CrossMark

Processing of X-ray snapshots from crystals in random orientations

Wolfgang Kabsch

Max-Planck-Institut für medizinische Forschung, Jahnstrasse 29, D-69120 Heidelberg, Germany

Correspondence e-mail:
kabsch@mpimf-heidelberg.mpg.de

A functional expression is introduced that relates scattered X-ray intensities from a still or a rotation snapshot to the corresponding structure-factor amplitudes. The new approach was implemented in the program *nXDS* for processing monochromatic diffraction images recorded by a multi-segment detector where each exposure could come from a different crystal. For images containing indexable spots, the intensities of the expected reflections and their variances are obtained by profile fitting after mapping the contributing pixel contents to the Ewald sphere. The varying intensity decline owing to the angular distance of the reflection from the surface of the Ewald sphere is estimated using a Gaussian rocking curve. This decline is dubbed 'Ewald offset correction', which is well defined even for still images. Together with an image-scaling factor and other corrections, an explicit expression is defined that predicts each recorded intensity from its corresponding structure-factor amplitude. All diffraction parameters, scaling and correction factors are improved by post-refinement. The ambiguous case of a lower point group than the lattice symmetry is resolved by a method reminiscent of the technique of 'selective breeding'. It selects the indexing alternative for each image that yields, on average, the highest correlation with intensities from all other images. Processing a test set of rotation images by *XDS* and treating the same images by *nXDS* as snapshots of crystals in random orientations yields data of comparable quality, clearly indicating an anomalous signal from Se atoms.

1. Introduction

The availability of free-electron lasers (FELs) as a source of ultrabright X-ray pulses of femtosecond duration has provided a new approach for collecting diffraction data that makes extremely small and radiation-sensitive samples accessible to structural studies. It is expected that this new X-ray source will strongly contribute to our knowledge of the large group of biological objects such as membrane proteins that are difficult to grow as macroscopic crystals.

Destruction of the irradiated object by a single pulse is much slower than the pulse duration, so that data can be collected at room temperature before the sample deteriorates owing to radiation damage. As a consequence many diffraction snapshots are required, each from a different sample in a random orientation, to yield a complete data set. For crystalline samples diffraction data are partially recorded on still images, which presents a challenging task to processing that traditional crystallographic programs were not designed to handle.

This has led to the development of new software suites: (i) *CrystFEL* (White *et al.*, 2012, 2013), which uses *MOSFLM*

(Leslie & Powell, 2007; Powell *et al.*, 2013) or *DirAx* (Duisenberg, 1992) for reflection indexing and merges individual intensity measurements (Kirian *et al.*, 2010, 2011) to estimate the integrated intensity of each unique reflection, and (ii) the *cctbx.xfel* (Sauter *et al.*, 2013; Hattne *et al.*, 2014) and *cctbx.spotfinder* (Zhang *et al.*, 2006) components within the *Computational Crystallography Toolbox* (*cctbx*). These programs rely on the Monte Carlo method for estimating reflection intensities, thereby assuming that other factors such as fluctuations in the incident-beam intensity, wavelength and spectrum as well as the irradiated volume of the specimen 'integrate out' upon averaging (White *et al.*, 2012; Kirian *et al.*, 2011). It has recently been demonstrated (Barends *et al.*, 2014) that a *de novo* crystal structure can be determined from X-ray free-electron laser data analyzed by *CrystFEL*. The data comprise about 60 000 indexed images out of 2.4 million snapshots from a lysozyme heavy-atom derivative that gives a strong anomalous signal from two Gd atoms per asymmetric unit (Barends *et al.*, 2014).

Another complication that is not encountered in traditional data collection by the rotation method is the ambiguity in the choice of unit-cell basis vectors in cases where the lattice symmetry is higher than the crystal symmetry. The problem arises because the reflections recorded by each snapshot are indexed independently and their partial intensities are too inaccurate for any decision-making based on correlations involving different exposures (White *et al.*, 2012). To circumvent this problem, *CrystFEL* generates a perfectly twinned data set of higher symmetry by merging all snapshots, which renders the subsequent structure solution more difficult. Recently, new methods for breaking the indexing ambiguity have been described (Brehm & Diederichs, 2014) and the detour *via* artificially twinned data sets appears to no longer be necessary.

Some of the problems in processing FEL snapshots had already been encountered more than 35 years ago when trying to solve the first virus crystal structures. Often, on account of radiation damage, the crystal had to be replaced after a single exposure covering only a small rotation range. This resulted in many partially recorded reflections and their complete intensities had to be estimated. A solution to this problem, the 'post-refinement' method, was developed (Schutt & Winkler, 1977; Rossmann *et al.*, 1979; Harrison *et al.*, 1985; Rossmann, 1985) to derive complete intensities from refined estimates for the fractions of observed intensity, the 'partiality'. For rotation images the partiality of each reflection can always be calculated as a function of orientation, the unit-cell metric, the mosaic spread of the crystal and a model of the reflection profile. The idea of the 'post-refinement' method is to improve these parameters after all images have been processed by using symmetry-related, fully recorded reflections for reference.

As an alternative to the processing of snapshots by Monte Carlo integration, the software package *nXDS* has been developed that resumes the old 'post-refinement' idea, now enriched by a model for the recorded snapshot intensities as a function of their corresponding structure-factor amplitudes

and also including the correct treatment of stills. The unique reflection intensities and diffraction parameters for all crystals are refined simultaneously to minimize the discrepancy between observed and expected spot centroids and recorded intensities. The indexing ambiguity is resolved by a method reminiscent of 'selective breeding'. It selects the indexing alternative for each image that yields, on average, the highest correlation with symmetry-related reflection intensities from all other images. The implementation of *nXDS* is derived from the standard *XDS* package (Kabsch, 2010*a*,*b*) in its present version, which includes the handling of multi-segment detectors.

2. Description of the diffraction experiment

Any right-handed orthonormal *laboratory* coordinate system $\{\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3\}$ can be chosen with its origin at the intersection point of the beam and the crystal. This reference system serves to specify the following.

(i) The monochromatic incident-beam wavevector $\mathbf{S}_0$ (wavelength λ, $|\mathbf{S}_0| = 1/\lambda$) pointing from the X-ray source towards the crystal.

(ii) The fixed rotation axis $\mathbf{m}_2$ passing through the origin of the *laboratory* system. During X-ray exposure the crystal is uniformly rotated by some angle Δ_φ around this axis. Still images are treated as a limiting case where $\Delta_\varphi = 0$ and the concept of a rotation axis loses its meaning.

(iii) The right-handed set of unit-cell basis vectors $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ of the single crystal and the associated reciprocal basis $\{\mathbf{b}_1^*, \mathbf{b}_2^*, \mathbf{b}_3^*\}$ such that any reciprocal-lattice point can be expressed as $\mathbf{p}_0^* = h\mathbf{b}_1^* + k\mathbf{b}_2^* + l\mathbf{b}_3^*$ where h, k, l are integers.

(iv) The right-handed orthonormal *detector* system $\{\mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3\}$ imagined to be fixed in the instrument and translated by the origin vector $\mathbf{D}_0$.

The detector consists of one or several rectangular, planar X-ray-sensitive segments. Their position and orientation is specified with respect to the *detector* system, which renders the specification independent of any detector movements. For a well calibrated instrument this greatly reduces the number of parameters that need to be refined during data processing.

A right-handed orthonormal *segment* system $\{\mathbf{d}_1', \mathbf{d}_2', \mathbf{d}_3'\}$ and an origin vector $\mathbf{d}_0'$ are specified for each segment with respect to the *detector* coordinate system. The X-ray-sensitive area consists of identical pixels of sizes Q_X, Q_Y (mm) along $\mathbf{d}_1'$ and $\mathbf{d}_2'$, respectively.

Expressing the segment origin $\mathbf{d}_0'$ in terms of $\{\mathbf{d}_1', \mathbf{d}_2', \mathbf{d}_3'\}$, the location $\mathbf{x}'$ of a segment pixel x, y with respect to the *detector* system is given by

$$\begin{aligned}
\mathbf{x}' &= xQ_X\mathbf{d}_1' + yQ_Y\mathbf{d}_2' + \mathbf{d}_0' \\
&= (x - x_0')Q_X\mathbf{d}_1' + (y - y_0')Q_Y\mathbf{d}_2' + F'\mathbf{d}_3'.
\end{aligned} \quad (1)$$

Thus, the *laboratory* coordinates x of the same pixel x, y are

$$\begin{aligned}
\mathbf{x} &= \mathbf{D}_0 + (\mathbf{D}_1|\mathbf{D}_2|\mathbf{D}_3)\mathbf{x}' \\
&= (x - x_0)Q_X\mathbf{d}_1 + (y - y_0)Q_Y\mathbf{d}_2 + F\mathbf{d}_3,
\end{aligned} \quad (2)$$

where

$$\mathbf{d}_\nu = (\mathbf{D}_1|\mathbf{D}_2|\mathbf{D}_3)\mathbf{d}'_\nu, \quad (\nu = 1, 2, 3) \tag{3}$$

and

$$x_0 = x'_0 - \mathbf{D}_0 \cdot \mathbf{d}_1/Q_X,$$
$$y_0 = y'_0 - \mathbf{D}_0 \cdot \mathbf{d}_2/Q_Y,$$
$$F = F' + \mathbf{D}_0 \cdot \mathbf{d}_3. \tag{4}$$

The segment plane is found at a distance $|F|$ from the crystal. F has a negative sign if $\mathbf{d}_3$ points towards the crystal.

