, and the larger the $Q_{\max}$, the larger is the noise. This behaviour can be understood intuitively by realizing that each Bragg peak contributes to the uncertainty in $\Delta\rho(\mathbf{r})$, due to scattered x-ray photon count fluctuations, by the same amount. In addition, (4.1.7) shows that, for a very small excitation probability, the noise can be very high. This high noise level due to the dependence on p^{-2} is lessened greatly in the case of using the non-Bragg scattering to determine the electron density distribution. It is intuitively thought that the higher the $Q_{\max}$ is, the more accurate is $\Delta\rho(\mathbf{r})$. This intuition is purely based on the property of Fourier transforms, since the scattering data with a higher $Q_{\max}$ will reveal a sharper contour for $\Delta\rho(\mathbf{r})$. Nevertheless, without the noises in the scattering data, $\Delta\rho(\mathbf{r})$ does not change from one system to another at a particular $Q_{\max}$.

To get a sense of the magnitude for the electron density fluctuation, we take $d = 3$, $I_0 = 1$ Å^{-2} s^{-1}, $p = 0.1$, $V = 1 \times 10^{18}$ Å^3, $\Delta t = 10^{-6}$ s, and $Q_{\max} = 2\pi$ Å^{-1}, yielding $\sqrt{\langle(\delta\Delta\rho(\mathbf{r}))^2\rangle} = 1.0$ Å^{-3}, which is large and comparable to the carbon atom average electron density of 0.6 Å^{-3}.

4.2. Non-Bragg scattering

In this subsection, we determine the fluctuation in determining the electron density profile due to the fluctuation in the non-Bragg scattering intensity. Integrating the non-Bragg scattering intensity over a small solid angle $\Delta\Omega$ around a particular $\mathbf{q}$, the measured scattering intensity is

$$I_{\mathrm{NB}}(\mathbf{q}) = Nr_0^2 \Delta t I_0 p(1-p)\Delta\Omega \left|\int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}}\, \mathrm{d}\mathbf{r}\right|^2. \tag{4.2.1}$$

When the scattering intensity has a small deviation $\delta I_{\mathrm{NB}}(\mathbf{q})$, the electron density will consequently have a small deviation $\delta\Delta\rho(\mathbf{r})$. Ignoring the second-order term in $\delta\Delta\rho(\mathbf{r})$, we

have

$$\delta_l I_{\mathrm{NB}}(\mathbf{q}) = N I_0 r_0^2 \Delta t p (1-p) \Delta\Omega \times \left(\int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}}\, \mathrm{d}\mathbf{r} \int_{\mathrm{cell}} \delta_l \Delta\rho(\mathbf{r}') \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{r}'}\, \mathrm{d}\mathbf{r}' + \mathrm{c.c.} \right). \tag{4.2.2}$$

In (4.2.2), we add a subscript l to indicate the lth complete data set, and c.c. to indicate the complex conjugate of the first term in the bracket. Summing over $\mathbf{q}$, we have

$$\sum_{\mathbf{q}} \delta_l I_{\mathrm{NB}}(\mathbf{q}) = 2N I_0 r_0^2 \Delta t p (1-p) \langle \Delta\Omega \rangle \left(\frac{2\pi}{\Delta q} \right)^d \int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \delta_l \Delta\rho(\mathbf{r}) \mathrm{d}\mathbf{r}. \tag{4.2.3}$$

In (4.2.3), Δq is the step size in $\mathbf{q}$, and $\langle \Delta\Omega \rangle$ is the average solid angle each pixel makes with the sample. Taking the ensemble average of the squared on both sides of (4.2.3), and noting that $\langle (\delta_l I_{\mathrm{NB}}(\mathbf{q}))^2 \rangle = \langle I_{\mathrm{NB}}(\mathbf{q}) \rangle$ and $\langle \delta_l I_{\mathrm{NB}}(\mathbf{q}) \delta_l I_{\mathrm{NB}}(\mathbf{q}')_{\mathbf{q} \neq \mathbf{q}'} \rangle = 0$, after carrying out the summation of $\langle I_{\mathrm{NB}}(\mathbf{q}) \rangle$ over $\mathbf{q}$, we have

$$\int_{\mathrm{cell}} \Delta\rho(\mathbf{r})^2 \mathrm{d}\mathbf{r} = 4N I_0 r_0^2 \Delta t p (1-p) \langle \Delta\Omega \rangle \left(\frac{2\pi}{\Delta q} \right)^d \times \left\langle \left(\int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \delta_l \Delta\rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \right)^2 \right\rangle. \tag{4.2.4}$$

The ensemble average of the right-hand side of (4.2.4) can be expressed as

$$\left\langle \left(\int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \delta_l \Delta\rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \right)^2 \right\rangle = \lim_{N_l \to \infty} \frac{\sum_{l=1}^{N_l} \int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \delta_l \Delta\rho(\mathbf{r}) \mathrm{d}\mathbf{r} \int_{\mathrm{cell}} \Delta\rho(\mathbf{r}') \delta_l \Delta\rho(\mathbf{r}') \mathrm{d}\mathbf{r}'}{N_l}. \tag{4.2.5}$$

To compute the right-hand side of (4.2.5), we divide the unit cell into divisions of volume of $(\pi/q_{\max})^d$ and use j to denote a particular division. Since for $l \neq l'$, $\delta_l \Delta\rho(\mathbf{r})$ and $\delta_{l'} \Delta\rho(\mathbf{r})$ are uncorrelated, and for $j \neq j'$, $\delta_l \Delta\rho(\mathbf{r}_j) \delta_l \Delta\rho(\mathbf{r}_{j'})$ is assumed to randomly distribute around zero, namely

$$\lim_{N_l \to \infty} \frac{\sum_{l=1; j \neq j'}^{N_l} \delta_l \Delta\rho(\mathbf{r}_j) \delta_l \Delta\rho(\mathbf{r}_{j'})}{N_l} = 0, \tag{4.2.6}$$

equation (4.2.5) becomes

$$\left\langle \left(\int_{\mathrm{cell}} \Delta\rho(\mathbf{r}) \delta_l \Delta\rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \right)^2 \right\rangle = \lim_{N_l \to \infty} \frac{\left(\frac{\pi}{q_{\max}} \right)^{2d} \sum_{j=1}^{N_j} \Delta\rho(\mathbf{r}_j)^2 \sum_{l=1}^{N_l} [\delta_l \Delta\rho(\mathbf{r}_j)]^2}{N_l} = \left(\frac{\pi}{q_{\max}} \right)^{2d} \sum_{j=1}^{N_j} \Delta\rho(\mathbf{r}_j)^2 \langle [\delta \Delta\rho(\mathbf{r}_j)]^2 \rangle, \tag{4.2.7}$$

where $N_j = V_{\mathrm{cell}}/(\pi/q_{\max})^d$. Therefore (4.2.4) becomes

$$\frac{\sum_{j=1}^{N_j} [\Delta\rho(\mathbf{r}_j)]^2 \langle [\delta \Delta\rho(\mathbf{r}_j)]^2 \rangle}{\sum_{j=1}^{N_j} [\Delta\rho(\mathbf{r}_j)]^2} = \frac{(\Delta q q_{\max})^d}{4N I_0 r_0^2 \Delta t p (1-p) \langle \Delta\Omega \rangle (2\pi^2)^d}. \tag{4.2.8}$$

The left-hand side of (4.2.8) is the averaged noise level squared, weighted by $\Delta\rho(\mathbf{r}_j)^2$. Equation (4.2.8) shows that without solvent and air scattering, and with no electronic noises, the weighted ensemble average of the noise squared of the determined electron density, under

a fixed pixel solid angle $\langle\Delta\Omega\rangle$, is proportional to Δq^3 for a three-dimensional (3D) sample, meaning that sampling the scattering intensity distribution more finely should reduce drastically the noise in the determined electron density profile. In the extreme case for a 3D sample, we have $\Delta q = 2\pi\sqrt{\langle\Delta\Omega\rangle}/\lambda$, and (4.2.8) can be written as

$$\frac{\sum_{j=1}^{N_j}\Delta\rho_{\rm NB}(\mathbf{r}_j)^2\langle[\delta\Delta\rho_{\rm NB}(\mathbf{r}_j)]^2\rangle}{\sum_{j=1}^{N_j}\Delta\rho_{\rm NB}(\mathbf{r}_j)^2} = \frac{q_{\max}^3\sqrt{\langle\Delta\Omega\rangle}}{4NI_0r_0^2\Delta tp(1-p)(\pi\lambda)^3}. \tag{4.2.9}$$

