Computational diffraction reveals long-range strains, distortions and disorder in molecular dynamics simulations of irradiated single crystals

Alexandre Boulle, Alain Chartier, Aurélien Debelle, Xin Jin, Jean-Paul Crocombette

Computational diffraction reveals long-range strains, distortions and disorder in molecular dynamics simulations of irradiated single crystals

A. Boulle[1], A. Chartier[2], A. Debelle[3], X. Jin[1,3], J. -P. Crocombette[4]

[1]Institut de Recherche sur les Céramiques, CNRS UMR 7315, Limoges, France
[2]CEA, DEN, DPC, SCCME, Gif-Sur-Yvette, France.
[3]Laboratoire de Physique des 2 Infinis Irène Joliot-Curie, Université Paris-Saclay, CNRS/IN2P3, UMR 9012, Université Paris-Saclay, Orsay, France.
[4]CEA, DEN, SRMP, Gif-sur-Yvette, France.

Abstract

Atomic scale simulations, and in particular molecular dynamics (MD), are key assets to model the behavior of the structure of materials under the action of external stimuli, say temperature, strain or stress, irradiation, etc. Despite the widespread use of MD in condensed matter science, some basic material characteristics remain difficult to determine. This is for instance the case of the long-range strain tensor, and its root-mean-squared fluctuations, in disordered materials. In this work, we introduce computational diffraction as a fast and reliable structural characterization tool of atomic scale simulation cells in the case of irradiated single crystals. As compared to most direct-space methods, computational diffraction operates in the reciprocal-space and is therefore highly sensitive to long-range spatial correlations. With the example of irradiated UO_2 single crystals, it is demonstrated that the normal strains, shear strains and rotations, as well as their root-mean-squared fluctuations (microstrain) and the atomic disorder, are straightforwardly and unambiguously determined. The methodology presented has been developed with efficiency in mind, in order to be able to provide simple and reliable characterizations operating either in real time, in parallel with other analysis tools, or operating on very large data sets.

1. Introduction

An issue common to many materials science studies is the understanding of the evolution of the nanoscale structure of materials under the action of external stimuli, say temperature, stress or strain, laser or ion irradiation, etc. In this context, molecular dynamics (MD) simulations have emerged as the method of choice to tackle that type of issues (Nordlund & Djurabekova, 2014; Krasheninnikov & Nordlund, 2010). Efficient algorithms have been developed to characterize the local structure of materials in multi-million atoms simulation cells, such as the Voronoi analysis, bond angle analysis, common neighbor analysis, dislocation extraction algorithm, crystallography cell deformation method, or the Voronoi cell deformation (VCD) method (Stukowski, 2012; Leonardi *et al.,* 2012) . These methods are particularly efficient to derive structural parameters such as the local strain, the atoms coordination, the defect densities, etc. Although the process of analyzing such large simulation cells can in principle be automated, thereby enabling so-called *in situ* computational microscopy (Zepeda-Ruiz *et al.,* 2017), it often remains a semi-automated process requiring manual fine tuning by expert scientists. This human-centered workflow is a time-consuming process that, *de facto,* constitutes a bottleneck when several hundreds, or thousands, of simulation data sets can be generated in the course of a few hours. Besides, structural parameters of primary importance, in particular the long-range strain tensor, are not straightforward to determine.

Over the past decades, a large number of studies have been dedicated to the determination of strain from atomic scale simulations (Zimmerman *et al.,* 2009; Mott *et al.,* 1992; Gullett *et al.,* 2008; Stukowski & Arsenlis, 2012; Stukowski *et al.,* 2009; Zhang *et al.,* 2015; Xiong *et al.,* 2019; Leonardi *et al.,* 2012). In these studies the local deformation tensor is obtained from the evolution of the coordinates of the atoms constituting a coordination shell between any two time steps of a MD trajectory ("kinematical" approach). Most approaches become problematic, with the noticeable exception of the VCD method (Leonardi *et al.,* 2012), in highly disordered regions of the cell where the coordination shells are incomplete, like in the close vicinity of surfaces, grain boundaries or dislocation cores. This is especially true for nanostructured materials or heavily damaged crystals, as those encountered during ion irradiation, where the level of disorder can bring materials into an amorphous state (Boulle & Debelle, 2016). Moreover, these methods inherently neglect long-range correlations in the atomic displacements, *i.e.* the displacement field around an atom in a given coordination shell is assumed to be independent from the displacement field around the next nearest neighbor, so that the strain averaged over the whole MD cell generally differs from the result that would have been obtained from a diffraction experiment (Xiong *et al.,* 2019).

A straightforward solution to this issue is to compute the "lattice strain" derived by the examination of the peak positions in a virtual X-ray diffraction (XRD) measurement, *i.e.* in the reciprocal space. A virtual XRD pattern can be easily generated using the Debye scattering equation (DSE) (Debye, 1915). Coupling the DSE with MD simulations is of common practice for the determination of the structure of nanostructured materials (Derlet *et al.*, 2005) (for a comprehensive review, see also (Gelisio & Scardi, 2016) and references therein). Conversely, the DSE remains only marginally used for the determination of strain and disorder in damaged materials, like those obtained by laser (Lin & Zhigilei, 2006) or ion (Chartier *et al.*, 2018; Soulié *et al.*, 2016; Debelle *et al.*, 2014) irradiation, for instance.

One important characteristic of the DSE is the powder assumption, which implies the same particle is observed with equal probability through any orientation in space. The resulting intensity profile corresponds to the spherical integration of the reciprocal space. All the structural and microstructural information about the system is contained in a single XRD pattern. The peak overlap inherent in that type of signals, especially for nanosized simulation cells might hinder an accurate evaluation of the peak positions and intensities and consequently an accurate determination of the strain. This issue is usually dealt with using whole powder-pattern fitting methods (such as the Rietveld method (Rietveld, 1969)) or more sophisticated whole powder-pattern modeling methods (WPPM) (Scardi & Leoni, 2002). It must be mentioned that advanced methods, such as WPPM, allow for an in-depth characterization of the microstructure, including not only the lattice strain, but also crystallite shapes and size distribution, dislocation and stacking fault densities, surface-induced disorder, etc. (Scardi *et al.*, 2015; Leonardi & Scardi, 2015; Rebuffi *et al.*, 2016; Leonardi & Bish, 2017). However, those fitting/modeling methods are relatively time consuming and require fine-tuning of the fitting parameters in order to avoid the occurrence of unphysical results. In this sense they are not very well adapted to deal with large data sets. Finally, all quantities derived from such an analysis are orientation-averaged, which prohibits the detection of spatial anisotropies and all information regarding rigid rotations are lost in the process.

In the present work, we propose to use single-crystal-like XRD patterns, or computational diffraction, instead of powder XRD patterns. By doing so we are able to isolate single XRD peaks which greatly facilitates the analysis. Since only an handful of peaks are considered, there is obviously a loss of information, especially regarding the detailed defect structure of the MD cells. However, for the sake of determining the long-range strain and disorder, we show below that this approach is largely sufficient, especially considering the fact, that the defect structure is more easily characterized using a direct analysis of the MD cells. Moreover, this makes the analysis more robust

and compatible with automated analysis workflows on very large data sets. The results presented below, corresponding to the analysis of several thousands of MD cells, have been obtained automatically without any manual intervention. It must be mentioned that the proposed approach has been designed with MD simulations from single crystals in mind. Blindly applying this computational diffraction approach to polycrystalline materials would probably yield erroneous results, mainly because in such cases large regions of the reciprocal space have to be investigated.
Finally, we prove that both the reciprocal-space and the real-space (that is, directly from the atomic coordinates) long-range strain determinations yield the same results if long-range correlations are correctly taken into account. The validity of our computational diffraction approach is tested on UO_2 MD cells in which we simulated ion-irradiation induced disorder and for which we have a good understanding of the defect structure and evolution (Chartier *et al.*, 2016; Jin *et al.*, 2020).

2. Theoretical background

Reciprocal space mapping, is a widespread experimental technique to analyze strain and various types of defects in epitaxial films or single crystals (Holý *et al.*, 1999). Reciprocal space maps (RSMs) correspond to two-dimensional sections of the reciprocal space. They are usually represented in a two-dimensional (Q_x, Q_z) plane, where Q_x and Q_z are the components of the scattering vector **Q**, respectively parallel and normal to the crystal surface, and they correspond to a section of the reciprocal space along this particular plane. Any atomistic model can be transformed into a RSM via the following equation (Channagiri *et al.*, 2015):

$$I(Q_x, Q_z) = \sum_{\Delta Q_y} \left| \sum_{j=1}^{N} f_j(\mathbf{Q}) \exp\left(i\mathbf{Q} \cdot \mathbf{r'_j}\right) \right|^2 \quad (1)$$

where $\mathbf{r}'_j$, f_j and N are the coordinate vector of the j^{th} atom, the atomic scattering factor of j^{th} atom in the simulation cell, and the total number of atoms in the cell, respectively. $Q_{x,y,z}$ are the three components of the scattering vector **Q**:

$$\mathbf{Q} = 2\pi \left(H\mathbf{a}^* + K\mathbf{b}^* + L\mathbf{c}^*\right) \quad (2)$$

where $\mathbf{a}^*$,$\mathbf{b}^*$,$\mathbf{c}^*$ are the reciprocal space basis vectors and H, K, L are the so-called continuous Miller indices that determine the magnitude and direction of the scattering vector. Integer values of H, K and L correspond to maxima of the intensity distribution, *i.e.* Bragg peaks (provided that the said values obey the selection rules dictated by the symmetry of the crystal). Selecting different *HKL* values allows one to probe different directions of the simulation cell, *via* the scalar product $\mathbf{Q}\cdot\mathbf{r}'_j$ in Eq. (1). Finally ΔQ_y is a narrow integration range perpendicular to the (Q_x, Q_z) plane aimed at smoothing the computed RSM (see Appendix A).