For accurate integration, the spot shape and extent are modelled as Gaussian distributions involving two parameters: the standard deviations of the reflecting range (mosaicity), σ_M, and the combined effects of beam divergence and mosaicity, σ_D.

3. Processing steps of *nXDS*

The program package *nXDS* was developed for automatic determination of scaled and fully corrected reflection intensities from monochromatic X-ray diffraction images, each from a different crystal in a random orientation. *nXDS* uses many ideas, routines and the overall structure of the rotation data-processing program *XDS* (Kabsch, 2010*a,b*) and includes numerous new routines for the efficient evaluation of a large number of snapshots. The algorithms developed for integration, scaling and post-refinement are described in the following.

The snapshots are processed by *nXDS* in seven steps, named *XYCORR*, *INIT*, *COLSPOT*, *POWDER*, *IDXREF*, *INTEGRATE* and *CORRECT*, which are called in succession by *nXDS*. Information is communicated between the steps by files, which allows the repetition of selected steps with a different set of input parameters without rerunning the whole program.

3.1. *XYCORR*

If necessary for the detector being used, this step computes lookup tables of spatial corrections for each detector pixel. In subsequent data-processing steps, when the true coordinates of a pixel with respect to the laboratory coordinate system are needed, the correction values for the X and Y coordinates are retrieved from the tables and added to the pixel's array coordinates in the data image. This step is essentially the same as that used by *XDS*.

3.2. *INIT*

Three lookup tables are determined here that are required by the subsequent processing steps for classifying pixels in the data images as untrusted, background or belonging to a diffraction spot. This step is virtually the same as that used by *XDS*.

3.3. *COLSPOT*

Strong diffraction spots occurring in the data images are located and their centroids are saved in a file. Compared with the corresponding step in *XDS* a simplification resulted for snapshots from the fact that there are no neighbouring images with contributions to the same spot.

3.4. *POWDER*

The origin of the *detector* system can often be found from a powder pattern generated from the spots located in the previous step. An incorrectly specified origin could lead to a misindexing of reflections, a risk that is particularly high for processing snapshots. As the use of multi-segment detectors can prevent direct recognition of the powder circles, the following method was devised.

A plane with the incident-beam vector $\mathbf{S}_0$ as its normal is constructed from a second vector $\mathbf{t} = (1, 1, 1)$. To assure that $\mathbf{t}$ is non-collinear with $\mathbf{S}_0$, one of its components is reset to 0, namely the component corresponding to the maximum absolute value of the components of $\mathbf{S}_0$. If this is not unique, the first occurrence of the maximum is taken. For example, $\mathbf{t} = (1, 0, 1)$ if the y coordinate is the first occurrence of the absolutely largest component of $\mathbf{S}_0$. A right-handed ortho-normal system $\{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3\}$ can then be defined by

$$\mathbf{t}_1 = \mathbf{t} \times \mathbf{S}_0/|\mathbf{t} \times \mathbf{S}_0|, \quad \mathbf{t}_2 = \mathbf{S}_0 \times \mathbf{t}_1/|\mathbf{S}_0 \times \mathbf{t}_1|, \quad \mathbf{t}_3 = \mathbf{S}_0/|\mathbf{S}_0|. \tag{5}$$

Now, for a spot located at pixel coordinates x, y in the *COLSPOT* step a scattering vector $\mathbf{S}$ can be calculated,

$$\mathbf{x} = (x - x_0)Q_X\mathbf{d}_1 + (y - y_0)Q_Y\mathbf{d}_2 + F\mathbf{d}_3,$$
$$\mathbf{S} = \mathbf{x}/\lambda|\mathbf{x}|, \tag{6}$$

which intersects the powder plane at unit distance at coordinates

$$\mathbf{S} \cdot \mathbf{t}_1/\mathbf{S} \cdot \mathbf{t}_3, \quad \mathbf{S} \cdot \mathbf{t}_2/\mathbf{S} \cdot \mathbf{t}_3. \tag{7}$$

In the ideal case the scattering vectors mark a set of concentric rings: the powder pattern. The centre of the rings should be at the tip of the incident beam, which is the image centre. Often the centre of the powder rings is instead found to be offset from its ideal position, which can be interpreted to result from an incorrect value for the origin of the detector coordinate system $\mathbf{D}_0$.

3.5. *IDXREF*

For each snapshot this step tries to find a crystal lattice that explains the observed diffraction spots by assigning indices to them and refines all parameters. The step returns a list of the successfully processed images and the indexed spots as well as the corresponding diffraction parameters. These parameters are refined for accurate prediction and integration of all reflections that are expected to occur in each snapshot.

Short real-space lattice vectors are determined by the method of Steller *et al.* (1997). Extraction of a reduced cell, Bravais lattice determination and indexing of the observed spots is performed as described by Kabsch (1993). The computational methods were taken from *XDS*, with the exception of the refinement procedure, which had to be adapted for handling stills as well, where the concept of a rotation axis has no meaning. This is achieved by using a

different refinement target function. Instead of a rotation angle about a fixed camera axis, the smallest angle is used by which a reciprocal-lattice point could reach the Ewald sphere by some unrestricted rotation. A closed analytical expression for this angle is derived below and it is subsequently shown how it is used in the refinement procedure.

3.5.1. Rotating a reciprocal-lattice point to the Ewald sphere. We would like to find the smallest rotation that moves a reciprocal-lattice point onto the Ewald sphere. Let $\mathbf{p}_0^*$ denote any arbitrary reciprocal-lattice vector and $\mathbf{S}_0$ denote the incident-beam wavevector of length $1/\lambda$ (λ is the wavelength) pointing from the X-ray source towards the crystal. Diffraction occurs along the wavevector $\mathbf{S}$ when the crystal is rotated so that $\mathbf{p}_0^*$ is changed into $\mathbf{p}^*$ on the Ewald sphere satisfying the Laue equations

$$\mathbf{S} = \mathbf{S}_0 + \mathbf{p}^*, \quad \mathbf{S}^2 = \mathbf{S}_0^2 \Longrightarrow \mathbf{p}^{*2} = -2\mathbf{S}_0 \cdot \mathbf{p}^* = \mathbf{p}_0^{*2}. \quad (8)$$

The distance vector $\boldsymbol{\Delta} = \mathbf{p}^* - \mathbf{p}_0^*$ depends on the rotation used and thus is not unique.

Extending earlier work (Kabsch, 1988b), a unique rotation can be found for a given $\mathbf{S}_0$, $\mathbf{p}_0^*$ that yields the shortest distance vector $\boldsymbol{\Delta} = \mathbf{p}^* - \mathbf{p}_0^*$, provided $|\mathbf{p}_0^*| < 2|\mathbf{S}_0|$ and $|\mathbf{p}_0^* \cdot \mathbf{S}_0| < |\mathbf{p}_0^*||\mathbf{S}_0|$.

The unique rotation can be obtained by minimizing the function

$$f(\boldsymbol{\Delta}) = \boldsymbol{\Delta}^2 + \mu_1[(\mathbf{S}_0 + \mathbf{p}_0^* + \boldsymbol{\Delta})^2 - \mathbf{S}_0^2] + \mu_2[(\mathbf{p}_0^* + \boldsymbol{\Delta})^2 - \mathbf{p}_0^{*2}], \quad (9)$$

where the two constraints on the solution are enforced by the Lagrange multipliers μ_1, μ_2. At the minimum the gradient of $f(\boldsymbol{\Delta})$ must vanish and the matrix of second derivatives must be positive definite,

$$\frac{1}{2}\frac{\partial f}{\partial \boldsymbol{\Delta}} = \boldsymbol{\Delta} + \mu_1(\mathbf{S}_0 + \mathbf{p}_0^* + \boldsymbol{\Delta}) + \mu_2(\mathbf{p}_0^* + \boldsymbol{\Delta})$$

$$= 0 \Longrightarrow \mathbf{p}^* = \frac{\mathbf{p}_0^* - \mu_1\mathbf{S}_0}{1 + \mu_1 + \mu_2},$$

$$\frac{1}{2}\frac{\partial^2 f}{\partial \boldsymbol{\Delta}\partial \boldsymbol{\Delta}} = (1 + \mu_1 + \mu_2)\mathbf{1} \Longrightarrow 1 + \mu_1 + \mu_2 > 0. \quad (10)$$

The Lagrange multipliers are adjusted so that the constraints are satisfied. At a vanishing gradient of f, the Laue equations are satisfied if

$$\mathbf{p}_0^{*2} = -2\mathbf{S}_0 \cdot \mathbf{p}^* = 2\frac{\mu_1\mathbf{S}_0^2 - \mathbf{S}_0 \cdot \mathbf{p}_0^*}{1 + \mu_1 + \mu_2}, \quad (11)$$

which implies

$$\frac{1}{1 + \mu_1 + \mu_2} = \frac{\mathbf{p}_0^{*2}}{2(\mu_1\mathbf{S}_0^2 - \mathbf{S}_0 \cdot \mathbf{p}_0^*)}. \quad (12)$$