Equation (4.2.9) shows that the smaller the pixel solid angle $\Delta\Omega$, the smaller the determined electron density noise level (achieved through much more computation time). For $N = 8\times10^{12}$, $q_{\max} = 2\pi$ Å^{-1}, $p = 0.1$, $\lambda = 1$ Å, $\langle\Delta\Omega\rangle = 10^{-4}$, $\Delta t = 10^{-6}$ s, and $I_0 = 1$ Å^{-2} s^{-1}, (4.2.9) gives that the weighted ensemble average of the electron density uncertainty is $\sqrt{\langle(\delta\Delta\rho_{\rm NB}(\mathbf{r}))^2\rangle_{\rm weighted}} = 0.74$ Å^{-3}, which is on the same level as the electron density fluctuation determined by using the difference Fourier map. Since the over-sampling ratio is proportional to $(\Delta q)^{-d}$, the noise in $\Delta\rho_{\rm NB}(\mathbf{r})$ is inversely proportional to the over-sampling ratio. Equation (4.2.8) suggests that if $\langle\Delta\Omega\rangle$ is infinitely small, the uncertainty in determining the electron density profile can be infinitely small as well, even if the detected non-Bragg scattering data are very noisy. Furthermore, (4.2.9) suggests to us the use of x-rays of longer wavelength to reduce the noise in the $\Delta\rho(\mathbf{r})$ determination.

In the presence of solvent (water) in the sample, the noise in the non-Bragg scattering contains an additional term associated with the solvent scattering noise, which cannot be subtracted. In this case, (4.2.8) is replaced by

$$\frac{\sum_{j=1}^{N_j}[\Delta\rho(\mathbf{r}_j)]^2\langle[\delta\Delta\rho(\mathbf{r}_j)]^2\rangle}{\sum_{j=1}^{N_j}[\Delta\rho(\mathbf{r}_j)]^2} = \frac{(\Delta qq_{\max})^d\left(1+\frac{N_{\rm W}q_{\max}^d}{N(2\pi^2)^dp(1-p)r_0^2}\times\frac{\int{\rm d}\mathbf{q}\ {\rm d}\sigma_{\rm W,molecule}(q)/{\rm d}\Omega}{\sum_{j=1}^{N_j}[\Delta\rho(\mathbf{r}_j)]^2}\right)}{4NI_0r_0^2\Delta tp(1-p)\langle\Delta\Omega\rangle(2\pi^2)^d}. \tag{4.2.10}$$

In (4.2.10), $N_{\rm W}$ is the number of water molecules in the sample, and ${\rm d}\sigma_{\rm W,molecule}(q)/{\rm d}\Omega$ is the water molecule x-ray scattering differential cross section in electron units. As demonstrated by (4.2.9), (4.2.10) suggests that, even with solvent scattering, the uncertainty in determining $\Delta\rho_{\rm NB}(\mathbf{r})$ is still inversely proportional to the over-sampling ratio, and for a 3D sample, the uncertainty is lower-bounded by how small $\langle\Delta\Omega\rangle$ is.

4.3. Noise level comparison

The crucial criterion in noise comparison should be that the time durations spent on collecting the entire sets of data for the Bragg and non-Bragg scattering be the same. However, to collect a complete set of data, the sample needs to be at many different orientations. Therefore, the total time spent in collecting a set of data may be different for Bragg and non-Bragg scattering.

Using the Laue scattering technique with incident x-rays of a broad wavelength spectrum from $\lambda_{\min}$ to $\lambda_{\max}$ (pink x-rays), ignoring the conic region around the rotation axis in reciprocal space [20], the number of times to rotate the sample to cover half of the possible q-space is (see figure 4) about $\pi/\angle$AOB (note that rotating the sample to move point A to point B does not mean that there is no q-space undetected except the cones between the two adjacent angular positions since the inner circle has a larger curvature). It is easy to obtain [21] that

$$\angle{\rm AOB} = \arcsin\frac{q_{\max}\lambda_{\max}}{4\pi} - \arcsin\frac{q_{\max}\lambda_{\min}}{4\pi}.$$

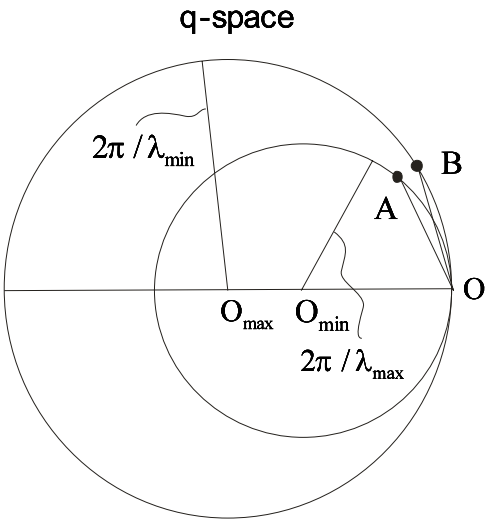

Figure 4. Diagram to show a cut in Ewald spheres corresponding to a Laue scattering with the largest x-ray wavelength of λ_{max} and the smallest wavelength of λ_{min}. OA = OB = q_{max}.

For a beamline offering a pink x-ray spectrum of 7 to 13 keV, $\lambda_{min} = 0.95$ Å and $\lambda_{max} = 1.78$ Å. Taking $q_{max} = 6$ Å^{-1}, we have $\angle$AOB $= 0.55$ rad. Therefore the number of sample angular positions in using the Laue Bragg scattering is about 6. In reality, the need to resolve overlapped Bragg peaks typically demands that $\angle$AOB $= 0.1$ rad, meaning that the number of sample angular positions to collect a complete set of data is about 31.

In measuring the non-Bragg scattering, at the start we may use only a few sample angular positions. Depending on the result of the data analysis, we may take data at finer angular positions. However, if the data analysis cannot be finished quickly to guide whether measurements at finer angular positions are needed, we may just take data at the positions between the Bragg peaks, meaning that the data are taken at a linear interval of $2\pi/V_{cell}^{1/3}$ in 3D q-space, corresponding to an angular step size of $2\pi/(q_{max}V_{cell}^{1/3})$. Therefore there are about $0.5q_{max}V_{cell}^{1/3}$ angular positions to cover half of the possible q-space. Taking $q_{max} = 6$ Å^{-1} and $V_{cell}^{1/3} = 50$ Å, we have $0.5q_{max}V_{cell}^{1/3} = 150$, which is about five times as many as the number used in the Laue Bragg scattering. Therefore, except at very low excitation probabilities ($p < 0.001$), the SNR in the electron density determination from using the non-Bragg scattering is lower than that from using the Laue Bragg scattering unless the over-sampling ratio is very high. Nevertheless, if $\Delta\rho(\mathbf{r})$ is confined to a much smaller space, we can afford to collect data with a much larger step size in reciprocal space.

Instead of using wide spectrum pink x-rays, if using monochromatic x-rays with Bragg scattering, the data collection relies on finite spectrometer resolution width and/or sample mosaicity. In this scenario, the sample rotation step size is often about 0.1° to 0.3°, and covering 180° means 1800 to 600 sample angular positions, leading to a significant longer time duration in data collection. Therefore, using time as a criterion for noise level comparison, this mode of data collection generates a lower SNR in electron density determination than that using the non-Bragg scattering.

