

Supporting Information

Mapping Morphological and Structural Properties of Lead Halide Perovskites by

Scanning Nanofocus XRD

Samuele Lilliu^{1,}, Thomas G. Dane², Mejd Alsari³, Jonathan Griffin¹, , Alexander T. Barrows¹, Marcus S. Dahlem⁴, Richard H. Friend³, David G. Lidzey¹, J. Emyr Macdonald⁵*

Dr. S. L. Author 1, Dr. J. G. Author 4, Dr. A. T. B. Author 5, Prof. Dr. D. G. L. Author 8
Department of Physics and Astronomy, University of Sheffield, Hicks Building, Hounsfield
Road, Sheffield S3 7RH, United Kingdom
E-mail: samuele_lilliu@hotmail.it, s.lilliu@sheffield.ac.uk

Dr. T. G. D. Author 2
European Synchrotron Radiation Facility, BP 220, Grenoble F-38043, France

Miss M. A. Author 3, Prof. Dr. Sir R. H. F. Author 7
Cavendish Laboratory, University of Cambridge, Madingley Road, CB3 0HE Cambridge,
United Kingdom

Prof. Dr. M. S. D. Author 6,
Masdar Institute of Science and Technology, PO Box 54224, Abu Dhabi, United Arab
Emirates

Prof. Dr. J. E. Macdonald Author 9
School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, CF24
3AA Cardiff, United Kingdom

1 Introduction

The main purpose of this Supporting Information is to illustrate in great detail the analysis of data obtained from Scanning Nanofocus XRD (nXRD). We also briefly report GI-WAXS measurements and provide instructions for accessing the nXRD raw data. The paragraphs illustrated in this document (quoted in *italic*) are here briefly summarized.

In *Grazing Incidence Wide Angle X-Ray Scattering (GI-WAXS)* we report on GI-WAXS measurements performed on the cold, medium and hot samples.

In *nXRD Measurements and Data Analysis* we describe the nXRD beamline setup, how the measurements were conducted, and technical details of the data analysis. The paragraph contains several subparagraphs. In *Spatial Resolution* we discuss the resolution limits of nXRD. In *Noise Removal and Average Diffraction Pattern* we discuss how images were denoised and how average diffraction patterns were constructed by averaging over the 10 201 images acquired in a single scan. In *Montages* we show how a large ‘montage’ image was constructed by concatenating the 10 201 patterns acquired in a single scan. We provide a code to rebuild the montage and navigate it. In *Remapping: from Pixel to Scattering Vector Coordinates* we discuss the calibration procedure and the image remapping from pixel to scattering vector coordinate in transmission geometry. We provide a rigorous definition of the scattering vector expressed in polar coordinates q and χ . We also comment on the resolution. In *Azimuthally Integrated Line Profiles* we show how the azimuthally integrated line profiles were constructed. The obtained line profiles were used in the Rietveld refinement illustrated in *Refinement*. In *Circular Region of Interest Around the Main Perovskite Double Peak* we discuss the region of the diffraction patterns where the analysis was performed. In *First Clustering Step* we show how diffraction spots in single diffraction patterns were isolated or clustered by employing pairwise Euclidean distance between the pixels. In *Preliminary Quiver Plot* we illustrate the concept of the ‘quiver plot’, which is a way of representing the χ coordinate of each detected peak in spatial coordinates. In *Second Clustering Step* we show how the quiver plot and the spatial coordinates were employed to perform a second level of clustering which allows us to isolate single grains based on their χ orientation. We then show a method to overcome residual distortion and isolate (002) grains from (110) grains. In *Maps* we discuss the information that can be extracted from single grains extracted from the second clustering step, in the form of local structural maps (e.g. lattice constant or diffraction spot intensity). We report these maps for the cold, medium and hot samples. In *Quiver Plots Magnifications* we show magnifications of the quiver plots illustrated in the main text.

In the following paragraph, *Raw Data Instructions* we provide instructions (and MATLAB code) on how to use the raw data available as supporting files to plot again the data discussed in the manuscript and in the supporting information.