Conservation of the length of the reciprocal-lattice point implies that

$$\mathbf{p}_0^{*2} = \mathbf{p}^{*2} = \frac{(\mathbf{p}_0^* - \mu_1\mathbf{S}_0)^2}{(1 + \mu_1 + \mu_2)^2} = \frac{(\mathbf{p}_0^* - \mu_1\mathbf{S}_0)^2\mathbf{p}_0^{*4}}{4(\mu_1\mathbf{S}_0^2 - \mathbf{S}_0 \cdot \mathbf{p}_0^*)^2}, \quad (13)$$

which leads to a quadratic equation for μ_1,

$$\mu_1^2 - 2\mu_1\frac{\mathbf{S}_0 \cdot \mathbf{p}_0^*}{\mathbf{S}_0^2} + \frac{4(\mathbf{S}_0 \cdot \mathbf{p}_0^*)^2 - \mathbf{p}_0^{*4}}{(4\mathbf{S}_0^2 - \mathbf{p}_0^{*2})\mathbf{S}_0^2} = 0. \quad (14)$$

The solution is

$$\mu_1 = \frac{\mathbf{S}_0 \cdot \mathbf{p}_0^*}{\mathbf{S}_0^2} \pm \frac{\mathbf{p}_0^{*2}}{2\mathbf{S}_0^2}\left[\frac{\mathbf{S}_0^2\mathbf{p}_0^{*2} - (\mathbf{S}_0 \cdot \mathbf{p}_0^*)^2}{\mathbf{S}_0^2\mathbf{p}_0^{*2} - \frac{1}{4}\mathbf{p}_0^{*4}}\right]^{1/2}. \quad (15)$$

From

$$\frac{1}{1 + \mu_1 + \mu_2} = \frac{\mathbf{p}_0^{*2}}{2(\mu_1\mathbf{S}_0^2 - \mathbf{S}_0 \cdot \mathbf{p}_0^*)} = \pm\left[\frac{\mathbf{S}_0^2\mathbf{p}_0^{*2} - \frac{1}{4}\mathbf{p}_0^{*4}}{\mathbf{S}_0^2\mathbf{p}_0^{*2} - (\mathbf{S}_0 \cdot \mathbf{p}_0^*)^2}\right]^{1/2} \quad (16)$$

it can be concluded that the positive sign refers to the minimum of $\boldsymbol{\Delta}^2$, while the negative sign corresponds to the maximum distance from the Ewald sphere. Using the abbreviations

$$A = \frac{1}{1 + \mu_1 + \mu_2} = \left[\frac{\mathbf{S}_0^2\mathbf{p}_0^{*2} - \frac{1}{4}\mathbf{p}_0^{*4}}{\mathbf{S}_0^2\mathbf{p}_0^{*2} - (\mathbf{S}_0 \cdot \mathbf{p}_0^*)^2}\right]^{1/2}$$

$$B = \frac{\mu_1}{1 + \mu_1 + \mu_2} = (A\mathbf{S}_0 \cdot \mathbf{p}_0^* + \mathbf{p}_0^{*2}/2)/\mathbf{S}_0^2, \quad (17)$$

the closest point $\mathbf{p}^*$ on the Ewald sphere reachable by a rotation of $\mathbf{p}_0^*$ is

$$\mathbf{p}^* = A\mathbf{p}_0^* - B\mathbf{S}_0, \quad \boldsymbol{\Delta} = \mathbf{p}^* - \mathbf{p}_0^*, \quad \mathbf{S} = \mathbf{S}_0 + \mathbf{p}^*. \quad (18)$$

Obviously, the vectors $\boldsymbol{\Delta}$, $\mathbf{S}$, $\mathbf{p}^*$ all lie in the plane defined by $\mathbf{p}_0^*$, $\mathbf{S}_0$. We conclude that $\mathbf{p}_0^*$ reaches the Ewald sphere at $\mathbf{p}^*$ by the shortest path if it is rotated about an axis parallel to the plane normal and passing through the origin of reciprocal space. The rotation axis is identical to the 'β-axis' (Schutt & Winkler, 1977) defined for each reflection to be perpendicular to both the incident-beam and diffracted-beam wavevectors.

In the case of the rotation method the crystal is forced to rotate about a fixed axis $\mathbf{m}_2$ which has a component ζ on the 'β-axis'. ζ is called the reflecting-range expansion factor (Kabsch, 2010b) because the angle to rotate a reflection to the Ewald sphere about $\mathbf{m}_2$ increases (by $1/\zeta$).

Finally, the angular deviation of the reciprocal-lattice point $\mathbf{p}_0^*$ from the Ewald sphere can be defined as the vector

$$\tau = \mathrm{rad}\boldsymbol{\Delta}/|\mathbf{p}_0^*|, \quad \mathrm{rad} = 180°/\pi. \quad (19)$$

3.5.2. Initial refinement. For each image, let j enumerate the n strong spots with observed centroids X_j', Y_j', detector segment identifying number s_j and reflection indices h_j, k_j, l_j. To each j, we denote the reciprocal-lattice point $\mathbf{p}_{0j}^* = h_j\mathbf{b}_1^* + k_j\mathbf{b}_2^* + l_j\mathbf{b}_3^*$, from which a diffracted beam wavevector $\mathbf{S}_j$, an angular deviation from the Ewald sphere τ_j and reflecting-range expansion factors ζ_j can be determined as described in the previous section. With the detector segment recording the strong spot at distance F^{s_j}, orientation $\mathbf{d}_1^{s_j}$, $\mathbf{d}_2^{s_j}$, $\mathbf{d}_3^{s_j}$ and origin $X_0^{s_j}$, $Y_0^{s_j}$, the residuals between the calculated and observed spot centroids are (segment specified in the *laboratory* system)

$$\Delta_X^j = X_0^{s_j} + F^{s_j}\mathbf{S}_j \cdot \mathbf{d}_1^{s_j}/\mathbf{S}_j \cdot \mathbf{d}_3^{s_j} - X_j',$$
$$\Delta_Y^j = Y_0^{s_j} + F^{s_j}\mathbf{S}_j \cdot \mathbf{d}_2^{s_j}/\mathbf{S}_j \cdot \mathbf{d}_3^{s_j} - Y_j'. \tag{20}$$

The goal of the refinement procedure is to find the detector parameters, incident-beam direction and unit-cell basis vectors that minimize the target function

$$E = w_X \sum_{j=1}^{n}(\Delta_X^j)^2 + w_Y \sum_{j=1}^{n}(\Delta_Y^j)^2 + w_\tau \sum_{j=1}^{n}\tau_j^2/[\tfrac{1}{3}(\zeta_j\Delta_\phi/2)^2 + \sigma_M^2]. \tag{21}$$

The positional part of the target function pushes the calculated spot positions towards the observed centroids, while the angular part minimizes the angular deviations from the Ewald sphere. The variance for the angular part is expected to increase with the size of the rotation range.

The residuals are expanded to first order in the parameter changes so that E becomes a quadratic function of these changes. Minimization then leads to a system of normal equations whose solution is used to update the parameters. During refinement, the cell metric obeys constraints imposed by the lattice symmetry by choosing an appropriate set of free parameters for representing the allowed changes of the basis vectors (details not shown).

The minimization procedure is repeated keeping the same set of weights w_X, w_Y, w_τ until convergence is reached. New weights are then determined so that

$$w_X = 1/\sum_{j=1}^{n}(\Delta_X^j)^2,$$
$$w_Y = 1/\sum_{j=1}^{n}(\Delta_Y^j)^2,$$
$$w_\tau = 1/\sum_{j=1}^{n}\tau_j^2/[\tfrac{1}{3}(\zeta_j\Delta_\varphi/2)^2 + \sigma_M^2]. \tag{22}$$

Refinement continues with the new weights until convergence is reached again. The whole refinement procedure is terminated upon convergence of the weights w_X, w_Y, w_τ.

3.6. INTEGRATE

Starting with the refined diffraction parameters for the successfully indexed data images, this step determines the recorded reflection intensities by two-dimensional profile fitting and saves the results on file for subsequent processing by the *CORRECT* step. No correction factors are applied except for compensating missing parts in the reflection profile owing to overlap with bad pixels or closely neighbouring reflections.

Similar to the integration procedure as implemented in *XDS* (Kabsch, 2010*b*), analysis of each image consists of determination of the reflection spot size and the mosaicity and refinement of the diffraction parameters using the strong spots. This is followed by determination of a two-dimensional reflection reference profile and integration by profile fitting for all expected reflections, including the weak reflections

3.6.1. Mapping image pixels to the Ewald sphere. As described earlier (Kabsch, 1988*b*, 2010*b*) and implemented in

the program *XDS* (Kabsch, 2010*a*), it is useful to represent the intensity distribution in a reflection-specific coordinate system $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ so that all reflections appear as if they had followed the shortest path through the Ewald sphere and were recorded on the surface of the sphere. This eliminates reflection-specific differences in the intensity profile caused by the oblique incidence of the diffracted beam on a flat detector segment or by crystal rotation around a fixed axis for a rotation data image.