5. Conclusions

In this paper, we first derived the ensemble averages of the non-Bragg scattering differential cross sections for a crystal having statistically uniformly distributed random substitution disorder. For the case of one molecule per unit cell, the non-Bragg scattering is just that for a

single pseudo-molecule of $\Delta\rho(\mathbf{r})$ multiplied by $Np(1-p)$. For the case of multiple molecules per unit cell, the non-Bragg scattering is the sum of the molecular scattering factors of $\Delta\rho(\mathbf{r})$ for all the molecules in the unit cell, resulting in an overlap of the scattering intensities. In dealing with the non-Bragg scattering overlap, we developed a Patterson verification method. We showed that the non-Bragg scattering can be easily applied to determine whether there is one excited state or multiple excited states, and the excitation probability. Using the non-Bragg scattering to determine the structure of $\Delta\rho(\mathbf{r})$ for crystals without solvent is possible using the traditional pump–probe approach, although the noise in determining $\Delta\rho(\mathbf{r})$ can be larger than that by utilizing the Laue Bragg scattering in the case of not-too-small excitation probabilities. In the presence of solvent, if there is only one excited state, we can shine the trigger laser continuously on to the sample and collect data with a time duration long enough to allow clear determination of $\Delta\rho(\mathbf{r})$. In the presence of solvent, if there are multiple excited states per molecule, we can shine the trigger laser continuously to obtain $|\sum_{n=1}^{N_e} p_n \Delta F_n(\mathbf{q})|^2$. Whether $|\sum_{n=1}^{N_e} p_n \Delta F_n(\mathbf{q})|^2$ can give us clear $\Delta\rho(\mathbf{r})$ at its excitation states needs further investigation. Caution is needed to make sure that the laser light will not heat up the sample much. In reaching the conclusions above, we assumed that the non-Bragg scattering associated with the thermal motions of the atoms and the unit cells and the non-Bragg scattering associated with crystal defects are unchanged before and after the trigger laser pulse.

We obtained the non-Bragg scattering intensity fluctuations associated with the randomness of the molecular excitations. We showed that under the condition of an infinitely sharp spectrometer resolution, the ensemble averages of the fluctuations are identical to the ensemble averages themselves, indicating very large fluctuations. However, in the case of a non-infinitely sharp spectrometer resolution, which is typical if the sample is not too small, the fluctuations from system to system are negligibly small, proportional to the square root of the total number of x-ray illuminated unit cells. The results of our computer Monte Carlo simulations on 1D crystal systems were shown to be consistent with our theory. For small samples, one needs to tune the spectrometer so that the resolution widths are large enough to suppress the fluctuations.

Finally, we considered the noise levels in determining $\Delta\rho(\mathbf{r})$ due to the Poisson noises associated with both the Bragg and the non-Bragg scattering. We showed that the noises in determining $\Delta\rho(\mathbf{r})$ through both the Bragg and the non-Bragg scattering can be significant. The noise in determining $\Delta\rho(\mathbf{r})$ associated with the non-Bragg scattering can be more than that associated with the Laue Bragg scattering but smaller than that through monochromatic Bragg scattering. Nevertheless, we realized the advantage of using the non-Bragg scattering to determine $\Delta\rho(\mathbf{r})$ due to the lack of systematic errors. Specifically, we found that the noise levels due to the scattering fluctuations through both the Bragg and the non-Bragg scattering are proportional to $q_{\max}^{d/2}$, and for the case of using non-Bragg scattering, the noise level in $\Delta\rho(\mathbf{r})$ is inversely proportional to the over-sampling ratio. Numerical studies are needed to provide support for these predications.

Acknowledgments

I thank Keith Moffat for proposing to me the study of non-Bragg scattering due to substitution disorder in molecular crystals. I also thank Spencer Anderson, Sudarshan Rajagopal, Vukica Srajer, Jianwei Miao, John C H Spence, Abraham Szoke for fruitful communications, Frantisek Pavelcik for sharing his insight on Patterson functions, and Tim Graber and Yusheng Chen for information regarding time-resolved crystallographic experiments using monochromatic x-rays. This work was supported by NIH grant RR07707 to Keith Moffat.

Appendix A. Derivation of the non-Bragg scattering intensity ensemble averages for more than one excited state with multiple molecules per unit cell

A.1. Two excited states with one molecule per unit cell

In this case, the non-Bragg scattering structure factor is

$$F_{\mathrm{NB}}(\mathbf{q}) = \int_{\mathrm{sample}} \left[\sum_{k_1=1}^{K_1} \Delta\rho_1(\mathbf{r} - \mathbf{R}_{n_1(k_1)}) + \sum_{k_2=1}^{K_2} \Delta\rho_2(\mathbf{r} - \mathbf{R}_{n_2(k_2)}) \right] \mathrm{d}\mathbf{r} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{r}}, \tag{A.1.1}$$

where $n_1(k_1)$ and $n_2(k_2)$ are the indices of the unit cells which are in the first and the second excited states, respectively, and K_1 and K_2 are the numbers of molecules in the first and the second excited states, respectively. Defining $\Delta F_1(\mathbf{q})$ and $\Delta F_2(\mathbf{q})$ as the Fourier transforms of $\Delta\rho_1(\mathbf{r})$ and $\Delta\rho_2(\mathbf{r})$, respectively, the non-Bragg scattering differential cross section is just (ignoring x-ray polarization and absorption)

$$\frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} = r_0^2 |\Delta F_1(\mathbf{q})|^2 \left| \sum_{k_1=1}^{K_1} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n_1(k_1)}} \right|^2 + r_0^2 |\Delta F_2(\mathbf{q})|^2 \left| \sum_{j_2=1}^{K_2} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n_2(k_2)}} \right|^2 + (\text{cross term}), \tag{A.1.2}$$

with the (cross term) defined as

$$\begin{aligned}(\text{cross term}) = {} & r_0^2 \Delta F_1(\mathbf{q}) \Delta F_2^*(\mathbf{q}) \sum_{k_1=1}^{K_1} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n_1(k_1)}} \sum_{k_2=1}^{K_2} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}_{n_2(k_2)}} \\ & + r_0^2 \Delta F_1^*(\mathbf{q}) \Delta F_2(\mathbf{q}) \sum_{k_1=1}^{K_1} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}_{n_1(k_1)}} \sum_{k_2=1}^{K_2} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n_2(k_2)}}. \end{aligned} \tag{A.1.3}$$

Using (2.1.13), we have

$$\langle (\text{crossterm}) \rangle = \frac{-r_0^2 \langle K_1 K_2 \rangle}{N-1} [\Delta F_1(\mathbf{q}) \Delta F_2^*(\mathbf{q}) + \Delta F_1^*(\mathbf{q}) \Delta F_2(\mathbf{q})]. \tag{A.1.4}$$

Therefore, the non-Bragg scattering differential cross section is

$$\begin{aligned}\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega}(\mathbf{q}) \right\rangle = {} & r_0^2 N p_1 (1-p_1) |\Delta F_1(\mathbf{q})|^2 + \Gamma N p_2 (1-p_2) |\Delta F_2(\mathbf{q})|^2 \\ & - \frac{r_0^2 \langle K_1 K_2 \rangle}{N-1} [\Delta F_1(\mathbf{q}) \Delta F_2^*(\mathbf{q}) + \Delta F_1^*(\mathbf{q}) \Delta F_2(\mathbf{q})]. \end{aligned} \tag{A.1.5}$$

Since $\langle K_1 K_2 \rangle = N(N-1) p_1 p_2$, (A.1.5) becomes

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 N p_1 |\Delta F_1(\mathbf{q})|^2 + r_0^2 N p_2 |\Delta F_2(\mathbf{q})|^2 - \frac{r_0^2}{N} |N p_1 \Delta F_1(\mathbf{q}) + N p_2 \Delta F_2(\mathbf{q})|^2. \tag{A.1.6}$$

Equation (A.1.6) can be generalized to account for the case of any number of excited states per molecule. The general form is, at off-Bragg peak positions,

$$\left\langle \frac{\mathrm{d}\sigma_{\mathrm{NB}}}{\mathrm{d}\Omega} \right\rangle = r_0^2 \sum_{n=1}^{N_e} N p_n |\Delta F_n(\mathbf{q})|^2 - \frac{r_0^2}{N} \left| \sum_{n=1}^{N_e} N p_n \Delta F_n(\mathbf{q}) \right|^2. \tag{A.1.7}$$

A.2. Multiple excited states, with each unit cell containing multiple molecules

For the case of N_e excited states per molecule and N_m molecules per unit cell, the non-Bragg scattering structure factor is

$$F_{NB}(\mathbf{q}) = \sum_{m=1}^{N_m} F_{NB,m}(\mathbf{q}), \tag{A.2.1}$$

where

$$F_{NB,m}(\mathbf{q}) = \sum_{n=1}^{N_e} \sum_{k_{m,n}=1}^{K_{m,n}} e^{-i\mathbf{q}(\mathbf{R}_{n_{m,n}(k_{m,n})}+\mathbf{r}_m)} \Delta F_{m,n}(\mathbf{q}). \tag{A.2.2}$$