2 Grazing Incidence Wide Angle X-Ray Scattering (GI-WAXS)

Details on the BM28 (XMaS, ESRF, Grenoble, France) beamline are available at www.xmas.ac.uk. Details on the experimental setup and the GI-WAXS geometry can be found in ref.[1] and some are reported here again for clarity.

The radiation coming from a bending magnet (critical energy $E_c = 9.8$ keV) was monochromatised using a fixed-exit, water-cooled, double crystal Si(111) monochromator, placed at 25 m from the source. A Rh-coated toroidal mirror was used to focus the monochromatic beam horizontally and vertically. The X-ray beam was focused to a beam spot size of $\sim 500 \times 100 \mu\text{m}^2$, resulting in beam footprint of $\sim 500 \times (100 / \sin(\alpha_i)) \mu\text{m}^2$, where α_i is the (out-of-plane) incident angle.

Cold, medium, and hot spun films were prepared as described in the experimental section and spin-coated at the beamline, annealed, and measured within one day. Here, samples were housed in a helium-filled custom made environmental chamber mounted on an 11-axis Huber diffractometer.^[1] A MAR 165 CDD detector (2048 rows \times 2048 columns) was placed on the diffractometer arm at a distance of 223 mm from the sample (calibrated as in ref. ^[2, 3]). Single diffraction patterns were obtained by illuminating the samples for 10 s with an X-ray beam having a wavelength of ~ 1.2384 Å. Data processing was performed using MATLAB software.^[4]

Figure S1 shows GI-WAXS diffraction patterns recorded for the cold, medium and hot spun MAPI films measured above the critical angle ($\alpha_c \approx 0.16^\circ$) at an out-of-plane incident angle equal to $\alpha_i \approx 0.3^\circ$, achieving a penetration depth of ~ 140 nm.