For a reciprocal-lattice point $\mathbf{p}_0^*$ the nearest point $\mathbf{S}$ on the Ewald sphere is obtained as described above, so that a reflection-specific coordinate system can be defined as

$$\mathbf{e}_1 = \mathbf{S} \times \mathbf{S}_0/|\mathbf{S} \times \mathbf{S}_0|,$$
$$\mathbf{e}_2 = \mathbf{S} \times \mathbf{e}_1/|\mathbf{S} \times \mathbf{e}_1|,$$
$$\mathbf{e}_3 = (\mathbf{S} + \mathbf{S}_0)/|\mathbf{S} + \mathbf{S}_0|. \tag{23}$$

It has its origin at the terminus of the diffracted-beam wave-vector $\mathbf{S}$ and therefore could move depending on the specific path of $\mathbf{p}_0^*$ through the Ewald sphere. The unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are tangential to the Ewald sphere, while $\mathbf{e}_3$ is perpendicular to $\mathbf{e}_1$ and $\mathbf{p}^* = \mathbf{S} - \mathbf{S}_0$.

Diffraction along wavevector $\mathbf{S}'$ in the neighbourhood of $\mathbf{S}$ is recorded at pixel X', Y' in the detector plane. This pixel is at a distance D from the crystal. With the detector at a distance F, an orientation $\mathbf{d}_1$, $\mathbf{d}_2$ and an origin X_0, Y_0 of the detector plane, we have

$$D = [(X' - X_0)^2 + (Y' - Y_0)^2 + F^2]^{1/2},$$
$$\mathbf{S}' = [(X' - X_0)\mathbf{d}_1 + (Y' - Y_0)\mathbf{d}_2 + F\mathbf{d}_3]/(\lambda D), \tag{24}$$

The corresponding coordinates ε_1, ε_2, ε_3 in the reflection-specific system on the Ewald sphere are then

$$\varepsilon_1 = \mathbf{e}_1 \cdot (\mathbf{S}' - \mathbf{S})/|\mathbf{S}|,$$
$$\varepsilon_2 = \mathbf{e}_2 \cdot (\mathbf{S}' - \mathbf{S})/|\mathbf{S}|,$$
$$\varepsilon_3 = \mathbf{e}_3 \cdot (\mathbf{p}_0^* - \mathbf{p}^*)/|\mathbf{p}^*|. \tag{25}$$

In the case of a rotation image, a reciprocal-lattice point $\mathbf{p}_0^*$ needs a larger angle $\varepsilon_3/|\mathbf{e}_1 \cdot \mathbf{m}_2|$ to reach the Ewald sphere because the movement is restricted by the fixed rotation axis $\mathbf{m}_2$. The reflecting-range expansion factor $\zeta = \mathbf{e}_1 \cdot \mathbf{m}_2$ corrects for this effect. It is closely related to the reciprocal Lorentz correction factor for rotation images,

$$L^{-1} = |\mathbf{m}_2 \cdot (\mathbf{S} \times \mathbf{S}_0)|/(|\mathbf{S}| \cdot |\mathbf{S}_0|) = |\zeta \cdot \sin \angle(\mathbf{S}, \mathbf{S}_0)|. \tag{26}$$

The reflection-specific coordinate system defined above sets up a one-to-one correspondence between the points of a region R of the $\varepsilon_1\varepsilon_2$ plane tangential to the Ewald sphere and a region R' of the $X'Y'$ plane of the detector segment. The Jacobian of the transformation is

$$J = \frac{\partial(\varepsilon_1, \varepsilon_2)}{\partial(X', Y')} = \frac{\partial\varepsilon_1}{\partial X'}\frac{\partial\varepsilon_2}{\partial Y'} - \frac{\partial\varepsilon_1}{\partial Y'}\frac{\partial\varepsilon_2}{\partial X'}. \tag{27}$$

From

$$\frac{\partial\varepsilon_i}{\partial X'} = \mathbf{e}_i \cdot \frac{\partial\lambda\mathbf{S}'}{\partial X'}, \quad \frac{\partial\varepsilon_i}{\partial Y'} = \mathbf{e}_i \cdot \frac{\partial\lambda\mathbf{S}'}{\partial Y'} \quad (i = 1, 2) \tag{28}$$

one finds

$$J = \left(\mathbf{e}_1 \cdot \frac{\partial \lambda \mathbf{S}'}{\partial X'}\right)\left(\mathbf{e}_2 \cdot \frac{\partial \lambda \mathbf{S}'}{\partial Y'}\right) - \left(\mathbf{e}_1 \cdot \frac{\partial \lambda \mathbf{S}'}{\partial Y'}\right)\left(\mathbf{e}_2 \cdot \frac{\partial \lambda \mathbf{S}'}{\partial X'}\right)$$
$$= (\mathbf{e}_1 \times \mathbf{e}_2) \cdot \left(\frac{\partial \lambda \mathbf{S}'}{\partial X'} \times \frac{\partial \lambda \mathbf{S}'}{\partial Y'}\right). \tag{29}$$

Using

$$\frac{\partial \lambda \mathbf{S}'}{\partial X'} = \frac{\mathbf{d}_1}{D} - \frac{(\lambda \mathbf{S}')(X' - X_0)}{D^2}, \quad \frac{\partial \lambda \mathbf{S}'}{\partial Y'} = \frac{\mathbf{d}_2}{D} - \frac{(\lambda \mathbf{S}')(Y' - Y_0)}{D^2} \tag{30}$$

one finds

$$\frac{\partial \lambda \mathbf{S}'}{\partial X'} \times \frac{\partial \lambda \mathbf{S}'}{\partial Y'} = \frac{F}{D^3} \lambda \mathbf{S}' \tag{31}$$

and the Jacobian reduces to the simple expression

$$J = \frac{F}{D^3}(\lambda \mathbf{S}') \cdot (\lambda \mathbf{S}). \tag{32}$$

Thus, if the region R' of the $X'Y'$ plane of the detector covers a single pixel, the corresponding area R in the $\varepsilon_1 \varepsilon_2$ plane is the mean value of $|J|$ times the pixel area,

$$R = \iint_R d\varepsilon_1 \, d\varepsilon_2 = \iint_{R'} |J| \, dX' \, dY'$$
$$\simeq |J| \iint_{R'} dX' \, dY' = |J| \, R'. \tag{33}$$

3.6.2. Intensity recorded by a detector pixel. Because of crystal mosaicity and beam divergence, the intensity of a reflection is smeared around the diffraction maximum. Following earlier work (Kabsch, 1988b, 2010b), the fraction of the total reflection intensity found in the volume element $d\varepsilon_1 d\varepsilon_2 d\varepsilon_3$ at ε_1, ε_2, ε_3 is modelled as the product of two functions:

$$\omega(\varepsilon_1, \varepsilon_2, \varepsilon_3)d\varepsilon_1 d\varepsilon_2 d\varepsilon_3 = \omega_{12}(\varepsilon_1, \varepsilon_2)d\varepsilon_1 d\varepsilon_2 \cdot \omega_3(\varepsilon_3)d\varepsilon_3. \tag{34}$$

The first function $\omega_{12}(\varepsilon_1, \varepsilon_2)$ describes the expected intensity profile of the reflections mapped to their specific ε_1, ε_2 plane tangential to the Ewald sphere. Initially, Gaussians are assumed with the parameter σ_D modelling the combined effects of beam divergence and mosaicity as

$$\omega_{12}(\varepsilon_1, \varepsilon_2) = \frac{\exp(-\varepsilon_1^2/2\sigma_D^2)}{(2\pi)^{1/2}\sigma_D} \cdot \frac{\exp(-\varepsilon_2^2/2\sigma_D^2)}{(2\pi)^{1/2}\sigma_D}. \tag{35}$$

σ_D is estimated from the observed variance in intensity of scattered rays for the strong reflections and provides information on the spot width. In a second step the initial Gaussian form for ω_{12} is replaced by the superposition of the profiles of all strong reflections and is used for definition of the final integration region included in profile fitting of all reflections predicted to occur in the diffraction image.

The second function, the rocking curve $\omega_3(\varepsilon_3)$, models the dependency of the reflection intensity on the angular distance from the surface of the Ewald sphere as a Gaussian with the mosaicity σ_M of the crystal as the standard deviation. An estimate of σ_M is found as described previously (Kabsch, 2010b). If $\tau = \text{rad}|\boldsymbol{\Delta}|/|\mathbf{p}_0^*|$ is the angular deviation of the reciprocal-lattice point $\mathbf{p}_0^*$ from the Ewald sphere, the fraction

of intensity recorded on the image, i.e. the partiality of the reflection, is

$$q = \int_{-\varepsilon}^{\varepsilon} \omega_3(\varepsilon_3 - \tau) \, d\varepsilon_3 = \int_{-\varepsilon}^{\varepsilon} \frac{\exp[-(\varepsilon_3 - \tau)^2/2\sigma_M^2]}{(2\pi)^{1/2}\sigma_M} \, d\varepsilon_3. \tag{36}$$

The integration extends over the rotation range Δ_φ of the spindle during exposure of the image multiplied by the reflecting-range expansion factor ζ, which corrects for the increased path length of the reflection through the Ewald sphere when rotated around a fixed axis, i.e. $\varepsilon = \zeta \cdot \Delta_\varphi/2$. Using the dimensionless variables $t = \tau/2^{1/2}\sigma_M$ and $z = \varepsilon/2^{1/2}\sigma_M$ the partiality of the reflection can also be given in terms of the error function erf:

$$q(t, z) = \int_{-z}^{z} \frac{\exp[-(t - z')^2]}{(\pi)^{1/2}} \, dz' = [\text{erf}(t + z) - \text{erf}(t - z)]/2. \tag{37}$$

Together with the above expression for the Jacobian J, the expected fraction of reflection intensity recorded by pixel X', Y' in the detector plane is, according to our model,

$$\iint_R \omega_{12}(\varepsilon_1, \varepsilon_2) \, d\varepsilon_1 \, d\varepsilon_2 \int_{-\varepsilon}^{\varepsilon} \omega_3(\varepsilon_3 - \tau) \, d\varepsilon_3 \simeq \omega_{12}(\bar{\varepsilon}_1, \bar{\varepsilon}_2)|\bar{J}|R'q. \tag{38}$$

Here, $\omega_{12}(\bar{\varepsilon}_1, \bar{\varepsilon}_2)$ and $\bar{J}$ denote mean values in the region R of the $\varepsilon_1 \varepsilon_2$ plane that maps to the region R' of the pixel at X', Y' in the plane of the detector.