In (A.2.2), $n_{m,n}(k_{m,n})$ is the index of a unit cell which has an mth molecule in its nth excited state, $K_{m,n}$ is the total number of mth molecules in nth excited state in the crystal, and $\mathbf{r}_m$ is the location of the mth molecule relative to its unit cell centre. Therefore the total non-Bragg scattering differential cross section is

$$\left\langle \frac{d\sigma_{NB}}{d\Omega} \right\rangle = r_0^2 \sum_{m=1}^{N_m} \left\langle |F_{NB,m}(\mathbf{q})|^2 \right\rangle + r_0^2 \sum_{m=1}^{N_m} \sum_{m'=1;m'\neq m}^{N_m} \langle F_{NB,m}(\mathbf{q}) F^*_{NB,m'}(\mathbf{q}) \rangle. \tag{A.2.3}$$

Since whether an mth molecule is excited is independent of whether an m'th molecule is excited if $m \neq m'$, at off-Bragg peak positions, the ensemble average of $\exp[-i\mathbf{q}(\mathbf{R}_{n_{m,n}(k_{m,n})} - \mathbf{R}_{n_{m',n'}(k_{m',n'})})]$ is zero when $m \neq m'$. Therefore

$$\langle F_{NB,m}(\mathbf{q}) F^*_{NB,m'}(\mathbf{q}) \rangle_{m\neq m'} = 0, \tag{A.2.4}$$

which gives

$$\langle I_{NB}(\mathbf{q}) \rangle = r_0^2 \sum_{m=1}^{N_m} \left\langle |F_{NB,m}(\mathbf{q})|^2 \right\rangle. \tag{A.2.5}$$

Hayakawa and Cohen studied the diffuse scattering from local atomic disorder for a crystal with multiple sublattices [22]. The sublattice in their study corresponds to a lattice formed by the molecules in a particular position in the unit cells in this paper. Assuming the atoms are always at their equilibrium positions and a complete random distribution for different kinds of atoms among the sublattices, we show that their result agrees with (A.2.5). Note that, in their formula, probabilities over a single sample were used, and for a particular sample the use of the probabilities breaks down when the number of unit cells multiplying the probabilities are small. In this extreme case, the concept of an ensemble has to be used. If there is only one excited state per molecule, (A.2.5) was also reached by Szoke [13].

Appendix B. Derivation of the non-Bragg scattering intensity fluctuations

B.1. One excited state, with one molecule per unit cell: without spectrometer resolution considerations

Using the definition (2.1.8) for the non-Bragg scattering lattice factor $L_{NB}(\mathbf{q})$, we have

$$\langle L^2_{NB}(\mathbf{q}) \rangle_{suben} = \left\langle \left(\sum_{k_1=1}^{K} e^{-i\mathbf{q}\mathbf{R}_{n(k_1)}} \sum_{k'_1=1}^{K} e^{i\mathbf{q}\mathbf{R}_{n(k'_1)}} \right) \left(\sum_{k_2=1}^{K} e^{-i\mathbf{q}\mathbf{R}_{n(k_2)}} \sum_{k'_2=1}^{K} e^{i\mathbf{q}\mathbf{R}_{n(k'_2)}} \right) \right\rangle_{suben}. \tag{B.1.1}$$

In (B.1.1), the subscript suben indicates a subensemble in which every system has K molecules excited. Since each particular system in the subensemble has an occurrence probability of

$1/C_N^K$, the right-hand side of (B.1.1) can be written as

$$\left\langle \sum_{\substack{k_1=1;k_1'=1\\k_2=1;k_2'=1}}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_1)}-\mathbf{R}_{n(k_1')}+\mathbf{R}_{n(k_2)}-\mathbf{R}_{n(k_2')})} \right\rangle_{\text{suben}}$$

$$= \frac{1}{C_N^K} \sum_{S_K} \sum_{\substack{k_1=1;k_1'=1\\k_2=1;k_2'=1}}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_1;S_K)}-\mathbf{R}_{n(k_1';S_K)}+\mathbf{R}_{n(k_2;S_K)}-\mathbf{R}_{n(k_2';S_K)})}, \tag{B.1.2}$$

where S_K indicates a particular system in the subensemble.

We exchange the order of summations in S_K and (k_1, k_1', k_2, k_2') on the right-hand side of (B.1.2) and note that

$$\sum_{k_1=1,k_1'=1,k_2=1,k_2'=1}^{K} \sum_{\substack{S_K\\(k_1,k_1',k_2,k_2')\in S_K}} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_1;S_K)}-\mathbf{R}_{n(k_1';S_K)}+\mathbf{R}_{n(k_2;S_K)}-\mathbf{R}_{n(k_2';S_K)})}$$

$$= \sum_{n_1=1;n_1'=1;n_2=1;n_2'=1}^{N} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n_1}-\mathbf{R}_{n_1'}+\mathbf{R}_{n_2}-\mathbf{R}_{n_2'})} \sum_{\substack{S_K\\(n_1;n_1';n_2;n_2')\in S_K}} . \tag{B.1.3}$$

Therefore the key to computing the fluctuation is to compute the summation of the number of systems under the constraint that the excited molecules in the unit cells of indices n_1, n_1', n_2, and n_2' belong to the same system.

Computing the system summation $\sum_{S_K}$ in (B.1.3) is straightforward. For example,

$$\sum_{\substack{S_K\\(n_1=n_1'=n_2=n_2')\in S_K}} = C_{N-1}^{K-1} = \frac{(N-1)(N-2)\cdots(N-K+1)}{(K-1)!}, \tag{B.1.4}$$

and

$$\sum_{\substack{S_K\\(n_1;n_1'\neq n_1;n_2\neq n_1,n_2\neq n_1';n_2'\neq n_1,n_2'\neq n_1',n_2'\neq n_2)\in S_K}} = C_{N-4}^{K-4}. \tag{B.1.5}$$

Carrying out the computation on the right-hand side of (B.1.3), we have, at off-Bragg peak positions,

$$\sum_{k_1=1,k_1'=1,k_2=1,k_2'=1}^{K} \sum_{\substack{S_K\\(k_1,k_1',k_2,k_2')\in S_K}} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_1;S_K)}-\mathbf{R}_{n(k_1';S_K)}+\mathbf{R}_{n(k_2;S_K)}-\mathbf{R}_{n(k_2';S_K)})}$$

$$= 2K^2\left(1-\frac{K}{N}\right)^2 C_N^K. \tag{B.1.6}$$

Therefore

$$\langle L_{\mathrm{NB}}^2(\mathbf{q})\rangle_{\text{suben}} = 2K^2\left(1-\frac{K}{N}\right)^2. \tag{B.1.7}$$

B.2. More than one excited state, with one molecule per unit cell: without spectrometer resolution considerations

To obtain $\langle|F_{\mathrm{NB}}|^4\rangle$, the key is to obtain $\langle|F_{\mathrm{NB}}|^4\rangle_{\text{suben}}$, where suben indicates a subensemble in which the crystals have $K_1, \ldots, K_{N_e}$ number of molecules in excited states $1, \ldots, N_e$,

respectively. We know

$$\left\langle |F_{\mathrm{NB}}|^4 \right\rangle_{\mathrm{suben}} = \left\langle \sum_{n_1=1}^{N_e} \Delta F_{n_1} \sum_{k_1=1}^{K_{n_1}} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_1;n_1;E)}} \sum_{n_1'=1}^{N_e} \Delta F_{n_1'}^* \sum_{k_1'=1}^{K_{n_1'}} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_1';n_1';E)}} \right.$$
$$\left. \times \sum_{n_2=1}^{N_e} \Delta F_{n_2}^* \sum_{k_2=1}^{K_{n_2}} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_2;n_2;E)}} \sum_{n_2'=1}^{N_e} \Delta F_{n_2'}^* \sum_{k_2'=1}^{K_{n_2'}} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_2';n_2';E)}} \right\rangle_{\mathrm{suben}} . \tag{B.2.1}$$