The structural information that can be accessed *via* the RSMs is summarized in the following equation (Holý *et al.*, 1999):

$$I(\mathbf{Q}) = \sum_{j=1}^{N}\sum_{k=1}^{N} f_j(\mathbf{Q}) f_k^*(\mathbf{Q}) \times \left\langle \exp\left\{ i\mathbf{Q}\left[\delta\mathbf{u}(\mathbf{r}_j) - \delta\mathbf{u}(\mathbf{r}_k)\right]\right\}\right\rangle \times \exp(i\left\{\mathbf{Q}^{\mathbf{T}}\left[\mathbf{I} + \nabla\mathbf{u}(\mathbf{r}_j)\right]\right\}\cdot\left[\mathbf{r}_j - \mathbf{r}_k\right]) \quad (3)$$

where **r** are the atomic coordinates in the unperturbed lattice, and **u(r)** are the deviations from the perfect lattice, *i.e.* the atomic displacement vectors, so that $\mathbf{r}' = \mathbf{r} + \mathbf{u}(\mathbf{r})$. The displacement **u(r)** can be further separated into two components, $\mathbf{u}(\mathbf{r}) = \nabla\mathbf{u}(\mathbf{r})\cdot\mathbf{r} + \delta\mathbf{u}(\mathbf{r})$, where $\nabla\mathbf{u}(\mathbf{r})\cdot\mathbf{r}$ describes the average response of the lattice to the presence of crystal defects. This term is a Taylor series expansion of the non-random displacement **u**, limited to the first order, which implicitly implies that the component of $\nabla\mathbf{u}(\mathbf{r})$ are much smaller than 1 (small strain approximation). $\nabla\mathbf{u}$ is the Jacobian of the displacement (the displacement gradient tensor, or distortion tensor), with components $e_{ij} = \partial u_i/\partial j$. The diagonal components of $\nabla\mathbf{u}$ correspond to the normal strains. The off-diagonal elements of the distortion tensor include both shear strain and rigid rotations, both being important parameters to derive from the MD cells[1]. The second term, $\delta\mathbf{u}(\mathbf{r})$, corresponds to random local deviations around the average displacement.

The first exponential on the right-hand side of Eq. (3) contains the above-mentioned random displacement term and gives rise to two distinct effects. The first effect is an attenuation of the coherent scattering quantified by the so-called static Debye-Waller (DW) factor (Warren, 1969; Krivoglaz, 1969). For a perfect crystal, $\delta\mathbf{u}(\mathbf{r}) = 0$, so that DW = 1 and the coherent intensity has its maximum value. On the contrary, for highly disordered crystals, or amorphized materials, $\delta\mathbf{u}(\mathbf{r}) >> 0$, so that DW $\rightarrow$ 0 and the coherent intensity vanishes. The intensity subtracted from the Bragg scattering is redistributed in the background with a reciprocal space distribution depending on the degree of correlation of the disorder (*via* the $\delta\mathbf{u}(\mathbf{r}_j) - \delta\mathbf{u}(\mathbf{r}_k)$ term): this second effect is referred to as diffuse scattering. It is recalled in Appendix B that, in the case of correlated displacements, diffuse scattering manifests itself as peak broadening which can be connected to the rms distortion tensor, *i.e.* the *microstrain* tensor.

The second exponential contains the effect of homogeneous (long-range) strain. The diffraction condition implies that the argument of this exponential be an integer multiple of 2π. If, in a strain-free crystal, diffraction occurs at $\mathbf{Q}_0$, the presence of homogeneous strain, *via* $(\mathbf{I} + \nabla\mathbf{u})$, shifts the Bragg peak from its strain-free position $\mathbf{Q}_0$, to a modified coordinate $\mathbf{Q}_0^{\mathrm{T}}\cdot(\mathbf{I} + \nabla\mathbf{u})^{-1}$, which, in the

1 The usual strain tensor is obtained from the distortion tensor *via* $(\nabla\mathbf{u} + \nabla\mathbf{u}^{\mathrm{T}})$ / 2. The off-diagonal components only contain shear strains.

limit of small strain reduces to $\mathbf{Q}_0^{\mathrm{T}}\cdot(\mathbf{I} - \nabla\mathbf{u})$. Depending on the direction of $\mathbf{Q}_0$, different components of $\nabla\mathbf{u}$ can be selected. For instance, a $\mathbf{Q}_0 = (0, 0, Q_z)^{\mathrm{T}}$ vector is transformed into $(-e_{zx}\,Q_z, -e_{zy}\,Q_z, (1 - e_{zz})Q_z)^{\mathrm{T}}$, where e_{ij} are the components of $\nabla\mathbf{u}$. Measuring the coordinates of the Bragg peak for different $\mathbf{Q}_0$ values, hence, in principle, allows to retrieve the complete distortion tensor.

3. Computational details

3.1. Molecular dynamics

MD simulations were performed using the Frenkel pairs accumulation (FPA) methodology (Chartier *et al.*, 2005; Crocombette *et al.*, 2006) in order to mimic ion-irradiation induced ballistic damages. This methodology avoids the calculation of complete displacement cascades and their accumulation by directly creating their final states, *i.e.*, point defects. It has been proven very efficient to simulate irradiation damages in different oxides (Chartier *et al.*, 2009; Catillon & Chartier, 2014) and graphite (Chartier *et al.*, 2018). In MD simulations, only uranium Frenkel pairs were created, which imply to measure the introduced disorder in displacement per uranium, dpU (Jin *et al.*, 2020). MD calculations were performed in a 26 × 26 × 26 nm^3 UO_2 cell (768 000 atoms, with periodic boundary conditions in the 3 dimensions) using a Morelon empirical potential (Morelon *et al.*, 2003) which exhibits a relevant responses to irradiations (Devanathan *et al.*, 2010). FPA was performed in the NPT ensemble at 0 pressure and at 300 K using a modified version of the large-scale atomic/molecular massively parallel simulator (LAMMPS) code (Plimpton, 1995). 800 uranium Frenkel pairs (Crocombette & Chartier, 2007) were created every 2 ps. We used variable time steps ranging from 0.1 up to 1 fs in order to handle correctly local increase of atoms velocities due to interstitials/vacancies abruptly created. In addition, these defects produce local increase of temperature and pressure that are handled here by Berendsen thermostat / barostat. Hence, target pressure and temperature can be reached back before each new FP introduction. During the simulation, 898 snapshots of the MD cell, covering a 0 – 7.85 dpU range, were saved for further processing.

In addition, we created cells with rotated crystallites contained 768000 atoms with the perfect fluorite structure. Each cell was divided into 8 identical cubic sub-cells, half of which were rotated by 2° along an axis defined by a vector $\mathbf{r}_{x,y,z}$, where the subscript indicate its orientation with respect to the cell (see also Fig. 1). We considered reflections with *HKL* = 002 and 004, and a rotation axis directed along the **x**, **y** and **z** direction of the cell.

3.2. Computation and analysis of the reciprocal space maps

RSMs were computed using Eq. (1). For a typical RSM, as those represented in Fig. 1 for instance, the direct evaluation of Eq. (1) using a naive implementation requires more than one hour to reach completion which is incompatible with the large number of MD cells to be analyzed. Equation (1) was therefore evaluated on a graphics processing unit (GPU) using the Python programming language together with the NumPy (van der Walt *et al.*, 2011) and the PyNX (Favre-Nicolin *et al.*, 2011) libraries. With the hardware used in this work (a Nvidia Quadro P5000, released in 2016, with a theoretical 64bits peak performance of 277 GFLOPS) the computation time drops to ~20 seconds for a single RSM (a more recent hardware is expected to provide dramatically improved performances). Because of the large number of snapshots to be analyzed, the data processing (including peak finding, peak fitting, etc.) is entirely automated using Python scripting as described below.

RSMs for *HKL* = 002 and 004 have been computed for 6 different orientations of the MD cell, that is with the [100], [010], [001] directions successively set parallel to **Q** vector and, for each direction, two 90° spaced azimuthal orientations have been considered (see figure S1, Supplementary Material). Doing so, for each *HKL*, all 6 off-diagonal components of the distortion tensor can be determined, and the 3 diagonal components (normal strains) are determined twice.

Each computed RSM was processed as follows:

- A peak finding algorithm detects all the maxima in the RSM and the corresponding (*H*, *L*) coordinates are saved. The peaks are indicated as circles in Fig. 2. In order to avoid the detection of minor peaks or interference fringes, a detection threshold of 30% of the maximum intensity has been defined. There is no scientific justification behind this value; it has been empirically determined so as to be low enough to record all high intensity peaks, and high enough in order to reject minor peaks. Decreasing this value does not change the conclusions but complicates the analysis with the occurrence of several low intensity peaks. On the other hand, increasing this value too much may lead to erroneous conclusions regarding the development of strains, since high intensity peaks might be missed. These coordinates are then used to compute the diagonal and off-diagonal components of the strain tensor, *i.e.*

$$e_{ii} = (L_0 - L)/L_0 \text{ and } e_{ij} = (H_0 - H)/L_0 \quad (4)$$

  where *i, j = x, y, z* are determined by the orientation of the MD cell. Where the subscript 0 indicates the virgin cell (0 dpU).

- For the maximum intensity peak, the integrated (coherent) intensity is computed by integrating the intensity along the row $H = H_{max}$, where H_{max} is the H coordinate of the most intense peak. For HKL = 002 and 004, and for weak disorder, $H_{max} = 0$, whereas deviations from $H_{max} = 0$ indicate rotations, hereinafter also denoted as misorientations, or shear strain, as discussed above. The integrated intensities are then used to compute the DW factor, which, correspond to the lowering of the coherently scattered amplitude relative to the virgin crystal; that is:

$$DW = \sqrt{\frac{I^{H_{max}}}{I_0^{H_{max}}}} \tag{5}$$

It is important to note that the atomic configurations from which the RSMs are generated are not time-averaged, *i.e.* they contain both thermal (dynamic, *d*) and irradiation induced (static, *s*) disorder, $\langle \delta u^2 \rangle = \langle \delta u_d^2 \rangle + \langle \delta u_s^2 \rangle$ (Inagaki *et al.*, 1983). Using the intensity of the unirradiated cell as a normalization factor allows to solely consider the irradiation-induced (static) disorder. Note that the thermal disorder turns out to be negligible as compared to the damage induced by irradiation.

- The envelope of the intensity distribution was fitted with a bivariate asymmetrical Gaussian distribution (dotted contour lines in Fig. 2):

$$g^{(q,r)}(H,L) = I_{max} \exp\left\{-4\ln 2\left[\frac{(H-H_{max})^2}{w_H^{(q,r)2}} + \frac{(L-L_{max})^2}{w_L^{(q,r)2}}\right]\right\} \tag{6}$$

where $w_{H,L}$ are the full-widths at half-maximum (FWHM) in the H and L directions and the superscripts q and r designate the lower and upper half of the width, that is:

$$w_{H,L}^{(q,r)} = \begin{cases} w_{H,L}^{(q)} & \text{if } H \le H_{max} \text{ or } L \le L_{max} \\ w_{H,L}^{(r)} & \text{otherwise.} \end{cases} \tag{7}$$

The FWHM in the H and L directions is then obtained from:

$$w_{H,L} = (w_{H,L}^{(q)} + w_{H,L}^{(r)})/2 \tag{8}$$

Notice that we here used a definition of the Gaussian distribution using FWHMs instead of standard deviations. Both are related *via* a $2\sqrt{2\ln 2}$ factor (see below). The corresponding FWHMs are used to determine the microstrains.