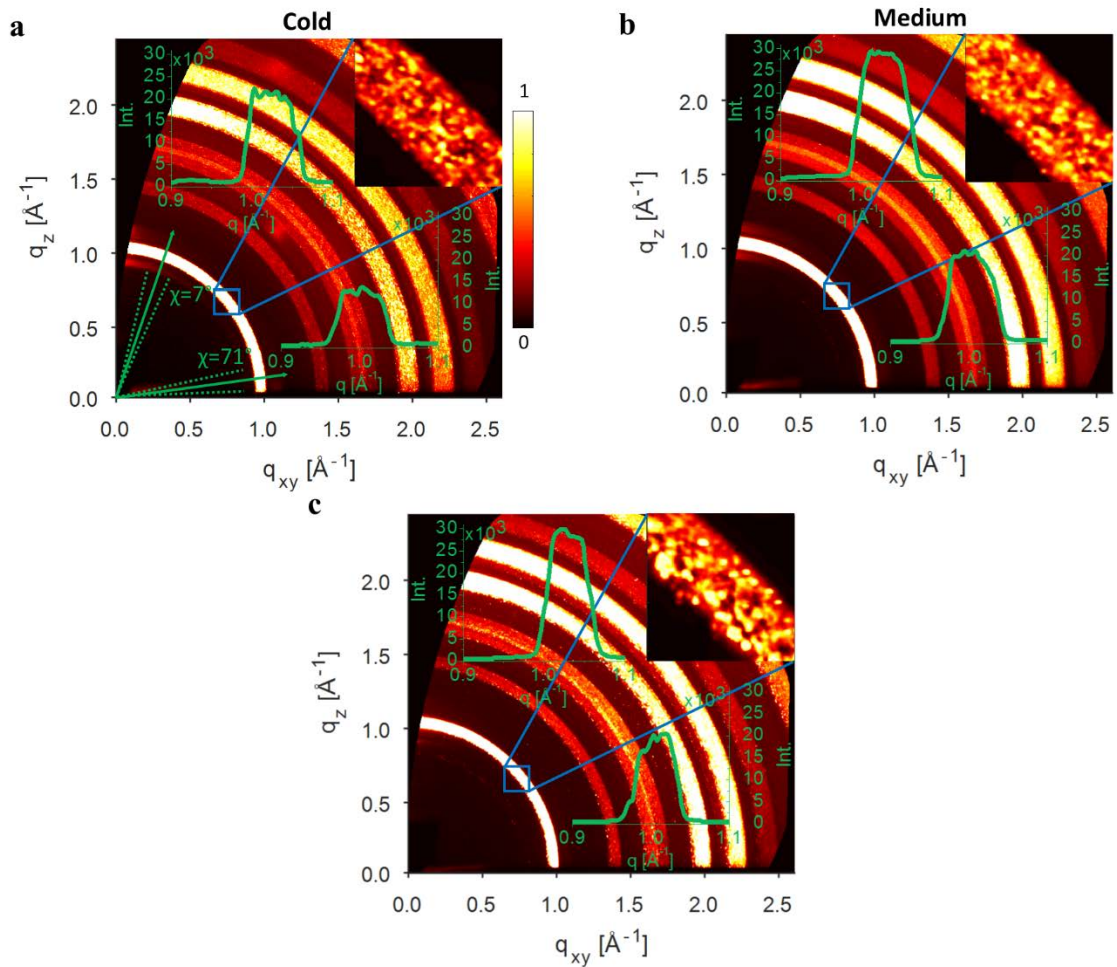


Figure S1. GI-WAXS diffraction patterns of cold (a), medium (b) and hot (c) spun MAPI films. The arrows in part a indicate the azimuthal angle (χ) of the cake slice used for extracting the out-of-plane line profile ($\chi = 7^\circ$) and the in-plane line profile ($\chi = 71^\circ$), which are both shown in each figure. The dotted lines indicate the cake slice aperture ($\chi = 10^\circ$) that is used in the line profile integration. The images in the top corner represent a zoom of the square region of interest indicated in blue. Diffraction patterns are normalized between 0 and 1 and are represented on a linear scale.

For the calculation of the GI-WAXS penetration depth we used the following expression:^[5]

$$\Lambda \approx \left[\sqrt{2}k \sqrt{\sqrt{(\alpha_i^2 - \alpha_c^2)^2 + 4\beta^2} - (\alpha_i^2 - \alpha_c^2)} \right]^{-1} \quad (1)$$

where $k = 2\pi/\lambda$ is the wavenumber (λ is the X-ray wavelength), α_c is the critical angle, α_i is the out-of-plane incident angle, and β is the imaginary part of the index of refraction. The index of refraction is defined as $n = 1 - \delta + i\beta$.^[5] The critical angle was calculated as $\alpha_c = \sqrt{2\delta}$. The δ and β parameters for $\text{CH}_3\text{NH}_3\text{PbI}_3$ at $\lambda = 1.2398 \text{ \AA}$ were estimated from the online toolbox available at ref. ^[6], and are equal to 3.76813×10^{-6} and 3.06721×10^{-7} , respectively. A plot of the penetration depth for these experiments is shown in **Figure S2**.

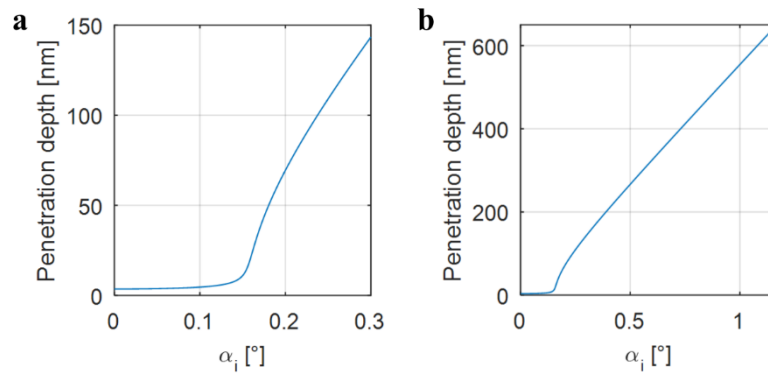


Figure S2. Calculated X-ray penetration depth in GI-WAXS vs incident angle. a, range of out-of-plane incidence angles used in the scans reported here (α_i). **b**, penetration depth calculated over a larger range of α_i . The incident angles required to probe all the bulk in the cold (~350 nm thick), medium (~500 nm thick), and hot sample (~650 nm thick) are 0.64° , 0.9° , and 1.17° respectively. Above 0.3° the sample edge in GI-WAXS diffraction patterns increasingly hides the in-plane region of the diffraction pattern.