In (B.2.1), K_{n_1} is the number of molecules in the n_1th excited state and $n(k_1; n_1; E)$ is the unit cell index which has the molecule in its n_1th excited state. ΔF_{n_1} is the Fourier transform of the electron density difference between a molecule in the n_1th excited state and the ground state. Denoting the total summation in (B.2.1) as $S_{\mathrm{E}}(n_1, n_1'; n_2, n_2')$, $S_{\mathrm{E}}(n_1, n_1'; n_2, n_2')$ can be separated into 15 terms as

$$\begin{aligned} S_{\mathrm{E}}(n_1, n_1'; n_2, n_2') &= S_{\mathrm{E}}(n_1, n_1' = n_1; n_2 = n_1, n_2' = n_1) \\ &+ S_{\mathrm{E}}(n_1, n_1' = n_1; n_2 = n_1, n_2' \neq n_1) + S_{\mathrm{E}}(n_1, n_1' = n_1; n_2 \neq n_1, n_2' = n_1) \\ &+ S_{\mathrm{E}}(n_1, n_1' = n_1; n_2 \neq n_1, n_2' = n_2) \\ &+ S_{\mathrm{E}}(n_1, n_1' = n_1; n_2 \neq n_1, n_2' \neq n_1, n_2' \neq n_2) \\ &\ldots \\ &+ S_{\mathrm{E}}(n_1, n_1' \neq n_1; n_2 \neq n_1, n_2 \neq n_1', n_2' \neq n_1, n_2' \neq n_1', n_2' \neq n_2). \end{aligned} \tag{B.2.2}$$

Each term in (B.2.2) is computed. For example,

$$S_{\mathrm{E}}(n_1, n_1' = n_1; n_2 \neq n_1, n_2' \neq n_1, n_2' \neq n_2) = \sum_{n_1=1}^{N_e} \sum_{n_2=1, n_2 \neq n_1}^{N_e} \sum_{\substack{n_2'=1, n_2' \neq n_1 \\ n_2' \neq n_2}}^{N_e} |\Delta F_{n_1}|^2 \Delta F_{n_2} \Delta F_{n_2'}^*$$
$$\times \left\langle \sum_{k_1=1}^{K_{n_1}} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_1;n_1;E)}} \sum_{k_1'=1}^{K_{n_1}} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_1';n_1;E)}} \sum_{k_2=1}^{K_{n_2}} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_2;n_2;E)}} \sum_{k_2'=1}^{K_{n_2'}} \mathrm{e}^{\mathrm{i}\mathbf{q}\mathbf{R}_{n(k_2';n_2';E)}} \right\rangle_{\mathrm{suben}} . \tag{B.2.3}$$

As we did in appendix B.1, the subensemble average in the brackets in (B.2.3) is what is denoted as an ensemble series summation defined by

$$\mathrm{ENSS}(l_1, l_1'; l_2, l_2'; (l_1, l_1') \in n_1, l_2 \in n_2, l_2' \in n_2'; n_2 \neq n_1, n_2' \neq n_1, n_2' \neq n_2)$$
$$= \sum_{l_1=1}^{N} \sum_{l_1'=1}^{N} \sum_{\substack{l_2=1 \\ l_2 \neq l_1 \\ l_2 \neq l_1'}}^{N} \sum_{\substack{l_2'=1 \\ l_2' \neq l_1 \\ l_2' \neq l_1' \\ l_2' \neq l_2}}^{N} \mathrm{e}^{-\mathrm{i}\mathbf{q}[\mathbf{R}(l_1) - \mathbf{R}(l_1') + \mathbf{R}(l_2) - \mathbf{R}(l_2')]} \sum_{\substack{S\{K_n; n=1,\cdots,N_e\} \\ (l_1, l_1') \in n_1 \\ l_2 \in n_2; l_2' \in n_2' \\ n_2 \neq n_1; n_2' \neq n_1, n_2' \neq n_2}} , \tag{B.2.4}$$

divided by $N_{\mathrm{sys}} = N!/\Pi_{n=1}^{N_e} K_n!$, which is the total number of systems in the subensemble. In (B.2.4), $S\{K_n; n = 1, \ldots, N_e\}$ means a system which has exactly K_n molecules in the nth excited state with $n = 1, \ldots, N_e$. $(l_1, l_1') \in n_1$ means that the molecules in the unit cells of indices l_1 and l_1' are in the n_1th excited state. If the molecules in the unit cells of indices l_1 and l_1' are at the same excited state n_1, l_1 and l_1' can be equal or unequal (if l_1 and l_1' belong to different excited states, then l_1 and l_1' cannot be equal since a molecule can only be at one excited state at a particular time). Therefore the ensemble series summation in (B.2.4) can be

Table B.1. Values of the non-Bragg parts of the series summations which are defined through an example in (B.2.10). N is the total number of molecules in the crystal.

Conditions	$\mathrm{SS}(l_1, 1; l_1', -1; l_2, 1; l_2', -1)$
$l_1' = l_1; l_2 = l_1; l_2' = l_1$	N
$l_1' = l_1; l_2 = l_1; l_2' \neq l_1$	$-N$
$l_1' = l_1; l_2 \neq l_1; l_2' = l_1$	$-N$
$l_1' = l_1; l_2 \neq l_1; l_2' = l_2$	$N(N-1)$
$l_1' = l_1; l_2 \neq l_1; l_2' \neq l_1, l_2' \neq l_2$	$-N(N-2)$
$l_1' \neq l_1; l_2 = l_1; l_2' = l_1$	$-N$
$l_1' \neq l_1; l_2 = l_1; l_2' = l_1'$	$-N$
$l_1' \neq l_1; l_2 = l_1; l_2' \neq l_1, l_2' \neq l_1'$	$2N$
$l_1' \neq l_1; l_2 = l_1'; l_2' = l_1$	$N(N-1)$
$l_1' \neq l_1; l_2 = l_1'; l_2' = l_1'$	$-N$
$l_1' \neq l_1; l_2 = l_1'; l_2' \neq l_1, l_2' \neq l_1'$	$-N(N-2)$
$l_1' \neq l_1; l_2 \neq l_1, l_2 \neq l_1'; l_2' = l_1$	$-N(N-2)$
$l_1' \neq l_1; l_2 \neq l_1, l_2 \neq l_1'; l_2' = l_1'$	$2N$
$l_1' \neq l_1; l_2 \neq l_1, l_2 \neq l_1'; l_2' = l_2$	$-N(N-2)$
$l_1' \neq l_1; l_2 \neq l_1, l_2 \neq l_1'; l_2' \neq l_1, l_2' \neq l_1', l_2' \neq l_2$	$2N(N-3)$

split into two ensemble series summations corresponding to the cases of $l_1 = l_1'$ and $l_1 \neq l_1'$:

$$\begin{aligned}\mathrm{ENSS}(l_1, l_1'; l_2, l_2'; (l_1,\ l_1') \in n_1, l_2 \in n_2, l_2' \in n_2'; n_2 \neq n_1, n_2' \neq n_1, n_2' \neq n_2)\\ = \mathrm{ENSS}(l_1, l_1' = l_1; l_2, l_2'; C) + \mathrm{ENSS}(l_1, l_1' \neq l_1; l_2, l_2'; C).\end{aligned} \tag{B.2.5}$$

For simplicity, we denoted C on the right-hand side of (B.2.5) as the condition specified on the left-hand side. It is easy to see that

$$\begin{aligned}\mathrm{ENSS}(l_1, l_1' = l_1; l_2, l_2'; C) = \frac{(N-3)!}{(K(n_1)-1)!(K(n_2)-1)!(K(n_2')-1)!}\\ \times \mathrm{SS}(l_1, 1; l_1' = l_1, -1; l_2 \neq l_1, 1; l_2' \neq l_1, l_2' \neq l_2, -1),\end{aligned} \tag{B.2.6}$$

where the series summation $\mathrm{SS}(\cdots)$ is defined as

$$\begin{aligned}\mathrm{SS}(l_1, 1; l_1' = l_1, -1; l_2 \neq l_1, 1; l_2' \neq l_1, l_2' \neq l_2, -1)\\ = \sum_{l_1=1}^{N} \sum_{l_1'=l_1}^{N} \sum_{\substack{l_2=1 \\ l_2 \neq l_1}}^{N} \sum_{\substack{l_2'=1 \\ l_2' \neq l_1 \\ l_2' \neq l_2}}^{N} \mathrm{e}^{\mathrm{i}\mathbf{q}[\mathbf{R}(l_1)-\mathbf{R}(l_1')+\mathbf{R}(l_2)-\mathbf{R}(l_2')]}.\end{aligned} \tag{B.2.7}$$

On the left-hand side of (B.2.7), in the brackets, positive one means $\exp(\mathrm{i}\mathbf{qR})$ and negative one means $\exp(-\mathrm{i}\mathbf{qR})$. Similarly

$$\begin{aligned}\mathrm{ENSS}(l_1, l_1' \neq l_1; l_2, l_2'; C) = \frac{(N-4)!}{(K_{n_1}-2)!(K_{n_2}-1)!(K_{n_2'}-1)!}\\ \times \mathrm{SS}(l_1, 1; l_1' \neq l_1, -1; l_2 \neq l_1, l_2 \neq l_1', 1; l_2' \neq l_1, l_2' \neq l_1', l_2' \neq l_2, -1).\end{aligned} \tag{B.2.8}$$

It is straightforward to compute the non-Bragg part of the series summations. We list the results in table B.1.