The FWHMs determined from the fitting procedure contain both the contribution of defects and the contribution of the finite cell size. In order to extract the sole contribution of defects broadening, the cell size effect must be deconvoluted from the total FWHM which, in the case of Gaussian function, is straightforwardly performed:

$$w_{defects} = \sqrt{w^2 - w_{size}^2} \tag{9}$$

where w_{size} is given by the FWHM at low disorder levels. In the situation where defects only induce microstrain (as is the case in the present study) the diagonal and off-diagonal components of microstrain tensor are :

$$\epsilon_{i,j} = w_{H,defects}/(2\sqrt{2\ln 2}L) \tag{10}$$

$$\epsilon_{i,i} = w_{L,defects}/(2\sqrt{2\ln 2}L) \tag{11}$$

The whole analysis is easily distributed over several computing units. With the CPU used in this work (dual Intel Xeon processor with 20 cores each), the procedure requires less than 5 seconds for the 898 RSMS.

Before proceeding to the next section, it is important to notice that, in the general case, a Gaussian function is known to be a rather poor description of XRD peak profile shapes, which often display heavy tails. Whereas the peak position (hence the derived strain) is robust against different peak shapes, parameters such as the peak integrated intensity and width are heavily dependent on the selected peak shape. In the present case, it turned out that the Gaussian approximation led satisfactory results and the discrepancies are critically discussed. Therefore, when discussing the DW factor and the microstrain tensor, we shall restrict ourselves to discuss their relative evolution and we do not attempt to extract absolute values from the analysis.

4. Results and discussion

4.1. Rotated crystalline domains

Before addressing the case of disordered UO_2 cells, we shall first consider a simpler example consisting of UO_2 cells containing crystalline domains half of which are rotated by 2° around the *x*, *y* or *z* axes. This is shown in Fig. 1g-i, respectively. For simplicity, a simple cubic lattice is depicted. We considered 002 and 004 reflections, for which the **Q** vector of the Bragg reflection can be written $(0, 0, Q_z)^T$. Under the action of each of the three rotations, it gets transformed into $[0, -Q_z \times \sin\omega, Q_z \times (1-\cos\omega)]^T$, $[Q_z \times \sin\omega, 0, Q_z \times (1-\cos\omega)]^T$ and $(0, 0, Q_z)^T$, respectively, ω being the misorientation angle. The corresponding RSMs are displayed in Fig. 1a-f. The RSMs are plotted as a function of $\Delta H = H - H_0$ and $\Delta L = L - L_0$, which are the deviations from the unperturbed crystal indices, H_0 and L_0.

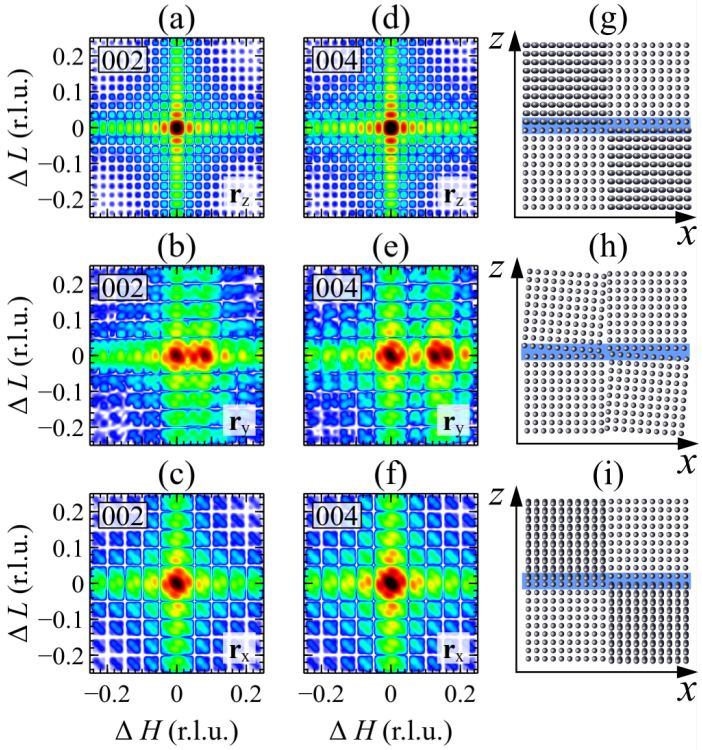

Fig. 1. RSMs of UO_2 cells divided into 8 rigid sub-cells where half of the sub-cells are rotated by 2°, with a rotation axis directed along the z (a,d), y (b,e) and x (c,f) direction. (a,b,c) : 002 reflection. (d,e,f): 004 reflection. The intensity is plotted on a logarithmic scale using a usual red-yellow-green-blue color scale. The axes are graduated in reciprocal lattice units. The corresponding schematic atomic structures are given in (g,h,i).

Let us consider the different cases in Fig. 1:

- Figures 1a,d correspond to the case $\mathbf{r}_z$ (rotation about the *z* axis). As indicated from the transformed **Q** vectors above, no modification is expected in this case and both RSMs correspond to the scattering from a perfect cubic crystal (a 2D Laue function). The interference fringes are related to the finite size of the MD cell and their period is given by the inverse of the number of UO_2 unit-cells in each direction, $1 / N_{uc} = 1 / 40 = 0.025$.
- Figures 1b,e correspond to the case $\mathbf{r}_y$. Since $\omega = 2°$ (0.035 rad), the transformed **Q** vector is close to $(- Q_z \times \omega, 0, Q_z)^T$, *i.e.* the crystallites affected by the rotation exhibit a Bragg peak shifted along the *H* direction, whereas the *L* coordinate remains unchanged. The result is a splitting of the Bragg peak, the magnitude of which being directly related to the rotation angle. The split for the 002 and 004 reflections is 0.07 and 0.14, respectively, consistently with a 2° misorientation (*i.e.* 0.035×2 and 0.035×4). Another visible feature is a broadening of the peak and of the interference fringes, as compared to Fig. 1a,b. This is due to the fact

that the sub-cells diffract incoherently, each containing N/8 atoms, hence the corresponding peak and fringes broadening (which is quantified by Scherrer's equation in the field of powder diffraction (Warren, 1969)). It can further be noticed that, contrarily to microstrain or misorientation effects, this "finite crystallite size" effect is independent on the magnitude of **Q**. This feature (Q-dependent *vs.* Q-independent effect) constitutes the basis of the size/microstrain separation in powder XRD line profile analysis methods (e.g. Williamson-Hall and Warren-Averbach methods (Warren, 1969)).

- Figures 1c,f correspond to the case $\mathbf{r}_x$. As for the previous case, the crystallites affected by the rotation have a **Q** vector close to $(0, -Q_z \times \omega, Q_z)^T$. However, contrarily to the previous case, because the RSMs result from an integration along the y-direction, they are not sensitive to the change in the Q_y component, therefore no splitting is observed. Nonetheless, an indirect indication of the presence of this rotation is the fact that, as for $\mathbf{r}_y$, the fringes are broader due to the smaller coherent domain size.

An alternate way of picturing the effect of crystallite misorientations on the RSMs is to observe the crystal structure projected in the (x, z) planes. Since the RSMs are computed with a $(0, 0, Q_z)^T$ vector, according to Eq. (3) we are solely sensitive to the z component of the atomic displacements:

- Figure 1g shows that when the rotation axis is parallel to z, no misorientation nor any discontinuity in the atomic z coordinates are observed (see colored regions), hence the RSMs exhibiting a characteristic perfect 2D Laue intensity distribution (Fig. 1a,d).
- Figure 1h shows the effect of $\mathbf{r}_y$ (normal to the plane of the figure). Here, both finite size and misorientation effects are clearly observed in the projected structure and in the RSMs (fig. 1b,e).
- Finally, Fig. 1i shows that, in the case of a $\mathbf{r}_x$ rotation axis (*i.e.* parallel to x), the atomic coordinates along the z axis exhibit abrupt variations both along the x and z axis (colored region), hence the observed finite size effect in the RSMs. On the contrary, no misorientation is observed, in agreement with Fig. 1c,f.

This simple examples reveals that a RSM computed for a single $00L$ reflection allows to straightforwardly derive the e_{zx} component of the distortion tensor using Eq. (4). The other e_{ij} components either have indirect effects (peak broadening and increased fringe spacing) or no effect at all. Although it is not explicitly illustrated in the example above, the e_{zz} component is readily

obtained by measuring the displacement of the Bragg peak along the Q_z direction (as evident from Eq. (4)).

Therefore, using 6 independent *HKL* reflections or, equivalently, 6 orientations of the cell allows to retrieve the full distortion tensor. In actual MD cells, the different components of the strain tensor occur simultaneously, with randomly oriented shear/rotations axis and with random distributions of normal strain and shear/rotation angles. The next sections illustrates how these different effects can be disentangled so as to determine the distortion tensor tensor (including the normal strains), the microstrain tensor, the disorder, etc. It should also be stressed that using only 6 *HKL* reflections does not allow one to get a complete picture of the defect structure in the MD cells, since defects may affect different reflections in different ways depending on their anisotropy, orientations, etc. We here only consider the strains and their evolution, whereas the detailed defect structure should be determined with a real-space analysis.

4.2. Defective UO_2 MD cells

It is now well documented that the disordering kinetics of UO_2 exposed to ion-irradiation can be divided in different stages corresponding to a different dominant defects in the MD cell (Chartier *et al.*, 2016; Jin *et al.*, 2020), namely Frenkel pairs (stage 1, from 0 to 0.09 dpU), Frank loops (stage 2, from 0.09 to 0.3 dpU), perfect loops (stage 3, from 0.3 to ~1.5 dpU) and dislocation lines (stage 4, above ~1.5 dpU), respectively. In these studies, it was also demonstrated that the different types of defects can be correlated with the evolution of both the average elongation strain and the average disorder. In the following, we apply the computational diffraction approach to these UO_2 cells and proceed to the determination of the complete distortion and microstrain tensors.