3.7. CORRECT

After resolving possible indexing ambiguities, the raw intensities of the (reindexed) reflections as obtained from the previous step are scaled and corrected by a modified post-refinement procedure and then saved in the final reflection output file for subsequent structure-solution software packages.

As described below, the new concept of the Ewald offset correction factor as a replacement for partiality links raw intensities to structure-factor amplitudes, rendering classical post-refinement applicable even to still snapshots.

3.7.1. Correction factors. A reflection intensity $\hat{I}$, as recorded on a snapshot and returned from the *INTEGRATE* step, is proportional to the 'true' intensity I (the squared structure-factor amplitude),

$$\hat{I} = CI, \quad C = \hat{C}T, \quad \hat{C} = QLPAO. \tag{39}$$

The correction factor C consists of a factor $\hat{C}$ that is not affected by possible indexing ambiguities and a factor T that accounts for different scale and resolution fall-off between the snapshots.

$\hat{C}$ is a product of factors modelling the Ewald offset correction (Q), Lorentz factor (L), polarization (P), air absorption (A) and sensor-thickness correction (O). These corrections are functions only of the diffraction parameters of the snapshot that recorded the reflection.

Ewald offset correction. The recorded intensities of the reflections of the same image would be directly comparable if

they simultaneously obeyed the Laue equations, *i.e.* $\tau = 0$ for all reflections. The idea is to correct the intensity by the estimated decline owing to the angular distance of the reflection from the surface of the Ewald sphere. Using the dimensionless variables $t = \tau/2^{1/2}\sigma_M$ and $z = \varepsilon/2^{1/2}\sigma_M$ ($\varepsilon = \zeta \cdot \Delta_\varphi/2$) this leads to the definition of the Ewald offset correction

$$Q(t, z) = q(t, z)/q(0, z) = \tfrac{1}{2}[\mathrm{erf}(t + z) - \mathrm{erf}(t - z)]/\mathrm{erf}(z),$$
$$Q(t, 0) = \lim_{z \to 0} Q(t, z) = \exp(-t^2), \qquad (40)$$

which is well defined even for still images ($z = 0$). $Q(t, z)$ can be calculated for each reflection from the incident-beam wavevector S_o, the reciprocal basis $\{b_1^*, b_2^*, b_3^*\}$, the mosaicity σ_M and the oscillation range Δ_φ.

Lorentz factor. Assuming an infinitely thin Ewald sphere, the Laue equations can be satisfied for a reciprocal-lattice point by a variation in the wavelength or the direction of the incident beam relative to the crystal. For an ideal crystal the scattered intensity is sharply concentrated around the direction of the diffraction maximum with a solid angle much smaller than the finite aperture of a detector pixel. Consequently, only integrated intensities can be observed: these are related to their squared structure factors by the Lorentz correction. Explicit forms of the corrections are available (see, for example, Zachariasen, 1945) for all conceivable methods of recording sharp diffraction maxima.

(i) *Rotation method.* For a fixed wavelength but a variable direction of incidence (accomplished by rotation of the crystal around a fixed axis and a fixed incident beam), the correction factor is

$$L = 1/|\zeta \sin(2\theta)| = |S||S_0|/|m_2 \cdot S \times S_0|, \quad 2\theta = \angle(S, S_0). \qquad (41)$$

For the special case that the rotation axis is perpendicular to both the incident and diffracted beams, *i.e.* $\zeta = 1$, the Lorentz correction simplifies to

$$L = 1/|\sin(2\theta)|. \qquad (42)$$

Still snapshots can be considered as a limiting case in which all reflections move infinitesimally through the Ewald sphere along their shortest routes.

(ii) *Laue method.* For a variable wavelength but a fixed direction of the incident beam, the correction factor is

$$L = 1/(2 \sin^2 \theta). \qquad (43)$$

(iii) *Powder method.* For a fixed wavelength but a variation of the incident beam with two degrees of freedom (accomplished by mosaic crystals), the correction factor is

$$L = 1/(4 \sin \theta). \qquad (44)$$

Lorentz correction is always a simple function of resolution and therefore does not affect the agreement among intensities of symmetry-related reflections.

Polarization. The intensity of scattering from the crystal is proportional to the polarization factor

$$P = \langle \sin^2 \varphi \rangle, \quad \varphi = \angle(E_0, S), \qquad (45)$$

where E_0 denotes the electrical field vector of the incident beam and S the diffracted beam wavevector. If n denotes the polarization plane normal and p the probability of finding the field vector E_0 in this plane,

$$P = \langle \sin^2 \varphi \rangle = p \sin^2 \varphi_1 + (1 - p) \sin^2 \varphi_2,$$
$$\varphi_1 = \angle(S_0 \times n, S),$$
$$\varphi_2 = \angle[(S_0 \times n) \times S_0, S]. \qquad (46)$$

For an unpolarized incident beam, the electrical field vector E_0 is found with equal probability pointing along $S_0 \times S$ or $(S_0 \times S) \times S_0$, so that

$$\varphi_1 = \pi/2, \quad \varphi_2 = \pi/2 - 2\theta, \quad 2\theta = \angle(S, S_0). \qquad (47)$$

This leads to

$$P = [1 + \cos^2(2\theta)]/2. \qquad (48)$$

The chosen parametrization allows description of the effect of polarization for most data-collection scenarios (Kahn *et al.*, 1982). Values for the two parameters n and p are provided by the user (not refined).

Air absorption. The recorded intensity is reduced from its 'true' value owing to air absorption of the diffracted beam by the factor

$$A = \exp(-\mu D), \qquad (49)$$

where μ denotes the fraction of intensity loss per millimetre and D is the distance (in millimetres) between the crystal and the position of the reflection spot on the detector segment. μ is a wavelength-dependent input constant (not refined).

Sensor-thickness correction. The recorded intensity is increased owing to the effect of the oblique incidence of the diffracted beam on a sensor of finite thickness. Let δ denote the thickness of the sensor and κ the fraction of intensity loss per millimetre. If the diffracted beam makes an angle ω with the segment normal, the probability of a photon penetrating the sensor undetected is $\exp(-\kappa\delta/\cos\omega)$. Thus, the correction

$$O = \frac{1 - \exp(-\kappa\delta/\cos\omega)}{1 - \exp(-\kappa\delta)} \qquad (50)$$

accounts for the oblique incidence of the diffracted beam. Values for δ and κ are provided by the user and are not refined.

Scaling and temperature factor. Using the parameters g and B, the factor

$$T = g \exp(B|p_0^*|^2) \qquad (51)$$

puts the 'true' intensity for reciprocal-lattice point p_0^* on the same scale as the observed one in the data image. The values of g and B adjust differences in beam intensity and in crystal disorder and volume in a resolution-dependent way. This correction factor is obtained by comparison of symmetry-equivalent reflections from different images and therefore depends on a consistent indexing choice. For resolving possible indexing ambiguities, correlations between equivalent reflection pairs from different snapshots must be computed (Brehm & Diederichs, 2014). To manage a

potentially large number of snapshots and reflections, an efficient solution was implemented in *nXDS*.

3.7.2. Efficient calculation of correlation factors. A procedure was developed for calculating correlation factors in which the number of operations is proportional to the total number of recorded reflections of the compared snapshots. The procedure consists of two parts: initialization of the data structures and calculation of the correlation factor between the unique intensity estimates from the compared snapshots.

Initialization.

(i) For the given space group and unit-cell parameters all possible unique reflections within a given resolution range are generated. A prime number is determined that is slightly larger than twice the number of possible unique reflections and is used to allocate space for a hash table of this size.

(ii) For each possible unique reflection a positive 64-bit integer is constructed from its unique indices and assigned a definite address in the table using the technique of hash-key transformations with quadratic probing to resolve key collisions (Wirth, 1976).

(iii) For each reflection from the snapshots the unique indices are determined and coded by their hash-table address, which is saved as an auxiliary reflection attribute. Thus, two reflections are symmetry-related only if they have identical hash-table addresses.

(iv) Four auxiliary arrays of the size of the hash table are allocated: two for each data set to be compared. They are needed for calculating unique intensities and their variances for the reflections of the two data sets.

Correlation.

(i) Unique intensities and variances are estimated from symmetry-related reflections of the first snapshot by updating the contents of their associated hash addresses in the first two auxiliary arrays.

(ii) For the intensity data from the second snapshot the procedure is repeated, this time updating the second two auxiliary arrays only if there is a positive entry from the first snapshot at the same hash address in the first two auxiliary arrays.

(iii) Pairs of corresponding unique reflection intensities are obtained easily by scanning the second two auxiliary arrays for a positive contents.