After computing each term in (B.2.2), we obtain that

$$\begin{aligned}
\langle |F_{\mathrm{NB}}|^4\rangle_{\mathrm{suben}} &= 2\sum_{n=1}^{N_e} K_n^2\left(1-\frac{K_n}{N}\right)^2(\Delta M_n)^2 \\
&+ \sum_{n=1}^{N_e}\sum_{n'=1,n'\neq n}^{N_e}\frac{[(N^2-N)-N(K_n+K_{n'})+2K_nK_{n'}]K_nK_{n'}\Delta M_n\Delta M_{n'}}{(N-1)(N-2)} \\
&+ \frac{2}{(N-1)(N-2)}\sum_{n_1=1}^{N_e}\sum_{\substack{n_2=1\\ n_2\neq n_1}}^{N_e}\Delta M_{n_1}K_{n_1}K_{n_2}\Bigg[(2K_{n_1}^2-2NK_{n_1}+N) \\
&\times (\Delta F_{n_1}\Delta F_{n_2}^* + \Delta F_{n_1}^*\Delta F_{n_2}) + \sum_{\substack{n_2'=1\\ n_2'\neq n_1\\ n_2'\neq n_1'}}^{N_e} K_{n_2'}(2K_{n_1}-N)\Delta F_{n_2}\Delta F_{n_2'}^*\Bigg] \\
&+ \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2=n_1;n_2'=n_1')\frac{K_{n_1}K_{n_1'}(2K_{n_1}K_{n_1'}-N)}{(N-1)(N-2)}\Delta F_{n_1}^2(\Delta F_{n_1'}^*)^2 \\
&+ \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2=n_1;n_2'\neq n_1,n_2'\neq n_1')\frac{2K_{n_1}^2K_{n_1'}K_{n_2'}}{(N-1)(N-2)} \\
&\times \Delta F_{n_1}^2\Delta F_{n_1'}^*\Delta F_{n_2'}^* + \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2=n_1';n_2'=n_1) \\
&\times \frac{K_{n_1}K_{n_1'}(N^2-N-NK_{n_1}-NK_{n_1'}+2K_{n_1}K_{n_1'})}{(N-1)(N-2)}\Delta M_{n_1}\Delta M_{n_1'} \\
&+ \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2=n_1';n_2'\neq n_1;n_2'\neq n_1') \\
&\times \frac{K_{n_1}K_{n_1'}K_{n_2'}(2K_{n_1'}-N)}{(N-1)(N-2)}\Delta F_{n_1}\Delta M_{n_1'}\Delta F_{n_2'}^* \\
&+ \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2\neq n_1,n_2\neq n_1';n_2'=n_1) \\
&\times \frac{K_{n_1}K_{n_1'}K_{n_2}(2K_{n_1}-N)}{(N-1)(N-2)}\Delta M_{n_1}\Delta F_{n_1'}^*\Delta F_{n_2} \\
&+ \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2\neq n_1,n_2\neq n_1';n_2'=n_1') \\
&\times \frac{2K_{n_1}K_{n_1'}^2K_{n_2}}{(N-1)(N-2)}\Delta F_{n_1}(\Delta F_{n_1'}^*)^2\Delta F_{n_2} \\
&+ \Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2\neq n_1,n_2\neq n_1';n_2'\neq n_1,n_2'\neq n_1',n_2'\neq n_2) \\
&\times \frac{2K_{n_1}K_{n_1'}K_{n_2}K_{n_2'}}{(N-1)(N-2)}\Delta F_{n_1}\Delta F_{n_1'}^*\Delta F_{n_2}\Delta F_{n_2'}^*. \qquad \text{(B.2.9)}
\end{aligned}$$

In (B.2.9), $\Delta M_{n_1} = |\Delta F_{n_1}|^2$, and $\Sigma_{\mathrm{E}}(\cdots)$ denotes the summation over the indices of excited states. For example,

$$\Sigma_{\mathrm{E}}(n_1;n_1'\neq n_1;n_2=n_1;n_2'=n_1')^2 = \sum_{n_1=1}^{N_e}\sum_{\substack{n_1'=1\\ n_1'\neq n_1}}^{N_e}\sum_{\substack{n_2=1\\ n_2=n_1}}^{N_e}\sum_{\substack{n_2'=1\\ n_2'=n_1'}}^{N_e}. \qquad \text{(B.2.10)}$$

Table B.2. Values of moments of $\{K_n\}$ which follow multinomial distributions. $\{K_n\}$ have the corresponding probabilities of $\{p_n\}$.

Moments	Conditions	Results
$\langle K_{n_1} K_{n_1'} K_{n_2} K_{n_2'}\rangle$	No n_1, n_1', n_2, n_2' are the same	$N^2(N-2)(N-3)\,p_{n_1} p_{n_1'} p_{n_2} p_{n_2'}$
$\langle K_{n_1}^2 K_{n_2}^2\rangle$	$n_1 \neq n_2$	$Np_{n_1} p_{n_2}[p_{n_1} p_{n_2}(N-1)(N-2)(N-3)$ $+(p_{n_1}+p_{n_2})(N-1)(N-2)+$ $N-1]$
$\langle K_{n_1}^2 K_{n_2} K_{n_2'}\rangle$	$n_1 \neq n_2, n_1 \neq n_2'$ $n_2 \neq n_2'$	$N(N-1)(N-2)\,p_{n_1} p_{n_2} p_{n_2'}[1+p_{n_1}(N-3)]$
$\langle K_{n_1}^3 K_{n_2}\rangle$	$n_1 \neq n_2$	$N(N-1)(N-2)\,p_{n_1}^2 p_{n_2}(Np_{n_1}+3-3p_{n_1})$
$\langle K_{n_1}^2 K_{n_2}\rangle$	$n_1 \neq n_2$	$N(N-1)\,p_{n_1} p_{n_2}(Np_{n_1}+1-2p_{n_1})$
$\langle K_{n_1} K_{n_2} K_{n_2'}\rangle$	No n_1, n_2, n_2' are the same	$N(N-1)(N-2)\,p_{n_1} p_{n_2} p_{n_2'}$
$\langle K_{n_1} K_{n_2}\rangle$	$n_1 \neq n_2$	$N(N-1)\,p_{n_1} p_{n_2}$
$\langle K_n^2\rangle$		$Np_n(Np_n+1-p_n)$

We use the results of multinomial distributions for $\{K_n\}$ listed in table B.2 and finally obtain

$$\langle |F_{\mathrm{NB}}|^4\rangle = 2(\langle |F_{\mathrm{NB}}|^2\rangle)^2. \tag{B.2.11}$$

In table B.2, p_{n_1} is the probability of a molecule being in its n_1th excited state.

B.3. More than one excited state, coupled with more than one molecule per unit cell: without spectrometer resolution considerations

The ensemble average of the fourth power of the magnitude of the non-Bragg x-ray scattering structure factor for the crystal is

$$\langle |F_{\mathrm{NB}}|^4\rangle = \sum_{\substack{n_1;n_1'\\ n_2;n_2'}}^{N_{\mathrm{e}}} \sum_{\substack{m_1;m_1'\\ m_2;m_2'}}^{N_{\mathrm{m}}} \langle F_{\mathrm{NB},n_1,m_1} F^*_{\mathrm{NB},n_1',m_1'} F_{\mathrm{NB},n_2,m_2} F^*_{\mathrm{NB},n_2',m_2'}\rangle, \tag{B.3.1}$$

where

$$F_{\mathrm{NB},n,m} = \Delta F_{n,m} \sum_{k_{n,m}=1}^{K_{n,m}} \mathrm{e}^{-\mathrm{i}\mathbf{q}[\mathbf{R}(k_{n,m};n,m)+\mathbf{r}_m]}. \tag{B.3.2}$$