We computed RSMs for *HKL* = 200, 020, 002, 400, 040 and 004, thereby setting the **Q** vector parallel to the **x**, **y** and **z** directions of the MD cell. Moreover, for each direction, two 90° spaced orientations around **Q** have been considered. RSMs of the 002 reflection at 4 selected disorder levels (0, 0.1, 0.3 and 3 dpU) are displayed in Fig. 2, together with the corresponding MD cells. It can be observed that, whereas it is difficult to get clear insights regarding strain and disorder affecting the cell from a visual inspection of the MD cells, the RSMs shows features that are readily interpreted. The undamaged cell (Fig. 2a, stage 1) corresponds to a 2D Laue function characteristic of a perfect crystal (Fig. 2e). At 0.1 dpU (stage 2), Fig. 2b, we observe a displacement of the Bragg peak towards lower *L* values (Fig. 2f) which, according to the previous section, is indicative of the development of tensile strain. At this damage level, it can be observed that the intensity spreads out of the specular direction characterized by $H = 0$, a feature that is characteristic of diffuse scattering

associated with atomic disorder. However, the coherent peak remains perfectly visible and aligned along the $H = 0$ row, which indicates that the level of disorder is low, and there is no evidence of misorientation or shear strain. At 0.3 dpU (stage 3), the coherent peak disappears and is replaced with a broad diffuse scattering peak, within which 3 peaks can be detected; this corresponds to the situation where the highest levels of disorder and strain are present. Finally, at 3 dpU (stage 4), the diffuse scattering coalesces back around the $H = 0$ coordinate, only one peak remains and interference fringes are reformed, which is indicative of a decrease in the level of disorder. The vertical shift of the reflection is also reduced, which demonstrates that the level of strain decreases as well. The Bragg peak is nonetheless broader than at 0 and 0.1 dpU and exhibits a fine structure, which may indicate the formation of rotated or sheared unit cells. In the following, we analyze the peak position, intensity and width in terms of strain and disorder.

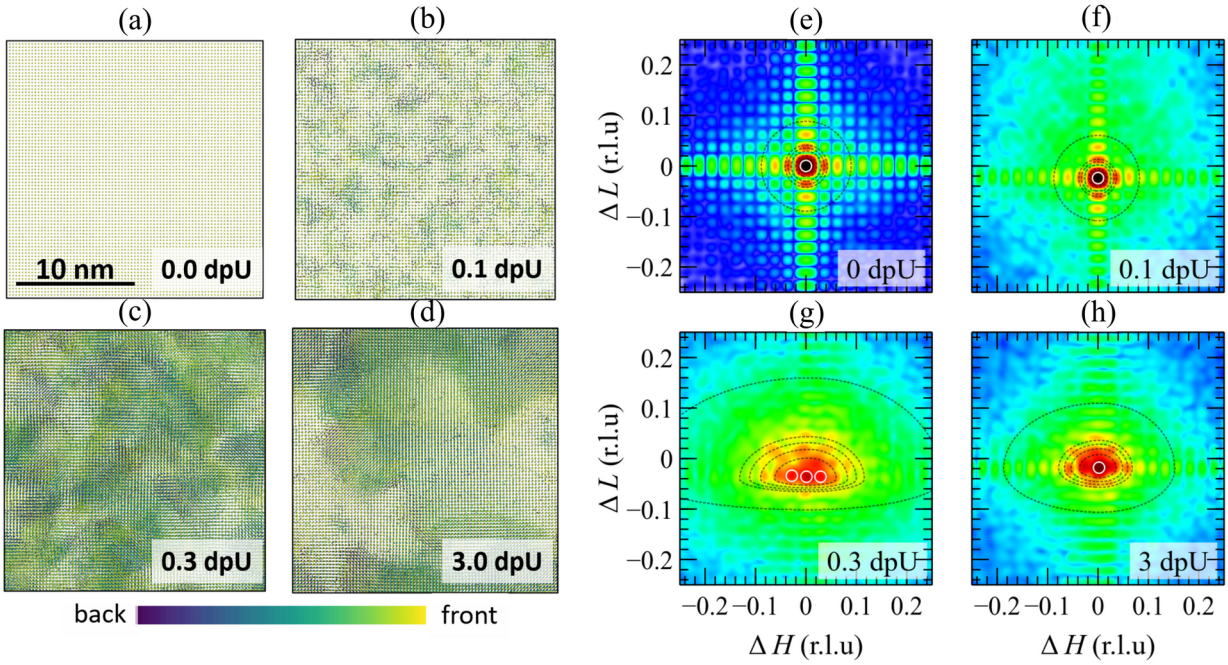

Fig. 2. (a-d) Snapshots of an UO_2 MD cell at different levels of induced disorder (expressed in displacement per uranium, dpU). The selected dpU levels correspond to boundaries between different regions of the damage build-up kinetic. For visualization purposes, only U atoms are displayed and the thickness normal to the figure is 5 nm. (e-h) RSMs of the 002 reflection of UO_2 MD cells for the same disorder levels The intensity is plotted on a logarithmic scale using a red-yellow-green-blue color scale. Axes are graduated in reciprocal lattice units. White circles indicate the location of the peak maxima, and dotted lines are iso-contour lines of the 2D fit with asymmetric Gaussian functions.

Normal strains and disorder

The normal strains (*i.e.* the diagonal elements of the distortion tensor) and the DW factor for the 002 RSMs from the 6 different cell orientations are given in Fig. 3 for different dpU levels. Let us first consider a single orientation, say, Fig. 3a. The circles represent the level of strain, deduced from the vertical peak coordinates. The symbol color indicates the intensity ratio of the peak from which the strain is determined relatively to the virgin crystal, using a usual red-yellow-green-blue color scale: a high intensity (red) corresponds to a large domain with low disorder, whereas a low intensity (blue) corresponds to disordered regions. The red line is the DW factor which, by definition, is the disorder-induced intensity lowering and is therefore completely correlated with the intensity ratio given by the circles color. The background colors represent different steps of the UO_2 disordering kinetics, corresponding to the different defect stage mentioned above.

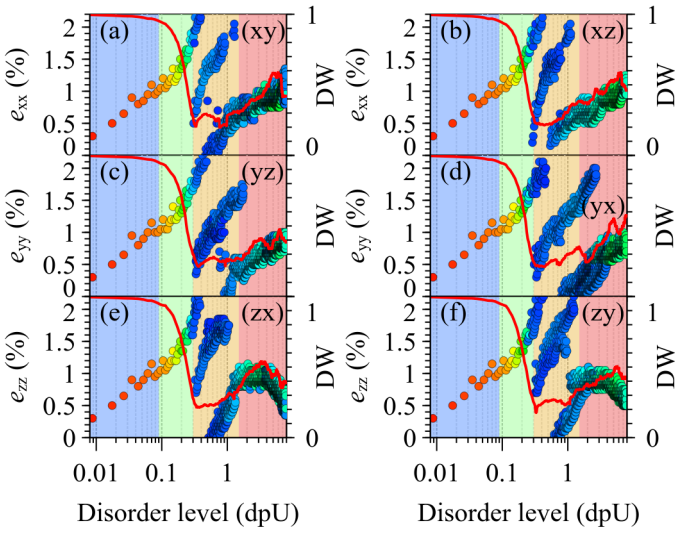

Fig. 3. Evolution of the strain (filled circles, left axis) and Debye-Waller factor (red line, right axis) with increasing dpU. The symbol color is proportional to the intensity ratio relatively to the virgin crystal. The different panels correspond to different orientations of the MD cell. Left column: azimuthal rotation of the MD cell is 0°. Rigth column: azimuthal rotation is 90° (see Fig. S1). (a,b): the x axis is set parallel to **Q**. *(c,d): the y axis is set parallel to* **Q**. *(e,f): the z axis is set parallel to* **Q**. *The intensity is integrated along z (a,d), y (b, e) and x (c, f). The background colors correspond to the different stages discussed in the text.*

Stage 1 (0 to 0.09 dpU, light blue) is characterized by a steep increase of the strain and a low level

of disorder. Stage 2 (0.09 to 0.3 dpU, light green) shows a continuous increase of the strain[2] and a considerable increase in disorder, as attested by the drop of the DW factor. In stage 3 (0.3 to ~1.5 dpU, light orange), the DW factor is plateauing whereas the strain exhibits a very perturbed behavior: for a given disorder level, several domains (with different levels of strain) are observed, giving rise to Bragg peaks splitted from the main peak (as illustrated in the previous section). In this region, the strain decreases from ~2% at 0.3 dpU to ~0.5% at 1.5 dpU. However, this decrease takes place abruptly at discrete dpU values. Although the strain drops at discrete dPU values, because of the ongoing disorder introduction, strain accumulation is continuing after the drop. Finally, in the last stage (above ~ 1.5 dpU), we observe a slight apparent healing of the material, as indicated by the increase of the DW factor. In this particular case the strain increases but, as further discussed below, its evolution depends on the orientation of the MD cell. Moreover, the exact extent of this last region also depends on the orientation.

Fig. 3 reveals that, in the first two stages, the strain evolve in a similar manner, irrespective of the orientation of the MD cell. However, within stage 3, the number of observed strained regions, their level of strain as well as the disorder level required to trigger the strain drop strongly depends on the orientation of the MD cell. This observation points to a Frank → perfect dislocation loop reaction mechanism (Chartier *et al.,* 2016; Jin *et al.,* 2020), which, every time it takes places, locally lowers (*i.e.* relaxes) the strain in the region surrounding the initial Frank loop, hence the abrupt character. Moreover, as mentioned above, even in partially relaxed regions of the cell, the level of strain continues to increase (up to ~ 1.5 dpU when complete relaxation takes place) as a consequence of the continuous disordering.

Above ~1.5 dpU, the strain exhibits an anisotropic behavior: whereas it increases along the [100] (Fig. 3a,b) and [010] directions (Fig. 3c,d), it decreases along [001] (Fig. 3e,f). This is due to the fact that, at high dpU, dislocations form a tangled network (Chartier *et al.,* 2016) that act as traps for the interstitials, and the trapping efficiency depends on the dislocation density in each direction. The disorder also exhibits an anisotropic behavior: on average, it decreases for all orientations, but for the [100] and [001] directions an increase is observed. It should here be reminded that during the MD simulations, the cells are allowed to swell or shrink in all three directions, so that the observed anisotropy is not related to dimensional constraints imposed on the cells.