Thus, the total number of operations for calculating the correlation factor between one snapshot and all others is only proportional to the total number of recorded reflections of the snapshots.

3.7.3. Indexing alternatives. For a given space group and unit-cell parameters, a reciprocal cell basis in some reference orientation (Kabsch, 1988*a*) is first defined and serves as a reference basis. As each image is indexed and integrated independently, it often happens that the reflection indices refer to different (reduced) cells. The problem is to find for each image the set of possible reindexing transformations that allow the reflections to be described in terms of a rotated version of the reference basis.

This set of reindexing transformations is determined for each image using the following procedure. From the reciprocal basis vectors used for indexing the reflections of the image in the *INTEGRATE* step, a reduced cell and its reciprocal cell are determined. The three reciprocal vectors thus obtained are considered as reflections that need to be indexed with respect to a rotated version of the reciprocal reference basis. Possible indices of the reduced-cell reciprocal vectors with respect to the reference basis are found by simply testing all possibilities involving indices absolutely smaller than 4. For each assignment of indices a residual error for the best superposition with the reference basis is determined (Kabsch, 1976, 1978) and is used as a measure of the quality of the indexing. Index assignments related by symmetry of the reference cell are omitted, so that a list of symmetry-independent interpretations remains. This list is sorted by increasing r.m.s. of the superposition with the reciprocal reference cell. Entries in the list with an r.m.s. larger than some multiple of that of the first item are omitted.

The list may be empty if no reasonable interpretation is found for the basis vectors of the image. In this case the image is omitted from further calculations.

3.7.4. Resolving the indexing ambiguity. Ideally, only one symmetry-independent solution remains, identified using the above procedure solely by geometrical considerations. However, for merohedral and pseudo-merohedral crystals, where the lattice symmetry is higher than the symmetry of the point group, more than one choice for the reindexing transformation exists. In *nXDS* the indexing ambiguity can be resolved by using one of two methods. Both methods rely on correlation factors between intensities that have been corrected by $\hat{C}$ for various effects as described above. Note that the scaling corrections T are not needed here.

Comparison with a reference. For each snapshot all possible reindexing transformations are tested and the indexing choice yielding the largest correlation factor with the given reference data set is selected. Here, symmetry-equivalent reflection intensities from the snapshot as well as from the reference data are merged separately prior to calculation of the correlation factor. Moreover, an initial scaling factor is determined at little additional computational effort that puts the intensities from each snapshot on the level of the reference.

Selective breeding. If no external reference data set is available, a solution is found by a method that is reminiscent of the technique of selective breeding. The method initiates a cyclic procedure with some arbitrary indexing transformation from the list of possibilities assumed by each snapshot. For each cycle the following steps are carried out.

(i) For each snapshot all of its possible indexing choices are tested in succession, with the reindexed reflection intensities treated as a hypothetical reference data set. The mean value of the correlation factors with all other snapshots is calculated, and the running number of the reindexing choice yielding the largest correlation factor is saved. If several choices result in the same value for the correlation maximum, the first one in the list is selected.

(ii) At the end of the cycle the list of optimal indexing choices just determined is used and replaces the previous selection. For some snapshots the new running number of the

reindexing choice may differ from that of the previous cycle. If none of the snapshots needs to be assigned to a different indexing choice than before, the cyclic procedure terminates.

This procedure usually terminates within ten cycles. The moderate amount of storage needed is proportional to the number of snapshots and to the size of the hash table. In addition one array is needed that keeps the hash-table address of the unique indices for each reflection of the whole set of images. This speeds up the computation of correlation factors between all pairs of snapshots. Within each cycle these computations can be carried out very efficiently and in parallel by a team of processors.

3.7.5. Scaling. We assume that all reflections have been consistently indexed. The reflection intensities of each image are still on different scales owing to differences in the intensity of the incident beam or irradiated crystal volume. Determination of the scale factors by the method described here is based on intensity estimates for the unique reflections occurring on each image,

$$\bar{I} = \sum_\nu \hat{I}_\nu \hat{C}_\nu / \hat{\sigma}_\nu^2 / \sum_\nu \hat{C}_\nu^2 / \hat{\sigma}_\nu^2, \quad \bar{\sigma}^2 = 1 / \sum_\nu \hat{C}_\nu^2 / \hat{\sigma}_\nu^2, \quad (52)$$

where ν enumerates the symmetry-equivalent reflections and $\hat{I}_\nu$ and $\hat{\sigma}_\nu$ the recorded intensities on the same image and their estimated standard deviations, respectively. $\hat{C}_\nu$ denotes the correction factors as described above. A scaling correction factor for each image is determined by least-squares minimization using common unique reflections with a positive intensity that occur in more than one image.

Let h enumerate the n_h different unique indices of the reflections involved in scaling and l enumerate the n_l images from which the reflections come. Let j enumerate the n unique reflections included. To each j, we denote intensity $\bar{I}_j > 0$, variance $\bar{\sigma}_j^2$, unique reflection index h_j and image l_j. The goal of the scaling procedure is to find factors $g_l > 0$ and mean intensities $I_h > 0$ by minimizing the target function (Hamilton *et al.*, 1965)

$$\psi(g, I) = \sum_{j=1}^n \frac{(\bar{I}_j - g_{l_j} I_{h_j})^2}{\bar{\sigma}_j^2}. \quad (53)$$

To guarantee success of the solution method for a large set of snapshots and possibly large variations in their scaling factors, the target function is modified by using logarithms,

$$\Psi(G, J) = \sum_{j=1}^n w_j (\hat{J}_j - G_{l_j} - J_{h_j})^2, \quad (54)$$

defining $\hat{J}_j = \ln \bar{I}_j$, $w_j = \bar{I}_j^2 / \bar{\sigma}_j^2$, $G_l = \ln g_l$ and $J_h = \ln I_h$. This target function is quadratic, with a constant matrix of second derivatives and positive diagonal elements:

$$\frac{\partial \Psi(G, J)}{\partial G_l} = -2 \sum_{j=1}^n w_j (\hat{J}_j - G_{l_j} - J_{h_j}) \delta_{ll_j},$$

$$\frac{\partial \Psi(G, J)}{\partial J_h} = -2 \sum_{j=1}^n w_j (\hat{J}_j - G_{l_j} - J_{h_j}) \delta_{hh_j},$$

$$\frac{\partial^2 \Psi(G, J)}{\partial G_{l'} \partial G_l} = \delta_{ll'} 2 \sum_{j=1}^n w_j \delta_{ll_j},$$

$$\frac{\partial^2 \Psi(G, J)}{\partial J_{h'} \partial J_h} = \delta_{hh'} 2 \sum_{j=1}^n w_j \delta_{hh_j}. \quad (55)$$

Therefore, unique directional minimizers can be defined by equating the gradients to zero:

$$\bar{G}_l = \sum_{j=1}^n w_j [\hat{J}_j - J_{h_j}(G)] \delta_{ll_j} / \sum_{j=1}^n w_j \delta_{ll_j},$$

$$J_h(G) = \sum_{j=1}^n w_j (\hat{J}_j - G_{l_j}) \delta_{hh_j} / \sum_{j=1}^n w_j \delta_{hh_j}. \quad (56)$$

The target function is monotonically reduced by a cyclic procedure of alternating minimizations along the J and G directions. The procedure is initiated at $G_l = 0$, $J_h(G = 0)$. In each following cycle new scaling factors are found as

$$G'_l - G_l = c(\bar{G}_l - G_l), \quad (57)$$

with the step size c chosen to maximize the reduction of the target function. Expansion of the quadratic target function at the point $\bar{G}, J(G)$ yields

$$\Psi[G, J(G)] = \Psi[\bar{G}, J(G)] + \sum_{j=1}^n w_j (G_{l_j} - \bar{G}_{l_j})^2$$

$$\Psi[G', J(G)] = \Psi[\bar{G}, J(G)] + \sum_{j=1}^n w_j (G'_{l_j} - \bar{G}_{l_j})^2 \quad (58)$$

and the resulting reduction in the target function is

$$\Psi[G, J(G)] - \Psi[G', J(G)] = [1 - (1 - c)^2] \sum_{j=1}^n w_j (G_{l_j} - \bar{G}_{l_j})^2. \quad (59)$$

Moving from G to G' changes the mean logarithmic intensities by

$$J_h(G') - J_h(G) = -\sum_{j=1}^n w_j (G'_{l_j} - G_{l_j}) \delta_{hh_j} / \sum_{j=1}^n w_j \delta_{hh_j}$$

$$= c[J_h(\bar{G}) - J_h(G)]. \quad (60)$$

Expansion of Ψ at point $G', J(G')$ yields

$$\Psi[G', J(G)] = \Psi[G', J(G')] + \sum_{j=1}^n w_j [J_{h_j}(G) - J_{h_j}(G')]^2. \quad (61)$$

Using the abbreviations

$$a = \sum_{j=1}^n w_j (\bar{G}_{l_j} - G_{l_j})^2, \quad b = \sum_{j=1}^n w_j [J_{h_j}(\bar{G}) - J_{h_j}(G)]^2, \quad (62)$$

the reduction in the target function at completion of one cycle is

$$\Psi[G, J(G)] - \Psi[G', J(G')] = 2ac - (a - b)c^2. \quad (63)$$

A finite value for the reduction requires $a > b$, which leads to an optimal step size $c = a/(a - b)$, so that the target function is reduced by $\Psi[G, J(G)] - \Psi[G', J(G')] = ac$.