Close inspection of the Debye-Scherrer rings^[7] evident in Figure S1 indicates that they are in fact composed by thousands of diffraction spots,^[1] as can be seen in the zoom in the insets. We plot further in-plane and out-of-plane line profiles at different incident angles in **Figure S3** and **Figure S4**. Interestingly, these measurements do not reveal any significant structural differences between the three samples nor any information on depth-related structural variation.

The only difference between the samples consists in the perovskite peak intensity, which is higher in the medium and hot samples. This corresponds to a higher degree of crystallinity in the medium and hot samples, which is confirmed by the nXRD measurements.

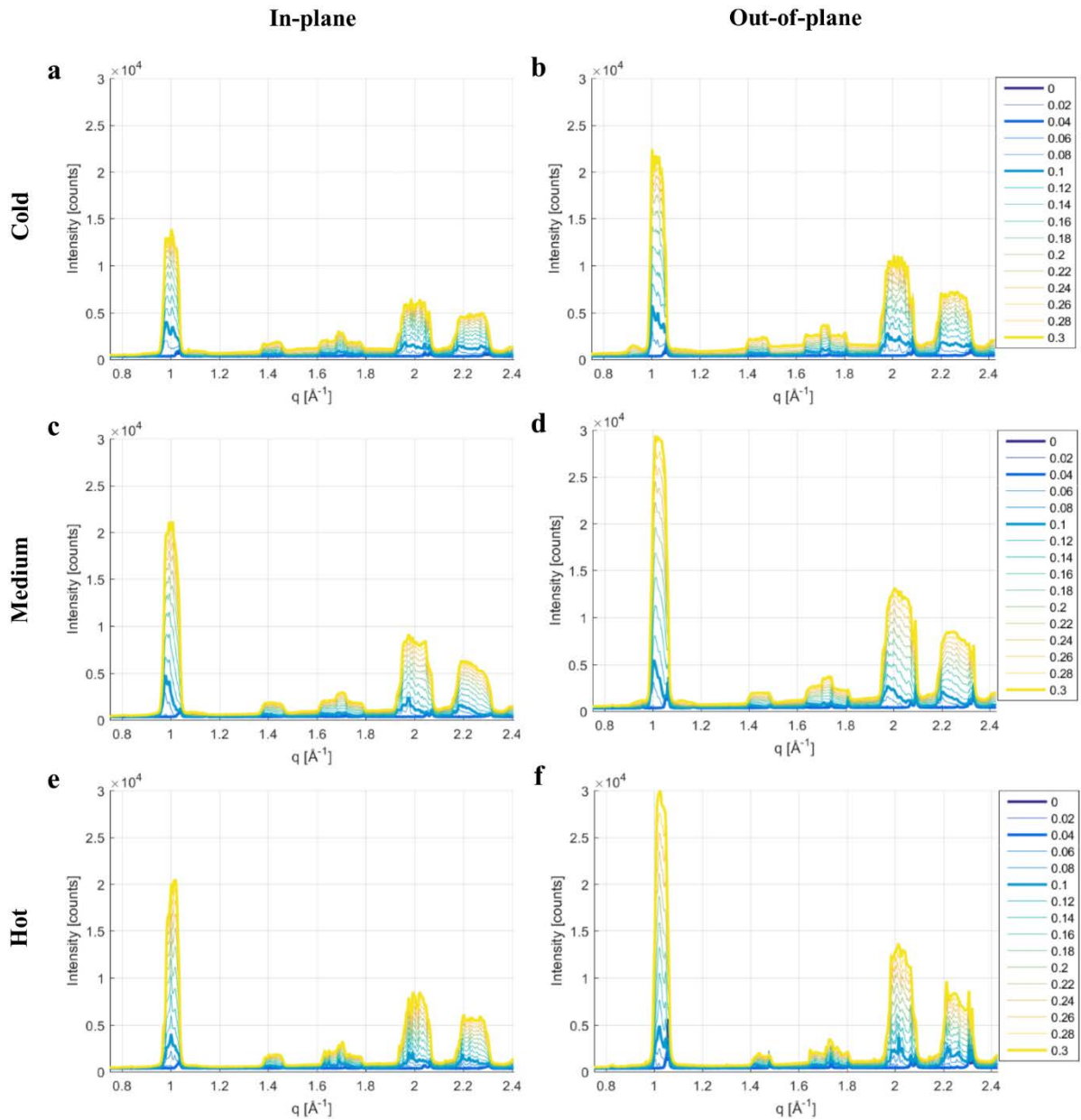


Figure S3. In-plane and Out-of-plane line profiles at different out-of-plane incident angles.

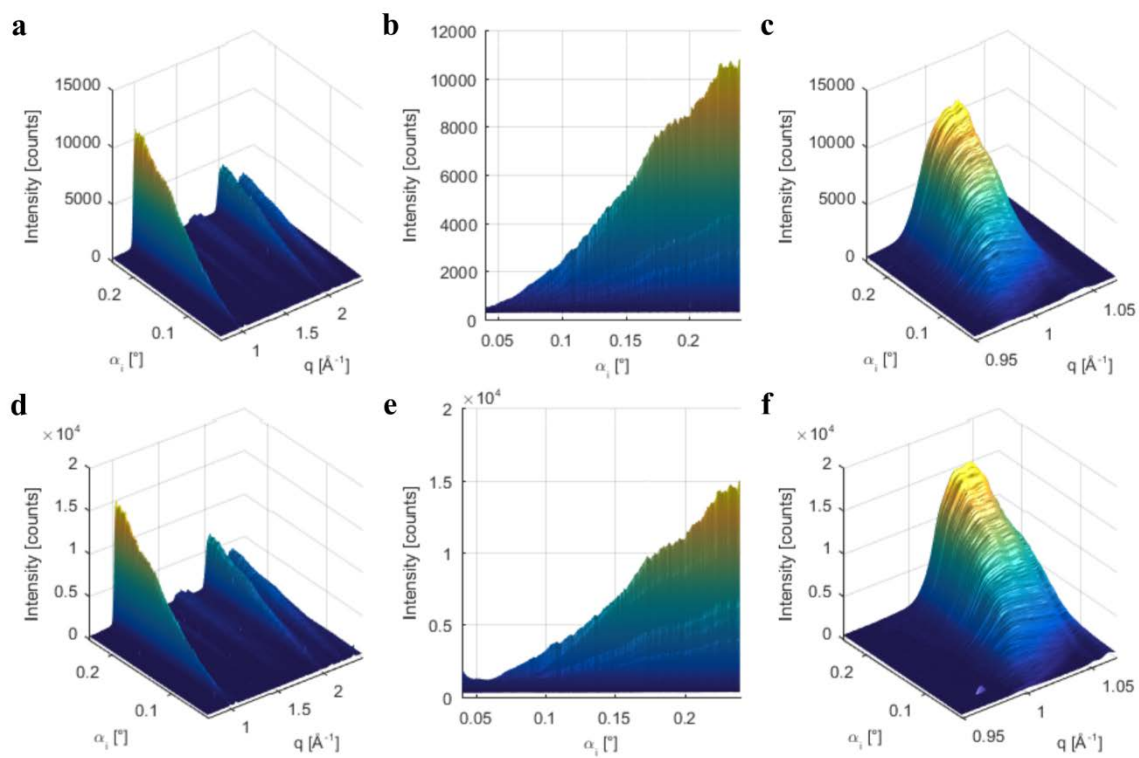


Figure S4. Fine α_i scan. **a, b, c** in-plane line profiles. **d, e, f** out-of-plane line profiles.