2 Although this is not obvious from the figure, because of the horizontal logarithmic scale, the slope in the first region is ~8.9%/dpU *vs.* ~5%/dpU. The strain build-up kinetic is reduced in region 2.

Misorientations and shear strains

Similarly to the normal strains, the misorientations and shear strains can be straightforwardly determined from the position of the peaks visible in the RSMs. Figure 4 shows the evolution with disorder level of the off-diagonal elements of the distortion tensor determined from the 002 RSMs (together with the DW factor, for comparison purposes). Up to 0.3 dpU, all components are equal to 0, indicating the absence of shear strains or misorientations. However, as soon as the Frank loops start to transform (in the third stage, above 0.3 dpU), a significant disorder is introduced as evidenced by the appearance of low intensity peaks corresponding to distinct strain/misorientation values randomly distributed in a ±1.5° range. These strained/rotated domains exactly correspond to those already observed in the case of the normal strains, but they are more clearly distinguished here because the difference in shear/misorientations is more pronounced than the differences in the normal strains. As for the normal strains, the correlation with the DW factor is evident.

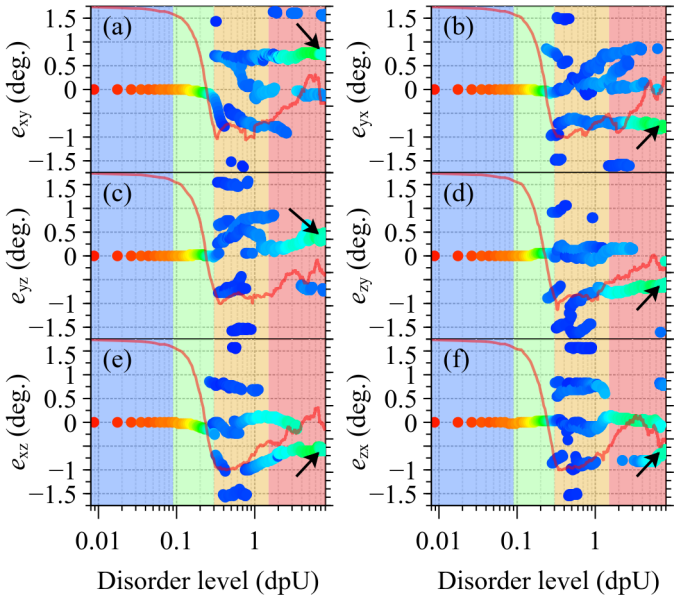

Fig. 4. Evolution of the different components of the strain tensor as a function of the disorder level. The symbol color is proportional to the intensity ratio relatively to the virgin crystal. (a, b): the x and y directions are respectively set parallel to **Q***, while the intensity is integrated along z. (c,d): the y and z directions are respectively set parallel to* **Q***, while the intensity is integrated along x. (e,f): the x and z directions are respectively set parallel to* **Q***, while the intensity is integrated along y. The background colors correspond to the different stages discussed in the text. The thin red line is the DW factor (right axis, graduated from 0 to 1).*

Although the disorder decreases in the last stage, some regions of the MD cells remain permanently rotated or sheared with seemingly randomly distributed values (0.5, 0.6, 0.8 and up to 1.5°). These random values are most likely due to the fact that the dislocations are randomly distributed within the MD cell. The number of misoriented domains is in the 1-3 range depending on the orientation of the MD cell. It can also be observed that in the (x, y) and the (y, z) planes, the disorder is mainly of rotational type, since the largest domains (indicated by the arrows in Fig. 4) are characterized by $e_{xy} = -e_{yx}$ and $e_{yz} = -e_{zy}$. On the contrary, in the (x, z) plane the domains exhibit shear strain, with $e_{xz} = e_{zx}$.

These results show that the dislocation network that forms in the late stages of the disordering process yields to the formation of distorted regions, rotated and sheared with respect to each other. However, contrarily to the case discussed in Fig. 1, where abrupt and incoherent interfaces were formed between the crystallites, eventually leading to finite-size broadening, there is no indication of such boundaries here, *i.e.* the transition from one rotated individual to another is continuous and there is no finite-size induced loss of coherence. This can be clearly confirmed from Fig. 5, which is a close-up of Fig. 2d taken at the intersection of different disordered domains. The occurrence of different peaks in the RSMs must hence not be interpreted as the presence of distinct crystallites, but is a consequence of distorted lattice planes. The finite size of the MD cell and, hence, the limited number of misoriented regions allows one to distinguish individual peaks in the RSMs.

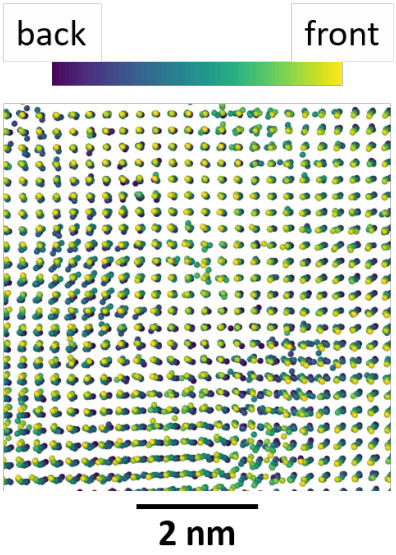

Fig. 5. Close-up in the center of an MD cell at 3 dpU. The atomic displacements are continuous across different regions of the cells.

Towards larger cells: microstrains

Because of the small size of the MD cells considered in this work ($26 \times 26 \times 26$ nm^3), individual peaks can be detected in the RSMs. However, due to the random distribution of shear/misorientation angles previously evidenced, it can be expected that in larger cells, the different peaks emanating from numerous distorted domains, will eventually overlap to form a continuous intensity distribution. This is the very definition of microstrain (see also Appendix B). Although, in the present study, we do not have such large cells, an estimation of the corresponding effect can be inferred by analyzing the overall envelope of the RSMs (dotted contour lines in Fig. 2). Figures 6a-b show the evolution of the FWHM of the intensity distribution, averaged over all orientations, along the *H* and *L* directions, respectively, for both 002 and 004 reflections. The inset in Fig. 6a shows an example of a simulation of an intensity profile along *H* (corresponding to the 3 dpU case of Fig. 2) from which the FWHM is derived. It can be noted that, because of the discontinuous nature of the peak to be simulated (exhibiting several sub-peaks), the simulation is far from perfect. Nonetheless, it is sufficient to get an estimation of the width of the envelope. Moreover, the reliability of the results should not be considered on the basis of a single simulation, but rather on the ensemble of several thousands of data points (6×898 points for each curve of Fig. 6) taking into account the statistical uncertainty given in Fig. 6.

The evolution of the FWHM is strikingly similar to the evolution of the DW factor. This is not unexpected since both parameters depend on the atomic displacement $\delta\mathbf{u}(\mathbf{r})$ in the MD cell. Up to 0.1 dpU, the FWHM is constant and non zero because of the finite size of the MD cells. The FWHM reaches a maximum at the beginning of stage 3 and then steadily decreases. It can be observed that, starting from 0.3 dpU, the width obtained for the 002 and 004 reflections are within a ratio of 2. The fact the width of the peaks scales with the length of the Q vector is a characteristic feature of strain/misorientation disorder (Boulle *et al.*, 2005). Below 0.3 dpU, the ratio differs from 2 because the overall FWHM results from the convolution of both the finite cell size (for which the FWHM ratio is 1) and the contribution of the disorder (for which the FWHM ratio is 2). This effect can be corrected by deconvolution (see section 3) and, as recalled in Appendix B, the components of the microstrain tensor, ε_{ij}, can be computed from the deconvoluted FWHM.

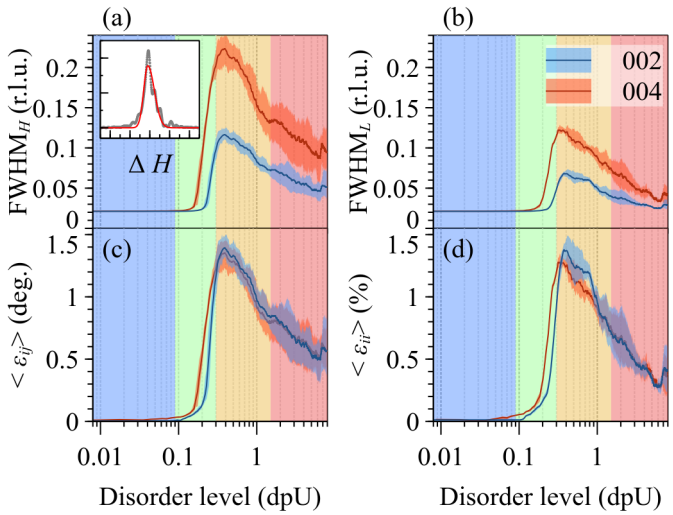

Fig. 6. Evolution, for both the 002 and 004 reflections, of the FWHM along the H (a) and L (b) directions, and the corresponding microstrains (c-d). Inset: example fit of the intensity distribution with an asymmetric Gaussian function (grey points: data points; red line: fit). The horizontal axis extends from -0.25 to 0.25 r.l.u. The colored bands correspond to the uncertainty given by the minimum and maximum values observed for each disorder level. The background colors correspond to the different stages discussed in the text.

The evolution of ε_{ij}, averaged over all orientations, is shown in Fig. 6c,d. The values obtained for the 002 and 004 reflections are identical, within the numerical uncertainty. This finding clearly demonstrates that no finite size effect contributes to the overall broadening (apart from the MD cell size), *i.e.* no incoherent grain boundaries are formed and the atomic coordinates vary smoothly when moving to one rotated/sheared domain to another, as inferred from Fig. 5. For the highest dpU, the residual ε_{ij} is ~0.6°, which means that, with a normal distribution of misorientations, the maximum deviation shall be around ±3 times the standard deviation, that is ±1.8°, which agrees well with values observed in Fig. 4. Similarly, the maximum strain fluctuations should be around ~ 3 × 0.4 = 1.2%, which also agrees with the values observed in Fig. 3. This results demonstrates that, although the MD cell has a limited spatial extension, the rms strain fluctuations that would be observed in a large crystal, sharing the same structural features, can be extrapolated by the analysis of the peak envelope. This extrapolation obviously does not allow to predict features that could take place if the MD simulations were actually performed using larger cells.