For the gradient, we have

$$\frac{\partial\Psi[G, J(G)]}{\partial G_l} = \left(\frac{2}{c}\right)\sum_{j=1}^{n} w_j(G_{l_j} - G'_{l_j})\delta_{ll_j}, \quad \frac{\partial\Psi[G, J(G)]}{\partial J_h} = 0.$$
(64)

Since the target function is non-negative, the procedure generates a bounded monotonically decreasing sequence converging to the minimum of the target function at vanishing gradient. As shown above, this also implies convergence of the sequence of scaling factors.

Obviously, the solution thus obtained is not unique because the target function does not change its value if one adds an arbitrary constant to the logarithmic scaling factor G_l and subtracts an appropriate constant vector from J_h. This amounts to just changing the common scale of all images in the original problem, which is of no importance as we are only interested in the relative scale.

It may happen that several sets of images exist that are not connected by common measurements. In this case the target function could be thought of consisting of a sum of the same type of target functions, one for each unconnected subset of images. Now there is an arbitrary common factor for each subset. Apparently, the presence of arbitrary common factors for each subset of images does not prevent convergence of the cyclic solution procedure described here.

3.7.6. Post-refinement. As mentioned above, the diffraction parameters of each image are refined in the *IDXREF* and *INTEGRATE* steps to minimize deviations between the observed and the predicted locations of the strong spots and to minimize their angular distance from the Ewald sphere. The angular part of the target function takes care of the fact that reflections can be visible only if they are close to the Ewald sphere. In addition, the distribution of τ angles thus obtained provides an initial guess for the crystal mosaicity σ_M. However, the angular part of the initial refinement target cannot account for the fact that very strong reflections can still be observed even if they are farther away from the Ewald sphere than the weaker reflections. This leads to a systematic bias in the initial parameter refinement so that strong reflections will be predicted to be closer to the Ewald sphere than they really are.

These deficiencies can be overcome when all images have been processed and intensity estimates for the recorded reflections are included in the refinement, which explains why this approach has been dubbed 'post-refinement'. The original idea (Schutt & Winkler, 1977; Rossmann *et al.*, 1979; Harrison *et al.*, 1985; Rossmann, 1985) is extended here to handle still snapshots as well when fully recorded, measured reflections are not available for comparison and the notion of 'partiality' loses its meaning. Its role is assumed by the Ewald offset correction Q defined above that is applicable for rotation images as well as stills. The 'post-refinement' variant implemented here considers the possible unique reflection intensities as free parameters that are to be refined along with the diffraction parameters of each snapshot (Bolotovsky *et al.*, 1998).

The goal of the refinement procedure is the minimization of the target function

$$E = w_X \sum_{j=1}^{n} w_j(\Delta_X^j)^2 + w_Y \sum_{j=1}^{n} w_j(\Delta_Y^j)^2 + w_I \sum_{j=1}^{n}[(\hat{I}_j - C_j I_{h_j})/\hat{\sigma}_j]^2.$$
(65)

Here again, j enumerates the n recorded reflections from all snapshots, Δ_X^j, Δ_X^j the residuals between the calculated and observed spot centroids (see §3.5.2) and $\hat{I}_j$, $\hat{\sigma}_j$ the recorded raw intensity and its standard deviation. If a spot j is strong enough so that a centroid could be determined then $w_j = 1$, otherwise $w_j = 0$.

Each spot j is associated with a reciprocal-lattice point $\mathbf{p}_{0j}^*$ close to the Ewald sphere. For each observation a correction factor C_j can be computed from the diffraction parameters that relates the recorded intensity $\hat{I}_j$ to a unique reference intensity I_{h_j}, where h_j denote the unique indices of the reciprocal-lattice points $\mathbf{p}_{0j}^*$.

The target function E depends on private parameters for each snapshot and global parameters and constants.

(i) *Private parameters.*
$\mathbf{S}_0$, the incident-beam wavevector.
$\{\mathbf{b}_1^*, \mathbf{b}_2^*, \mathbf{b}_3^*\}$, the reciprocal cell basis vectors.
g, the scaling factor for intensities.
B, the isotropic temperature factor.
σ_M, the mosaicity.
detector parameters.
(ii) *Global parameters.*
I_h, the squared structure-factor amplitudes for the possible unique reflections h (up to some global constant irrelevant in this context).
(iii) *Global constants.*
$\mathbf{n}$ and p, the polarization plane normal and the degree of polarization.
μ, the fraction of intensity loss per millimetre in air.
δ, the thickness of the detector sensor.
κ, the fraction of intensity loss per millimetre in the sensor.
$\mathbf{m}_2$ and Δ_φ, the rotation axis and oscillation range ($\Delta_\varphi = 0$ for 'stills').

Each refinement round starts with the determination of the unique reflection intensities I_h, keeping the current parameter values constant. Minimization of the target function yields

$$I_h = \sum_{j=1}^{n} \delta_{hh_j}\hat{I}_j C_j/\hat{\sigma}_j^2 \Big/ \sum_{j=1}^{n} \delta_{hh_j} C_j^2/\hat{\sigma}_j^2.$$
(66)

The weights are then calculated as

$$w_X = 1/\sum_{j=1}^{n} w_j(\Delta_X^j)^2,$$

$$w_Y = 1/\sum_{j=1}^{n} w_j(\Delta_Y^j)^2,$$

$$w_I = 1/\sum_{j=1}^{n}[(\hat{I}_j - C_j I_{h_j})/\hat{\sigma}_j]^2$$
(67)

Table 1
Comparison of data processing with *XDS* and *nXDS*.

Resolution (Å)	15–2	15–6	6–4	4–2.1	2.1–2
XDS					
Reflections	356854	12414	31981	264002	48457
Multiplicity	7.8	7.7	7.8	7.8	7.7
$\langle I/\sigma(I)\rangle$	38.6	90.9	83.3	35.6	12.2
$CC_{1/2}$ (%)	100.0	100.0	100.0	100.0	99.2
CC_{ano} (%)	76	98	96	71	34
nXDS					
Reflections	3409453	56009	231474	2610196	511774
Multiplicity	74.6	37.9	57.4	76.8	82.7
$\langle I/\sigma(I)\rangle$	16.9	24.9	30.2	16.8	6.7
$CC_{1/2}$ (%)	99.9	99.8	99.9	99.8	97.2
CC_{ano} (%)	41	84	72	38	18

Table 2
Wall-clock times (s) for processing with *XDS* and *nXDS*.

Step	*XDS*	*nXDS*	Comments
XYCORR	1	1	
INIT	172	124	
COLSPOT	82	1864	I/O limited
POWDER	—	7	
IDXREF	3	1206	
INTEGRATE	2022	1704	
CORRECT	24	43	Using a reference
		808	Using selective breeding

Table 3
Consistent indexing by selective breeding.

Generation	Replaced misfits
1	10045
2	9740
3	6380
4	114
5	0

and are kept throughout the refinement round. The whole procedure is terminated upon convergence of the weights w_X, w_Y, w_I.

During a refinement round the diffraction parameters for each snapshot are corrected iteratively to minimize the target function until convergence is reached. The residuals are expanded to first order in the parameter changes so that E becomes a quadratic function of these changes. Minimization then leads to a system of normal equations whose solution is used for updating the parameters. The gradients of Δ_X^j, Δ_Y^j and C_j are computed from analytic expressions (not shown).

Fortunately, these calculations are independent for each snapshot and can be performed in parallel by a team of processors. Moreover, the memory requirements are almost negligible even when the refinement of detector parameters is included.

4. Example of data processing with *nXDS*

As an example to demonstrate the quality of data processed by *nXDS* in comparison to conventional data reduction by *XDS*, 20 000 consecutive rotation images were collected at 100 K from a crystal of a selenomethionine-labelled double mutant of the RNA-processing factor SCAF8 (Becker *et al.*, 2008). Each image covers a rotation range Δ_φ of 0.02° and is treated by *nXDS* as a snapshot taken from a randomly oriented crystal. The images were collected on beamline X10SA at the Swiss Light Source, Villigen, Switzerland at a wavelength of 0.9779 Å, slightly above the Se *K* edge. The images were recorded by a PILATUS 6M pixel detector (Dectris AG, Baden, Switzerland) located at 300 mm distance.

The crystal has $P4_3$ space-group symmetry, which is lower than the 422 lattice symmetry, implying a twofold indexing ambiguity.

The processing results are summarized in Table 1. The upper part refers to the evaluation of the images by *XDS* as conventional rotation data. The lower part shows the corresponding quantities as obtained from *nXDS*. Here, reflections were only included if their Ewald offset correction was larger than 0.7. A total of 356 854 reflections in the resolution range 15–2 Å were integrated by *XDS*, so that for each unique reflection almost eight symmetry-related reflections are available. Because contributions to each reflection are also recorded by adjacent images, it is not surprising that *nXDS* found almost ten times more reflections, which is consistent with the correspondingly higher multiplicity of observations.