It is easy to see that only under one of the following conditions is the ensemble average on the right-hand side of (B.3.1) not zero: (i) $m_1 = m_1' = m_2 = m_2'$; (ii) $m_1 = m_1'$, $m_2 = m_2'$, $m_1 \neq m_2$; (iii) $m_1 \neq m_1'$, $m_2 = m_1'$, $m_2' = m_1$. Therefore (B.3.1) is just

$$\langle |F_{\mathrm{NB}}|^4\rangle = \sum_{m=1}^{N_{\mathrm{m}}} \langle |F_{\mathrm{NB},m}|^4\rangle + 2\sum_{m_1=1}^{N_{\mathrm{m}}} \sum_{\substack{m_2=1\\ m_2\neq m_1}}^{N_{\mathrm{m}}} \langle |F_{\mathrm{NB},m_1}|^2 |F_{\mathrm{NB},m_2}|^2\rangle. \tag{B.3.3}$$

In (B.3.3) $F_{\mathrm{NB},m}$ is defined as the summation of $F_{\mathrm{NB},n,m}$ over all the excited states. Since whether the m_1th molecule is excited is independent of whether the m_2 th molecule is excited if $m_1 \neq m_2$,

$$\langle |F_{\mathrm{NB},m_1}|^2 |F_{\mathrm{NB},m_2}|^2\rangle_{m_1\neq m_2} = \langle |F_{\mathrm{NB},m_1}|^2\rangle \langle |F_{\mathrm{NB},m_2}|^2\rangle. \tag{B.3.4}$$

Using the result for $\langle |F_{\mathrm{NB},m}|^4\rangle$ developed in the section B.2, we have

$$\langle |F_{\mathrm{NB}}|^4\rangle = 2\sum_{m=1}^{N_{\mathrm{m}}} (\langle |F_{\mathrm{NB},m}|^2\rangle)^2 + 2\sum_{m_1=1}^{N_{\mathrm{m}}} \sum_{\substack{m_2=1\\ m_2\neq m_1}}^{N_{\mathrm{m}}} \langle |F_{\mathrm{NB},m_1}|^2\rangle \langle |F_{\mathrm{NB},m_2}|^2\rangle, \tag{B.3.5}$$

namely

$$\langle |F_{\mathrm{NB}}|^4 \rangle = 2 \left(\sum_{m=1}^{N_{\mathrm{m}}} \langle |F_{\mathrm{NB},m}|^2 \rangle \right)^2 . \tag{B.3.6}$$

B.4. $\langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2 \rangle$ *computation: with spectrometer resolution considerations*

To compute the lattice factor squared $\langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2 \rangle$ with spectrometer resolution considerations, for mathematical simplicity, we assume a Gaussian-type spectrometer resolution function

$$R(\mathbf{q}) = \Delta \exp\left(-\frac{q_x^2}{\sigma_x^2} - \frac{q_y^2}{\sigma_y^2} - \frac{q_z^2}{\sigma_z^2} \right), \tag{B.4.1}$$

where Δ is a normalization factor. For the simple case of one molecule per unit cell and one excited state per molecule, with the resolution function defined by (B.4.1), for the subensemble of K molecules excited, we have

$$\begin{aligned} \langle (L_{\mathrm{NB,Re}(\mathbf{q})})^2 \rangle_{\mathrm{suben}} = \Bigg\langle \Bigg(& \sum_{k_1=1,k_1'=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_1)} - \mathbf{R}_{n(k_1')})} \otimes R(\mathbf{q}) \Bigg) \\ & \times \Bigg(\sum_{k_2=1,k_2'=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_2)} - \mathbf{R}_{n(k_2')})} \otimes R(\mathbf{q}) \Bigg) \Bigg\rangle_{\mathrm{suben}} , \end{aligned} \tag{B.4.2}$$

where $\otimes$ denotes a convolution. Carrying out the convolutions in (B.4.2), we have

$$\begin{aligned} & \langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2 \rangle_{\mathrm{suben}} \\ & = \Bigg\langle \Bigg(\sum_{k_1=1,k_1'=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_1)} - \mathbf{R}_{n(k_1')}) - \frac{1}{4}\left[\sigma_x^2 (\mathbf{R}_{n(k_1)} - \mathbf{R}_{n(k_1')})_x^2 + \sigma_y^2 (\mathbf{R}_{n(k_1)} - \mathbf{R}_{n(k_1')})_y^2 + \sigma_z^2 (\mathbf{R}_{n(k_1)} - \mathbf{R}_{n(k_1')})_z^2\right]} \Bigg) \\ & \times \Bigg(\sum_{k_2=1,k_2'=1}^{K} \mathrm{e}^{-\mathrm{i}\mathbf{q}(\mathbf{R}_{n(k_2)} - \mathbf{R}_{n(k_2')}) - \frac{1}{4}\left[\sigma_x^2 (\mathbf{R}_{n(k_2)} - \mathbf{R}_{n(k_2')})_x^2 + \sigma_y^2 (\mathbf{R}_{n(k_2)} - \mathbf{R}_{n(k_2')})_y^2 + \sigma_z^2 (\mathbf{R}_{n(k_2)} - \mathbf{R}_{n(k_2')})_z^2\right]} \Bigg) \Bigg\rangle_{\mathrm{suben}} . \end{aligned} \tag{B.4.3}$$

The average in the subensemble is just

$$\begin{aligned} & \langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2 \rangle_{\mathrm{suben}} \\ & = \frac{1}{C_N^K} \Bigg(\sum_{n_1=1,n_1'=1}^{N} \mathrm{e}^{\mathrm{i}\mathbf{q}(\mathbf{R}_{n_1} - \mathbf{R}_{n_1'}) - \frac{1}{4}\left[\sigma_x^2 (\mathbf{R}_{n_1} - \mathbf{R}_{n_1'})_x^2 + \sigma_y^2 (\mathbf{R}_{n_1} - \mathbf{R}_{n_1'})_y^2 + \sigma_z^2 (\mathbf{R}_{n_1} - \mathbf{R}_{n_1'})_z^2\right]} \Bigg) \\ & \times \Bigg(\sum_{n_2=1,n_2'=1}^{N} \mathrm{e}^{\mathrm{i}\mathbf{q}(\mathbf{R}_{n_2} - \mathbf{R}_{n_2'}) - \frac{1}{4}\left[\sigma_x^2 (\mathbf{R}_{n_2} - \mathbf{R}_{n_2'})_x^2 + \sigma_y^2 (\mathbf{R}_{n_2} - \mathbf{R}_{n_2'})_y^2 + \sigma_z^2 (\mathbf{R}_{n_2} - \mathbf{R}_{n_2'})_z^2\right]} \Bigg) \\ & \times \sum_{\substack{S_K \\ (n_1;n_1';n_2;n_2') \in S_K}} . \end{aligned} \tag{B.4.4}$$

To continue, we need the following results. The first is

$$\begin{aligned} \sum_{n=1}^{N} \mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_n - \frac{1}{4}\left[\sigma_x^2 (\mathbf{R}_n - \mathbf{R}_{n'})_x^2 + \sigma_y^2 (\mathbf{R}_n - \mathbf{R}_{n'})_y^2 + \sigma_z^2 (\mathbf{R}_n - \mathbf{R}_{n'})_z^2\right]} & = \exp\left(-\mathrm{i}\mathbf{q}\mathbf{R}_{n'} - \frac{q_x^2}{\sigma_x^2} - \frac{q_y^2}{\sigma_y^2} - \frac{q_z^2}{\sigma_z^2} \right) \\ & \times \frac{\pi^{d/2}}{\sigma_x \sigma_y \sigma_z V_{\mathrm{cell}}}, \end{aligned} \tag{B.4.5}$$

where $\mathbf{R}_{n'}$ is away from the boundary of the sample, and the second is

$$\sum_{n=1}^{N}\sum_{n'=1}^{N} \mathrm{e}^{-\frac{1}{2}[\sigma_x^2(\mathbf{R}_n-\mathbf{R}_{n'})_x^2+\sigma_y^2(\mathbf{R}_n-\mathbf{R}_{n'})_y^2+\sigma_z^2(\mathbf{R}_n-\mathbf{R}_{n'})_z^2]} = 2^{d/2}N\frac{\pi^{d/2}}{\sigma_x\sigma_y\sigma_z V_{\mathrm{cell}}}. \tag{B.4.6}$$

Using (B.4.5) and (B.4.6), we reach

$$\langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2\rangle_{\mathrm{suben}} = \left(1+\frac{(2\pi)^{d/2}}{N\sigma_x\sigma_y\sigma_z V_{\mathrm{cell}}}\right)\left(K-\frac{K^2}{N}\right)^2, \tag{B.4.7}$$

namely,

$$\langle (L_{\mathrm{NB,Re}}(\mathbf{q}))^2\rangle_{\mathrm{suben}} = \left(1+\frac{(2\pi)^{d/2}}{N\sigma_x\sigma_y\sigma_z V_{\mathrm{cell}}}\right)(\langle L_{\mathrm{NB}}(\mathbf{q})\rangle_{\mathrm{suben}})^2. \tag{B.4.8}$$

Appendix C. Sharpening a Patterson peak

Since the Patterson function of an atom has a peak width about twice the size of that atom and the number of peaks in a Patterson function for a molecule is $N(N-1)/2$, where N is the total number of atoms in the molecule (two peaks may have the same location), it is necessary to sharpen the Patterson peaks to reduce peak overlaps. The traditional way of sharpening a Patterson function leads to point-like peaks and is reported to have spurious peaks and is sensitive to truncation errors [23]. Here we shall introduce a simple technique to sharpen Patterson function peaks with a controlled sharpness level.