4.3. Real-space *vs.* reciprocal-space determinations

An important question to address is how are these reciprocal-space – based strain measurements connect with more direct (*i.e.* real-space) determinations performed from the MD cells[3]. As an example, we shall consider here the results obtained for the average normal strain $<e_{ii}>$ and the average microstrain $<\varepsilon_{ii}>$. The former is given by averaging the values displayed in Fig.3a-f, and the latter is given in Fig. 6d. At this point, the mathematical details of the methodology are beyond the scope of the paper, so we here only provide the most important results. The reader may refer to Appendix C where all relevant details are given.

As mentioned in the Introduction, the local strain determined from the MD cell is sensitive to displacements affecting the first neighbors, or the atoms of the first unit-cell. As a consequence, long-range, spatially correlated atomic displacements are not captured in this measurement, even if the local strain is averaged over the whole MD cell (Xiong *et al.,* 2019). To circumvent this issue, we make use of the directional pair distribution function (D-PDF) introduced by Leonardi *et al.* (Leonardi *et al.,* 2013), which corresponds to the probability of finding a pair of atoms separated by a distance comprised between z and $z + \Delta z$ along a given direction. In accordance with the previous section we here choose these directions to be [100], [010] and [001].

Figure 7a displays the D-PDFs for selected fluences (0, 0.1, 0.3 and 3 dpU), averaged over the three equivalent <100> directions. The effect of strain and disordering is clearly seen from the shift, the damping and the broadening of the peaks. The D-PDF exhibits a series of peaks, the position of which corresponds to the n^{th} neighbor distance, and the width of which corresponds to the rms fluctuations of this distance.

The plot of the peak position $<D>$ *vs.* n (Fig. 7b, left axis), exhibits a seemingly linear behavior, the slope of which allows to determine the state of strain (see Appendix C). The change in slope, consecutive to strain, is more easily seen for large neighbor distances (see inset in Fig. 7b). The corresponding strain is shown in Fig. 7c, together with the range of values obtained from the analysis of computational diffraction (gray shaded region). A remarkably good agreement is obtained between computational diffraction and the real-space measurement. It can also be noticed that when the strain is determined from the first neighbor only (which corresponds to a plain spatial average of the local strain), correct values are obtained for levels of strain lower than ~0.8%, *i.e.* when the dilatation of the unit-cell is solely induced by point defects (both at low and high dpU

3 It must be noticed that these measurements differ from the "macroscopic" swelling of the MD cell, in particular as soon as the concentration of interstitials and vacancies are no longer equal.

levels). As soon as spatial correlations are present, the strain computed from the first neighbor is incorrect by up to a factor 2.

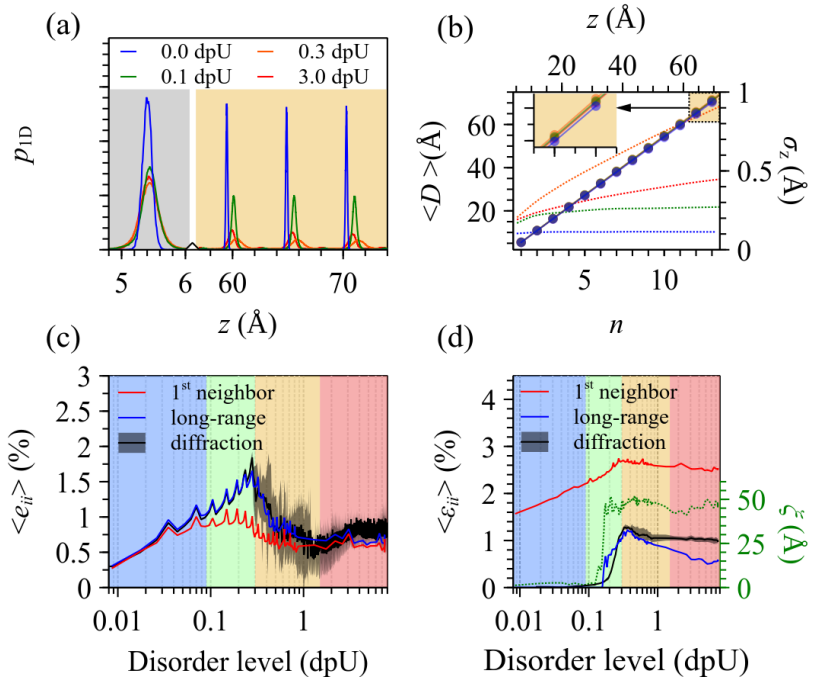

Fig. 7. (a) D-PDF along the <001> direction for increasing disorder. The first peak is of the D-PDF is highlighted (gray area). Higher distances peaks are shown in the light yellow area. (b) left axis: evolution the D-PDF peak position versus the neighbor index (circles and continuous line). Right axis: evolution of the standard deviation of the D-PDF peaks versus the neighbor distance (Warren's plot, dotted lines). The different disorder levels plotted use the same color coding as in (a). (c) strain determined from the <D> vs. n plot (blue line) compared with the values obtained from the RSMs (black line and gray area). The red line indicates the value obtained from the first neighbor distance. (d) left axis: microstrain determined from the σ_z vs. z plot (blue line) compared with the values obtained from the RSMs (black line and gray area). The red line indicates the value obtained from the first neighbor distance. Right axis: evolution of the correlation length (green dotted line).

Let us now consider the plot of the width of the D-PDF peaks σ_z *vs*. z. (Fig. 7b, right axis), which is also known as Warren's plot (Warren & Averbach, 1950; Leonardi *et al.*, 2013). The works of Leonardi *et al.* (Leonardi *et al.*, 2013; Leonardi & Bish, 2017) have demonstrated that real-space and reciprocal-space measurements lead to the same results if correlated and uncorrelated disorder are correctly distinguished from each other in the analysis. In these works, uncorrelated disorder originates from the distorted atomic structure at the grain boundaries, which do not contribute to XRD line broadening. Removing surface atomic layers from the MD cells allowed to have perfect agreement between XRD-derived and D-PDF-derived values of the microstrain. In the present case, as demonstrated earlier, there are no grain boundaries, and we have to rely on a different approach, using the following phenomenological equation to describe the Warren's plots (Boulle *et al.*, 2005):

$$\sigma_z(z) = \sigma_\infty \left\{ 1 - \exp\left[-\left(\frac{z}{\xi} \right)^{1/w} \right] \right\}^{H_u/w} \quad (12)$$

where w defines the width of the transition region between the correlated and the uncorrelated regimes, and H_u (the Hurst exponent) defines how $\sigma_z(z)$ scales with z in the correlated regime; for $H_u = 1$, the scaling is linear as described in Appendix C. Equipped with Eq. 12 it is possible to model the evolution of the standard deviation of the peaks of the D-PDF as a function of the neighbor distance, σ_z *vs.* z , and hence to retrieve the values σ_∞ and ξ. From these values the microstrain, ε_{zz}, is obtained with σ_∞ / ξ (Boulle *et al.*, 2005). The parameters H_u and w can also be obtained from the simulation, although we do not discuss their evolution in details in this paper[4].

In the non disordered sample (Fig. 7b, blue curve), the width of the peak remains constant whatever the neighbor distance, indicating a purely uncorrelated disorder and the broadening originates from thermal displacements. For 0.1 dpU (Fig. 7b, green curve), the disorder is dominated by point defects and it can be observed that the value of σ_z increases as compared to 0 dpU but remains constant over the whole z range (except for the two first neighbors where a small increase is detectable). The disorder is still essentially uncorrelated. The situation is drastically different for 0.3 and 3 dpU where a neat increase of σ_z is observed for increasing the neighbor distance, hence indicating spatial correlations.

The determination of σ_∞ and ξ requires a careful data analysis. In particular, at low disorder (below 0.1 dpU) the disorder is induced by point defects and is largely uncorrelated. In such a situation the disorder is constant and equal to σ_∞ and the correlation length should be zero. However, $\xi = 0$ is numerically undefined in Eq. 12 and, for those cases, a minimum value of 1 has been set for the fit. Subsequently, we set a threshold value for ξ below which it is considered as 0 and the corresponding microstrain is 0 as well. Above the threshold value, the microstrain is normally given by $\varepsilon_{zz} = \sigma_\infty / \xi$. The threshold value has been set to 25 Å; this value is justified below. The corresponding results are shown in Fig. 7d, left axis, together with the range of values obtained by the reciprocal space analysis (gray shaded region). Although the agreement is less perfect than for the homogeneous strain, it is still remarkably good (this is further discussed below).

The green dotted curve shows the evolution of the correlation length ξ. It is observed that, before the formation of dislocations loops, this value is essentially 0, whereas a sharp increase is observed around 0.15 dpU, *i.e.* when the presence of dislocations loops becomes significant. The inflection occurs for a value of $\xi \sim 25$ Å, this value has therefore been chosen as threshold value for the

4 It can be noticed that, in the correlated regime, Eq. 12 reduces to Eq. 4 in (Adler & Houska, 1979) where our H_u equals their $r + 1$ exponent.

calculation of ε_{zz}. The correlation length saturates at a value of 50 Å , *i.e.* the atomic displacements are correlated over ~1/4 of the MD cell as long as dislocations are present.

It can further be mentioned that the maximum disorder, σ_∞ (not shown), can be used as an overall measure of disorder in the MD cell. The DW can be computed from σ_∞ within the Gaussian approximation, see, for instance (Trueblood *et al.,* 1996). However, in the case of irradiated materials, the disorder has been proven to be largely non-Gaussian (Boulle & Debelle, 2016) so that this determination of the DW factor may lead erroneous results.

We now focus on the result that would have been obtained by neglecting spatial correlations, *i.e.* by computing the microstrain from the first neighbor only (Fig. 7d, blue curve). As can be observed, this measurement largely overestimates the microstrain in the whole investigated range, consistently with other studies (Stukowski *et al.,* 2009; Xiong *et al.,* 2019). The reason for this discrepancy is that peak broadening is sensitive to correlated displacements, *via* the $\langle \exp\{i\mathbf{Q}[\delta\mathbf{u}(\mathbf{r}_j) - \delta\mathbf{u}(\mathbf{r}_k)]\}\rangle$ term in Eq. 3. This term yields results similar to those obtained with the real space average only if the covariance of $\delta\mathbf{u}$ is equal to 0 (*i.e.* no correlations). Moreover this measurement does not provide any indication as to the presence of spatial correlations in the defect structure as evidenced by the lack of any significant feature in the 0.1 – 0.2 dpU range.