After merging symmetry-related reflection intensities one might expect that the mean signal-to-noise ratio $\langle I/\sigma(I)\rangle$ would come out about the same regardless of whether the images were processed by *XDS* or *nXDS*. This is not the case: the mean signal-to-noise ratio is higher by a factor of almost three for the results from *XDS*. The lower accuracy of *nXDS* presumably results from two-dimensional instead of three-dimensional profile fitting and the lack of other corrections not carried out yet by this version of *nXDS*. The nearly perfect correlation factors $CC_{1/2}$ (Karplus & Diederichs, 2012) between intensities of symmetry-related reflections obtained from processing by both programs reflect the excellent quality of the data images. The presence of anomalous scatterers is clearly indicated in both processing results by the highly significant value for the anomalous correlation CC_{ano}. Finally, the reflection intensities obtained from both programs are in excellent agreement, showing a correlation coefficient of 98%.

Data processing was carried out by a 12-core machine with 16 GB memory running under Linux. Making use of the hyperthreading capability, up to 24 threads were employed for processing the images stored on a local disk. Elapsed wall-clock times for each step are listed in Table 2. *COLSPOT* uses a very fast spot-finding procedure but spends most of its time waiting for the next image to arrive. On average only three out of 24 threads were active. *COLSPOT* is a time-consuming step in *nXDS* because each image had to be analyzed for diffraction spots since the knowledge that the images comprise a rotation data set was not used. In contrast, *XDS* only requires spots from a small fraction of the images for recognizing the crystal lattice and for accurate refinement of the diffraction parameters. For the same reason, the *IDXREF* step in *nXDS* takes much longer than in *XDS*. In the *INTEGRATE* step *nXDS* is somewhat faster than *XDS* because of the reduced

Table 4
Phasing statistics.

	XDS	nXDS
SHELXD		
Correct solutions for 100 trials	100	96
PATFOM	21.41	14.38
CC$_{all}$ (%)	56.07	38.31
CC$_{weak}$ (%)	34.51	21.99
SHELXE		
CC$_{free}$ (%)	71.53	62.25
Heavy-atom peaks (map units σ)		
Strongest (σ)	44	34
Weakest (σ)	19	13
Average (σ)	36	27
Highest noise (σ)	10	8

overhead in control when only single images are involved. Compared with *XDS*, the *CORRECT* step of *nXDS* takes longer because of the much larger number of reflections. Furthermore, additional computations are required by the selective breeding procedure when no reference data set is available. According to Table 2, the breeding procedure required about 14 min to resolve the twofold indexing ambiguities for 20 000 snapshots and a total of 3.4 million reflections.

Details of the procedure are shown in Table 3 as the number of misfitting snapshots. In the first two generations both indexing alternatives are nearly randomly distributed. A small fluctuation towards one choice builds up in the third generation and quickly dominates the population. After five generations a homogeneous population is obtained.

Phasing of single anomalous diffraction data was performed for the *XDS* and *nXDS* processed data sets using the *SHELXC/D/E* program suite (Sheldrick, 2010). The results are summarized in Table 4. In brief, the marker-atom structure factors were estimated from pre-merged data using *SHELXC*. Subsequently, the selenium substructure was determined in a search of 100 trials for ten putative sites while applying a high-resolution cutoff at 2.5 Å. For both data sets, *SHELXD* identified 14 positions, of which seven showed significantly higher occupancy when compared with the less significant positions. This is in good agreement with the expected eight selenium positions (Becker *et al.*, 2008). Phases were further improved by density modification using *SHELXE*; eight sites were refined to significant occupancy and the phases obtained resulted in excellent electron-density maps with high pseudo-free correlation coefficients CC$_{free}$ (Table 4) and the correct enantiomorphic setting. Peak heights at the eight heavy-atom sites are highly significant and are well above the largest noise peak. Although phasing for both data sets was unambiguous, the data set processed with *nXDS* showed slightly lower correlation coefficients, Patterson figures of merit (PATFOM) and heavy-atom peak heights.

5. Conclusion

This study describes a new approach for processing a large set of snapshots from randomly oriented crystals that does not rely on the Monte Carlo method of integration. Instead, the concept of the Ewald offset correction factor was devised to overcome difficulties arising from the use of partiality for modelling reflection intensities recorded by snapshots. The new approach has been implemented in the program *nXDS* that has borrowed many ideas and routines from the rotation data-processing package *XDS* as well as from the powerful post-refinement technique that has been in widespread use for several decades.

The implemented Ewald offset correction relies on a Gaussian model for the rocking curve, assuming a sufficiently large crystal whose shape transform can be ignored. In this case the exact functional form of the curve is not critical for reflections sufficiently close to the Ewald sphere. In the test data set, a reflection was only included if its Ewald correction factor was larger than 0.7.

As shown for the test case with fine-sliced rotation images of excellent quality, *nXDS* delivers results almost approaching those obtained by *XDS* and is able to retrieve the anomalous signal from a selenomethionine-labelled protein crystal. The source of the lower accuracy of *nXDS* is not yet clear. It could result from two-dimensional instead of three-dimensional profile fitting and the omission of information from weak contributions to reflections further away from the Ewald sphere that are used only by *XDS*. In fact, a small improvement in overall data quality by 0.9% in $\langle I/\sigma(I)\rangle$ and 2.9% in CC$_{ano}$ was observed upon the inclusion of weaker contributions when the minimum required Ewald offset correction was lowered from 0.8 to 0.7.

So far, no 'real' FEL data have been processed by *nXDS*. These data typically vary for each snapshot in wavelength, bandwidth and crystal parameters. Although *nXDS* allows some of these to change for each snapshot, program modifications are likely to become necessary when the incident-beam bandwidth can no longer be substituted by a mean wavelength or if the shape transform of the crystals cannot be ignored. Presently, *nXDS* can only accept images that *XDS* can read. Work is in progress to adapt the package to also handle the detectors used at FEL beamlines and to make *nXDS* and its documentation available from the internet.

References

Barends, T. R. M., Foucar, L., Botha, S., Doak, B., Shoeman, R. L., Nass, K., Koglin, J. E., Williams, G. J., Boutet, S., Messerschmidt, M. & Schlichting, I. (2014). *Nature (London)*, **505**, 244–247.
Becker, R., Loll, B. & Meinhart, A. (2008). *J. Biol. Chem.* **283**, 22659–22669.
Bolotovsky, R., Steller, I. & Rossmann, M. G. (1998). *J. Appl. Cryst.* **31**, 708–717.
Brehm, W. & Diederichs, K. (2014). *Acta Cryst.* D**70**, 101–109.

research papers

Duisenberg, A. J. M. (1992). *J. Appl. Cryst.* **25**, 92–96.

Hamilton, W. C., Rollett, J. S. & Sparks, R. A. (1965). *Acta Cryst.* **18**, 129–130.

Harrison, S. C., Winkler, F. K., Schutt, C. E. & Durbin, R. M. (1985). *Methods Enzymol.* **114**, 211–237.

Hattne, J. *et al.* (2014). *Nature Methods*, **11**, 545–548

Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.

Kabsch, W. (1978). *Acta Cryst.* A**34**, 827–828.

Kabsch, W. (1988*a*). *J. Appl. Cryst.* **21**, 67–72.

Kabsch, W. (1988*b*). *J. Appl. Cryst.* **21**, 916–924.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.

Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.

Kahn, R., Fourme, R., Gadet, A., Janin, J., Dumas, C. & André, D. (1982). *J. Appl. Cryst.* **15**, 330–337.

Karplus, A. & Diederichs, K. (2012). *Science* **336**, 1030–1033.

Kirian, R. A., Wang, X., Weierstall, U., Schmidt, K. E., Spence, J., Hunter, M., Fromme, P., White, T., Chapman, H. N. & Holton, J. (2010). *Opt. Express*, **18**, 5713–5723.

Kirian, R. A., White, T. A., Holton, J. M., Chapman, H. N., Fromme, P., Barty, A., Lomb, L., Aquila, A., Maia, F. R. N. C., Martin, A. V., Fromme, R., Wang, X., Hunter, M. S., Schmidt, K. E. & Spence, J. C. H. (2011). *Acta Cryst.* A**67**, 131–140.

Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.

Powell, H. R., Johnson, O. & Leslie, A. G. W. (2013). *Acta Cryst.* D**69**, 1195–1203.

Rossmann, M. G. (1985). *Methods Enzymol.* **114**, 237–280.

Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.

Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. (2013). *Acta Cryst.* D**69**, 1274–1282.

Schutt, C. & Winkler, F. K. (1977). *The Rotation Method in Crystallography*, edited by U. W. Arndt & A. J. Wonacott, pp. 173–186. Amsterdam: North-Holland.

Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.

Steller, I., Bolotovsky, R. & Rossmann, M. G. (1997). *J. Appl. Cryst.* **30**, 1036–1040.

White, T. A., Barty, A., Stellato, F., Holton, J. M., Kirian, R. A., Zatsepin, N. A. & Chapman, H. N. (2013). *Acta Cryst.* D**69**, 1231–1240.

White, T. A., Kirian, R. A., Martin, A. V., Aquila, A., Nass, K., Barty, A. & Chapman, H. N. (2012). *J. Appl. Cryst.* **45**, 335–341.

Wirth, N. (1976). *Algorithms + Data Structures = Programs*, pp. 264–274. New York: Prentice Hall.

Zachariasen, W. H. (1945). *Theory of X-Ray Diffraction in Crystals.* New York: Dover.

Zhang, Z., Sauter, N. K., van den Bedem, H., Snell, G. & Deacon, A. M. (2006). *J. Appl. Cryst.* **39**, 112–119.