The sharpened electron density for a molecule can be expressed as

$$\rho_{\mathrm{s}}(\mathbf{r}) = \beta_{\mathrm{s}}^3\sum_{k=1}^{K}\sum_{j_k=1}^{N_k}\rho_k[\beta_{\mathrm{s}}(\mathbf{r}-\mathbf{R}_{n(j_k)})], \tag{C.1}$$

where $\rho_k(\mathbf{r})$ is the electron density for an atom of type k. N_k is the number of type-k atoms in the molecule, and $\mathbf{R}_{n(j_k)}$ is the position of the j_k th type-k atom. In (C.1) β_{s} is a positive number. When β_{s} is larger than one, (C.1) produces a sharpened electron density profile. Therefore, the Patterson function generated by $\rho_{\mathrm{s}}(\mathbf{r})$ has peak widths inversely proportional toβ_{s}. The Patterson function for the sharpened electron density is

$$P_{\mathrm{s}}(\mathbf{u}) = \int_{\infty}\rho_{\mathrm{s}}(\mathbf{r})\rho_{\mathrm{s}}(\mathbf{r}+\mathbf{u})\mathrm{d}\mathbf{r}, \tag{C.2}$$

namely,

$$P_{\mathrm{s}}(\mathbf{u}) = \frac{1}{(2\pi)^3}\int_{\infty}\tilde{\rho}_{\mathrm{s}}(\mathbf{q})\tilde{\rho}_{\mathrm{s}}(-\mathbf{q})\mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{u}}\mathrm{d}\mathbf{q}, \tag{C.3}$$

where $\tilde{\rho}_{\mathrm{s}}(\mathbf{q})$ is the Fourier transform of $\rho_{\mathrm{s}}(\mathbf{r})$, namely

$$\tilde{\rho}_{\mathrm{s}}(\mathbf{q}) = \sum_{k=1}^{K}\sum_{j_k=1}^{N_k}\tilde{\rho}_k\left(\frac{\mathbf{q}}{\beta_{\mathrm{s}}}\right)\mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(j_k)}}. \tag{C.4}$$

Since, for different atoms, $\tilde{\rho}_k(\mathbf{q})$ is roughly proportional to the atomic number when $|\mathbf{q}|$ is not larger than 10 $\mathrm{\AA}^{-1}$, for light atoms of carbon, oxygen, and nitrogen [24], (C.4) can be written as

$$\tilde{\rho}_{\mathrm{s}}(\mathbf{q}) = \tilde{\rho}_1(\mathbf{q}/\beta_{\mathrm{s}})\sum_{k=1}^{K}\sum_{j_k=1}^{N_k}Z_k/Z_1\mathrm{e}^{-\mathrm{i}\mathbf{q}\mathbf{R}_{n(j_k)}}, \tag{C.5}$$

where Z_k is the atomic number for the type-k atom. Therefore,

$$\tilde{\rho}_s(\mathbf{q})\tilde{\rho}_s(-\mathbf{q}) = \tilde{\rho}_1(\mathbf{q}/\beta_s)\tilde{\rho}_1(-\mathbf{q}/\beta_s)\left|\sum_{k=1}^{K}\sum_{j_k=1}^{N_k} Z_k/Z_1 e^{-i\mathbf{q}\mathbf{R}_{n(j_k)}}\right|^2. \tag{C.6}$$

In (C.6), the term $|\sum_{k=1}^{K}\sum_{j_k=1}^{N_k} Z_k/Z_1 e^{-i\mathbf{q}\mathbf{R}_{n(j_k)}}|^2$ can be obtained through the x-ray scattering differential cross section of the molecule through

$$\left|\sum_{k=1}^{K}\sum_{j_k=1}^{N_k} Z_k/Z_1 e^{-i\mathbf{q}\mathbf{R}_{n(j_k)}}\right|^2 = \frac{\frac{d\sigma(\mathbf{q})}{d\Omega}}{\Gamma|\tilde{\rho}_1(\mathbf{q})|^2}. \tag{C.7}$$

Therefore the sharpened Patterson function can be computed by

$$P_s(\mathbf{u}) = \frac{1}{(2\pi)^3 r_0^2}\int_{\infty}\frac{|\tilde{\rho}_1(\mathbf{q}/\beta_s)|^2 \times \frac{d\sigma(\mathbf{q})}{d\Omega}}{|\tilde{\rho}_1(\mathbf{q})|^2} e^{-i\mathbf{q}\mathbf{u}}\,d\mathbf{q}. \tag{C.8}$$

References

[1] Drenth J 1994 *Principles of Protein X-ray Crystallography* (New York: Springer)
[2] Moffat K 1989 *Annu. Rev. Biophys. Chem.* **18** 309
[3] Henderson R and Moffat K 1971 *Acta Crystallogr.* B **27** 1414
[4] Terwilliger T C and Berendzen J 1995 *Acta Crystallogr.* D **51** 609
[5] Guinier A 1963 *X-ray Diffraction* (San Francisco, CA: Freeman)
[6] Moffat K 1998 *Acta Crystallogr.* A **54** 833
[7] Warren B E 1990 *X-ray Diffraction* (New York: Dover)
[8] Miao J, Charalambous P, Kirz J and Sayre D 1999 *Nature* **400** 342
[9] Miao J, Ishikawa T, Johnson B, Anderson E H, Lai B and Hodgson K O 2002 *Phys. Rev. Lett.* **89** 088303
[10] Miao J, Ishikawa T, Anderson E H and Hodgson K O 2003 *Phys. Rev.* B **67** 174104
[11] He H, Marchesini S, Howells M, Weierstall U, Hembree G and Spence J C H 2003 *Acta Crystallogr.* A **59** 143
[12] Marchesini S, He H, Chapman H N, Hau-Riege S P, Noy A, Howells M R, Weierstall U and Spence J C H 2003 *Phys. Rev.* B **68** 140101(R)
[13] Skoze A 1999 *Chem. Phys. Lett.* **313** 777
[14] Fienup J R 1982 *Appl. Opt.* **21** 2758
[15] Hansen J P and McDonald I R 1986 *Theory of Simple Liquid* (New York: Academic)
[16] Caspar D L D, Clarage J, Salunke D M and Clarage M 1988 *Nature* **332** 659
[17] Hura G, Russo D, Glaeser R M, Head-Gordon T, Krack M and Parrinello M 2003 *Phys. Chem. Chem. Phys.* **5** 1981
[18] Treacy M M and Gibson J M 1996 *Acta Crystallogr.* A **52** 212
[19] Srajer V, Teng T, Ursby T, Pradervand C, Ren Z, Adachi S, Schildkamp W W, Bourgeois D, Wulff M and Moffat K 1996 *Science* **274** 1726
[20] Dauter Z 1999 *Acta Crystallogr.* D **55** 1703
[21] Ren Z, Bourgeois D, Helliwell J R, Moffat K, Srajer V and Stoddard B L 1999 *J. Synchrotron Radiat.* **6** 891
[22] Hayakawa M and Cohen J B 1975 *Acta Crystallogr.* A **31** 635
[23] Viterbo D 2002 *Fundamentals of Crystallography* ed C Giacovazzo (Oxford: Oxford University Press) pp 413–501
[24] Brown P J 1999 *(IUCr) International Table for Crystallography* vol C (New York: Academic) pp 548–89