Let us now consider the discrepancy observed between the XRD-derived and the D-PDF-derived values in Fig. 7d. One possible reason for this discrepancy is that the real-space measurements requires a more complex numerical analysis (including a fitting procedure with Eq. 12), which is, hence, more prone to numerical errors than computational diffraction. A more likely reason lies in the assumption made regarding the shape of the XRD profiles. Throughout the analysis, the peak shapes were assumed to be Gaussian. The comparison between the real-space and the reciprocal-space analysis shown in Fig. 7b, demonstrates that this assumption is acceptable up to 1 dpU. Above this value, both measurements start to deviate from each other. Interestingly this region corresponds to the moment where perfect dislocation lines are starting to form, and it is known that randomly distributed dislocation lines yield non Gaussian peak profiles (Adler & Houska, 1979). This can be confirmed by inspecting the value of the Hurst exponent, H_u. In the range 0.3 – 1 dpU, its average value is 0.75, whereas it drops to 0.39 above 1 dpU. The former value corresponds to an XRD shape intermediate between a Gaussian (H_u = 1) and a Lorentzian (H_u = 0.5) function, whereas the latter is close to the Lorentzian peak shape (Adler & Houska, 1979; Boulle *et al.,* 2005). This shows the limit on the utilization of the Gaussian approximation and, depending on the situation, more flexible shapes could be considered (like Voigt or pseudo-Voigt functions).

To conclude this part, we have demonstrated that computational diffraction allows to determine the components of the homogeneous and heterogeneous strain tensors. The consistency of the results have been checked by comparing the values (of the average of diagonal elements) obtained by diffraction with those obtained from a real-space approach using D-PDF. Finally, a significant advantage of computational diffraction, as compared to the real-space approach is that the off-diagonal elements would be far more complicated to assess in the latter case.

5. Conclusions

We presented a computational diffraction method for a rapid and reliable determination of the distortion and microstrain tensors, as well as the associated disorder, in atomic scale simulations of disordered crystals. The analysis is based on a projected 3D Fourier transform followed by a 2D peak position and shape analysis. With an adapted hardware, and depending on the size of the simulation cell, the whole process takes a few seconds for a given cell orientation. The main advantage of the computational diffraction approach, as compared to usual real-space approaches, is that the normal strains and distortions, microstrains and disorder (DW factor) are straightforwardly determined from the peak positions, widths and intensities in the reciprocal space maps. Moreover, contrarily to most real-space approaches, computational diffraction includes the effect of long-range correlations in the atomic displacement field. Furthermore, as compared to the Debye scattering equation, the 3D distortion tensor can be reliably determined without orientational averaging and without necessitating advanced XRD pattern modeling.

The relevance of computational diffraction has been demonstrated in a self-consistent way by comparing characteristic parameters (such as the normal strains and microstrains) with those obtained from real space measurements. Since the whole analysis can be completely automated, it can easily be envisioned for on-the-fly data analysis where MD simulations snapshots are analyzed as soon as they are produced, allowing one to extract the data in real time without the need to save the simulation cells themselves. This *in situ* computational diffraction approach (similar to the *in situ* computational microscopy introduced in (Zepeda-Ruiz *et al.*, 2017)) allows one to significantly compress the amount of data to be stored. The efficiency of computational diffraction is also perfectly suited to analyze very large data sets in an automated way.

Appendix A: Effect of Q_y integration

In most actual laboratory XRD experiments, x-ray beams are in general only collimated in one plane (the plane in which the photons are detected), whereas they are divergent in the orthogonal plane (this geometry is schematically depicted in Fig. S2, Supplementary material). This results in a loss of spatial coherence in the direction perpendicular to the detection plane and an intensity integration as given by Eq. (13), see also (Channagiri *et al.*, 2015).

In the present work we take advantage of this effect to approximate a Monte-Carlo averaging in order to wipe out interference features that are not usually observed in experiments:

$$\langle I(Q_x, Q_z)\rangle = \int dQ_y \cdot |E(\mathbf{Q})|^2 \approx \sum_{\Delta Q_y=-N_y/2}^{N_y/2} \left|E\left[(Q_x, Q_z)_{\Delta Q_y}\right]\right|^2 \tag{13}$$

where E is the usual scattered X-ray amplitude $\sum_{j=1}^{N} f_j(\mathbf{Q}) \exp(i\mathbf{Q}\cdot\mathbf{r}'_\mathbf{j})$, N_y is the number of (Q_x, Q_z) sections involved in the summation, the appropriate value of which can be determined empirically by increasing N_y to determine the the convergence of the sum (Fig. S3, Supplementary material). All calculations presented in the current work are computed using the configuration with 50 sections. An interesting analogy that can be made from this equation is that, in the diffraction experiment, the crystal can be viewed as divided into unit-cell-thick (*x, z*) planes, parallel to the detection plane. The intensity diffracted from each (*x, z*) plane is given by the squared modulus of the amplitude scattered from each plane, and the intensities are summed to form the total scattered intensity. This operation introduces randomness in an otherwise fixed defect configuration.

Appendix B: Effect of microstrain

The first exponential of Eq. 3 is also referred to as the correlation function (Holý *et al.*, 1999) , $G(\mathbf{r}_j, \mathbf{r}_k)$, [or the strain Fourier coefficient (Warren, 1969) in the field of powder diffraction] and it determines the intensity and the shape of the diffuse scattering distribution. The important feature to notice is that $G(\mathbf{r}_j, \mathbf{r}_k)$ is a function of the *local displacement difference function* $\delta\mathbf{u}(\mathbf{r}_j) - \delta\mathbf{u}(\mathbf{r}_k)$: it encodes the degree of correlation between local displacements, $\delta\mathbf{u}$, from one site to another (Boulle *et al.*, 2005). In the case of correlated disorder, the individual displacements add up, which yields a destruction of the long range order. In, such a case, $G(\mathbf{r}_j, \mathbf{r}_k)$ rapidly decreases for increasing $\mathbf{r}_j - \mathbf{r}_k$ pair distances (Krivoglaz, 1969). The displacement difference can therefore safely be expanded into a first order Taylor series so that the correlation function finally writes:

$$G(\mathbf{r_j}, \mathbf{r_k}) = \langle \exp\{i\left[\mathbf{Q}\nabla\delta\mathbf{u}(\mathbf{r})\right]\cdot(\mathbf{r_j}-\mathbf{r_k})\}\rangle \tag{14}$$

$\nabla\delta\mathbf{u}(\mathbf{r})$ is the Jacobian of the local displacements. In order to illustrate how $G(\mathbf{r}_j, \mathbf{r}_k)$ affects the scattered intensity, let us consider the same two extreme cases as for the coherent scattering.

For a perfect crystal, $\delta\mathbf{u}(\mathbf{r}) = 0$ so that $G(\mathbf{r}_j, \mathbf{r}_k) = 0$, and there is no diffuse scattering, *i.e.* DW = 1. On the contrary for highly disordered crystals, $\delta\mathbf{u}(\mathbf{r}) >> 0$, DW$\rightarrow$ 0, the coherent intensity vanishes and the shape of the intensity distribution is then entirely governed by the correlation function, which is itself a function of the statistical defect distribution and the associated displacement fields; increasing correlated disorder, for instance by increasing the defect density, yields a broadening of the intensity peak. An exact solution for the correlation can be worked out for selected defects such as point defects, defect clusters, dislocation loops,etc (Dederichs, 1971, 1973; Ehrhart *et al.*, 1982; Iida *et al.*, 1988; Larson, 2019) but this is not the topic of the present article.

For illustration purposes, we shall assume that the statistical distribution of the local displacement difference function obeys a multivariate Gaussian distribution. The correlation function then writes

$$\langle \exp\{i\,[\mathbf{Q}\nabla\delta\mathbf{u}(\mathbf{r})]\cdot(\mathbf{r_j}-\mathbf{r_k})\}\rangle = \exp\left[-\frac{1}{2}(\mathbf{r_j}-\mathbf{r_k})^{\mathbf{T}}\boldsymbol{\Sigma}_{\mathbf{Q}\delta\mathbf{u}}(\mathbf{r_j}-\mathbf{r_k})\right] \tag{15}$$

where $\Sigma_{\mathbf{Q}\delta\mathbf{u}}$ is 3×3 tensor whose components are related to the microstrain tensor. For instance, using a $\mathbf{Q} = (0, 0, Q_z)^T$ vector and assuming, that the displacements δu_x, δu_y and δu_z are statistically independent, the tensor writes:

$$\boldsymbol{\Sigma}_{\mathbf{Q}\delta\mathbf{u}} = \begin{pmatrix} (\epsilon_{zx}Q_z)^2 & 0 & 0 \\ 0 & (\epsilon_{zy}Q_z)^2 & 0 \\ 0 & 0 & (\epsilon_{zz}Q_z)^2 \end{pmatrix} \tag{16}$$

in which case, Eq. (15) reduces to the product of three Gaussian functions with standard deviations $\varepsilon_{zx}\,Q_z$, $\varepsilon_{zy}\,Q_z$, $\varepsilon_{zz}\,Q_z$, where $\varepsilon_{i,j}$ are the components of the microstrain tensor. In other words, measuring the width of the peak in the RSM allows to determine the heterogeneous strain, which is the basis of all XRD line width analysis techniques related to Williamson & Hall pioneering work (Williamson & Hall, 1953). It is also clearly established that this approach is an oversimplification that relies on strong assumptions (Gaussian distribution of strain, statistically independent displacements). It nonetheless allows to get reasonable orders of magnitude.

Appendix C: Real-space *vs.* reciprocal-space determinations of strain

To address this task we first suggest to rewrite the equation of the total diffracted intensity (Eq. 3), by introducing the n^{th} neighbor index, $n = k - j$. Without loss of generality we shall focus on the Q_z intensity distribution of reflection with $\mathbf{Q}_0 = (0, 0, Q_0)^T$. The conclusions drawn remain perfectly

valid for any reflection or any direction, although the derivation of the corresponding real space quantities (in particular the shear/misorientational components of the strain tensors) would not be as straightforward. The corresponding intensity writes:

$$I(Q_z) = \sum_j \sum_n f_j(Q_z) f^*_{j+n}(Q_z) \times \langle \exp\{iQ_z\,[\delta u_z(z_j) - \delta u_z(z_{j+n})]\}\rangle \times \exp(i\,\{Q_z\,[1 + e_{zz}(z_j)]\} \cdot [z_j - z_{j+n}]) \tag{17}$$

The term in the angular brackets can be reduced to

$$\langle \exp\{iQ_z\,[\delta u_z(z_j) - \delta u_z(z_{j+n})]\}\rangle = \exp\left[-\frac{1}{2}Q_z^2\,\langle \delta u_z(z_j)\rangle^2\right] \tag{18}$$

where we implicitly assumed a Gaussian distribution to perform the average; more general distributions can be considered (Boulle *et al.*, 2005; Boulle & Debelle, 2016), but this does not change the conclusions. The strain values, e_{zz}, and the associated microstrain, ε_{zz}, computed from the MD data using these equations exactly correspond to those given in Fig. 4 and Fig. 7.

As detailed in section 2, we write the *z* coordinate of all atoms as

$$z' = z + e_{zz} \cdot z + \delta u_z(z) \tag{19}$$

and we make use of the D-PDF introduced in (Leonardi *et al.*, 2013), $p_{1D}(z)$, which corresponds to the probability of finding a pair of atoms separated by a distance comprised between *z* and $z + \Delta z$ along the [001] direction. The D-PDF can directly be evaluated from the MD cells (see below). This function exhibits maxima at the average n^{th} neighbor distance (Fig. 7a):

$$\langle D(n)\rangle = \left\langle \left(z'_j - z'_{j+n}\right)\right\rangle_j \tag{20}$$

where the average is taken over all atoms *j* in the cell. Using Eq. (19), the n^{th} neighbor distance can be rewritten

$$\langle D(n)\rangle = n\,\langle d(1)\rangle\,(1 + e_{zz}) \tag{21}$$

where $d(1)$ is the first neighbor distance in a disorder-free MD cell. The lattice strain can hence be directly deduced from the slope of the peak positions in the D-PDF *vs.* the neighbor index *n* (Fig. S3b). Moreover, comparing Eq. (21) with the last exponential of Eq. (17) reveals that the strain measured from calculated intensity should be strictly equal to the strain computed from Eq. (21).

Similarly, the variance of the peaks of the D-PDF is:

$$\left\langle D(n)^2 \right\rangle = \left\langle \left(z'_j - z'_{j+n} \right)^2 \right\rangle_j = \left\langle \delta u_z(z)^2 \right\rangle_j \tag{22}$$

which exactly corresponds to the term contained in Eq. (18). Depending on the state of correlation of the displacements, different behavior are observed. If all displacements are independent and uncorrelated, the width of the peaks of the D-PDF is constant and equal to $\sqrt{\langle \delta u_z^2 \rangle}$. On the contrary if all displacements add up, $\delta u_z(z) = \sum_{\zeta \leq z} \delta u_\zeta$, then the width of peaks of the D-PDF increase as $n\sqrt{\langle \delta u_z^2 \rangle}$. Intermediate behaviors require the introduction of an additional quantity, the correlation length ξ, which defines the length up to which the displacement are correlated; for neighbors separated by distances larger than ξ, correlation is lost and the disorder saturates to a constant value, σ_∞. A phenomenological equation capturing such a behavior is given by Eq. 12.

Finally, to obtain the D-PDF from the MD cell, we calculated the distribution of distances between neighbours considering for each atom a centred square bar oriented in the direction of interest. We chose square bars with a basal section 1 $Å^2$ and a length of 80 Å as it turned out to be sufficient to capture the main trends.

References

Adler, T. & Houska, C. R. (1979). *J. Appl. Phys.* **50**, 3282–3287.

Boulle, A. & Debelle, A. (2016). *Phys Rev Lett*. **116**, 245501.

Boulle, A., Guinebretière, R. & Dauger, A. (2005). *J. Phys. Appl. Phys.* **38**, 3907–3920.

Catillon, G. & Chartier, A. (2014). *J. Appl. Phys.* **116**, 193502.

Channagiri, J., Boulle, A. & Debelle, A. (2015). *J Appl Cryst*. **48**, 252.

Chartier, A., Catillon, G. & Crocombette, J.-P. (2009). *Phys Rev Lett*. **102**, 155503.

Chartier, A., Meis, C., Crocombette, J.-P., Weber, W. J. & Corrales, L. R. (2005). *Phys. Rev. Lett.* **94**, 025505.

Chartier, A., Onofri, C., Van Brutzel, L., Sabathier, C., Dorosh, O. & Jagielski, J. (2016). *Appl. Phys. Lett.* **109**, 181902.

Chartier, A., Van Brutzel, L. & Pageot, J. (2018). *Carbon*. **133**, 224–231.

Crocombette, J.-P. & Chartier, A. (2007). *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **255**, 158–165.

Crocombette, J.-P., Chartier, A. & Weber, W. J. (2006). *Appl. Phys. Lett.* **88**, 051912.

Debelle, A., Boulle, A., Chartier, A., Gao, F. & Weber, W. J. (2014). *Phys Rev B*. **90**, 174112.

Debye, P. (1915). *Ann. Phys.* **351**, 809–823.

Dederichs, P. H. (1971). *Phys. Rev. B*. **4**, 1041–1050.

Dederichs, P. H. (1973). *J. Phys. F Met. Phys.* **3**, 471–496.

Derlet, P. M., Van Petegem, S. & Van Swygenhoven, H. (2005). *Phys. Rev. B*. **71**, 024114.

Devanathan, R., Brutzel, L. V., Chartier, A., Guéneau, C., Mattsson, A. E., Tikare, V., Bartel, T., Besmann, T., Stan, M. & Uffelen, P. V. (2010). *Energy Environ. Sci.* **3**, 1406–1426.

Ehrhart, P., Trinkaus, H. & Larson, B. C. (1982). *Phys. Rev. B*. **25**, 834–848.

Favre-Nicolin, V., Coraux, J., Richard, M.-I. & Renevier, H. (2011). *J. Appl. Crystallogr.* **44**, 635–640.

Gelisio, L. & Scardi, P. (2016). *Acta Crystallogr. Sect. Found. Adv.* **72**, 608–620.

Gullett, P. M., Horstemeyer, M. F., Baskes, M. I. & Fang, H. (2008). *Model. Simul. Mater. Sci. Eng.* **16**, 015001.

Holý, V., Pietsch, U. & Baumbach, T. (1999). *High Resolut. X-Ray Scatt. Multilayers Thin Films Springer Tracts Mod. Phys.* **149**,.

Iida, S., Larson, B. C. & Tischler, J. Z. (1988). *J. Mater. Res.* **3**, 267–273.

Inagaki, M., Sasaki, Y. & Sakai, M. (1983). *J. Mater. Sci.* **18**, 1803–1809.

Jin, X., Boulle, A., Chartier, A., Crocombette, J.-P. & Debelle, A. (2020). *Acta Mater.* **201**, 63–71.

Krasheninnikov, A. V. & Nordlund, K. (2010). *J. Appl. Phys.* **107**, 071301.

Krivoglaz, M. A. (1969). Theory of X-Ray and Thermal Neutron Scattering by Real Crystals Springer US.

Larson, B. C. (2019). *Crystals*. **9**, 257.

Leonardi, A. & Bish, D. L. (2017). *Acta Mater.* **133**, 380–392.

Leonardi, A., Leoni, M., Li, M. & Scardi, P. (2012). *J. Nanosci. Nanotechnol.* **12**, 8546–8553.

Leonardi, A., Leoni, M. & Scardi, P. (2013). *J. Appl. Crystallogr.* **46**, 63–75.

Leonardi, A. & Scardi, P. (2015). *Front. Mater.* **1**, 1–10.

Lin, Z. & Zhigilei, L. V. (2006). *Phys. Rev. B*. **73**, 184113.

Morelon, N.-D., Ghaleb, D., Delaye, J.-M. & Brutzel, L. V. (2003). *Philos. Mag.* **83**, 1533–1555.

Mott, P. H., Argon, A. S. & Suter, U. W. (1992). *J. Comput. Phys.* **101**, 140–150.

Nordlund, K. & Djurabekova, F. (2014). *J. Comput. Electron.* **13**, 122–141.

Plimpton, S. (1995). *J. Comput. Phys.* **117**, 1–19.

Rebuffi, L., Troian, A., Ciancio, R., Carlino, E., Amimi, A., Leonardi, A. & Scardi, P. (2016). *Sci. Rep.* **6**, 20712.

Rietveld, H. M. (1969). *J. Appl. Crystallogr.* **2**, 65–71.

Scardi, P., Leonardi, A., Gelisio, L., Suchomel, M. R., Sneed, B. T., Sheehan, M. K. & Tsung, C.-K. (2015). *Phys. Rev. B*. **91**, 155414.

Scardi, P. & Leoni, M. (2002). *Acta Cryst A*. **58**, 190.

Soulié, A., Menut, D., Crocombette, J.-P., Chartier, A., Sellami, N., Sattonnay, G., Monnet, I. & Béchade, J.-L. (2016). *J. Nucl. Mater.* **480**, 314–322.

Stukowski, A. (2012). *Model. Simul. Mater. Sci. Eng.* **20**, 045021.

Stukowski, A. & Arsenlis, A. (2012). *Model. Simul. Mater. Sci. Eng.* **20**, 035012.

Stukowski, A., Markmann, J., Weissmüller, J. & Albe, K. (2009). *Acta Mater.* **57**, 1648–1654.

Trueblood, K. N., Bürgi, H.-B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U. & Abrahams, S. C. (1996). *Acta Crystallogr. A*. **52**, 770–781.

van der Walt, S., Colbert, S. C. & Varoquaux, G. (2011). *Comput. Sci. Eng.* **13**, 22–30.

Warren, B. E. (1969). X-Ray Diffraction.

Warren, B. E. & Averbach, B. L. (1950). *J. Appl. Phys.* **21**, 595–599.

Williamson, G. K. & Hall, W. H. (1953). *Acta Metall.* **1**, 22–31.

Xiong, S., Lee, S.-Y. & Noyan, I. C. (2019). *J. Appl. Crystallogr.* **52**, 262–273.

Zepeda-Ruiz, L. A., Stukowski, A., Oppelstrup, T. & Bulatov, V. V. (2017). *Nature*. **550**, 492–495.

Zhang, L., Jasa, J., Gazonas, G., Jérusalem, A. & Negahban, M. (2015). *Comput. Methods Appl. Mech. Eng.* **283**, 1010–1031.

Zimmerman, J. A., Bammann, D. J. & Gao, H. (2009). *Int. J. Solids Struct.* **46**, 238–253.