It is likely, however, that lattice strain and sample size effects or, equivalently, beam footprint effects (see **Figure S5**) play a significant role in peak broadening. Note that beam footprint effect could be minimized as suggested by Smilgies et al.^[8]. All these effects add up, making it difficult to quantify the lattice constants and the domain size of this class of polycrystalline materials using GI-WAXS. Note, however, that GI-WAXS is an extremely powerful technique when tracking relative changes (orientation and peak position) of individual diffraction spots in-situ during a thermal annealing.^[1, 9, 10]

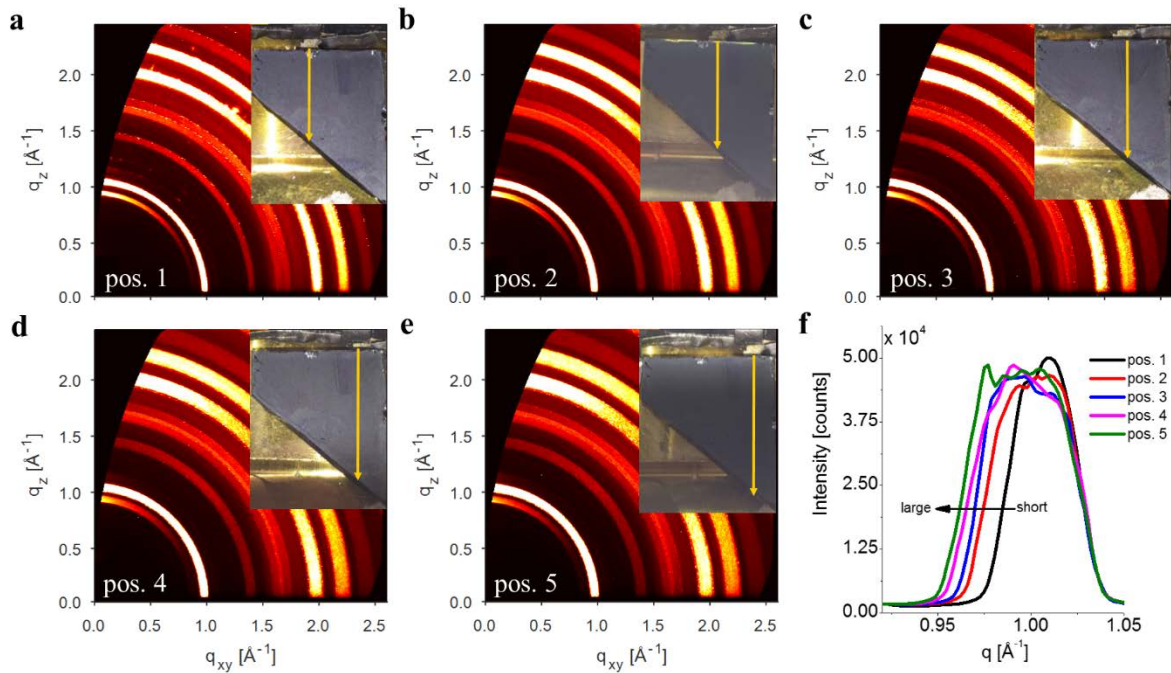


Figure S5. GI-WAXS diffraction patterns from a perovskite films (medium spun) deposited on a substrate with a trapezoidal shape and measured at different positions of the trapezoid. Positions are shown in the insets. **a**, position 1 (~1 cm, short). **b**, position 2 (~1.2 cm). **c**, position 3 (~1.3 cm). **d**, position 4 (~1.5 cm). **e**, position 5 (~1.6 cm, large). NOTE: These are estimated distances. **f**, in-plane line profiles ($\chi = 83^\circ$, $\Delta\chi = 7^\circ$) showing the main perovskite peaks at $q \approx 1 \text{\AA}^{-1}$ (as well as the others, not shown here) broaden as the probed sample length becomes longer. NOTE: the line profiles were slightly shifted in q to match their right edge with the line profile pos. 1. The FWHM for the line profiles at 0.3° are: 0.0408 (pos. 1), 0.0514 (pos. 2), 0.0557 (pos. 3), 0.0606 (pos. 4), and 0.0632 (pos. 5) \AA^{-1} . Therefore, the main perovskite peak peaks at $\alpha_i = 0.3^\circ$ are thin (FWHM = 0.0408\AA^{-1}) when a short section (~1 cm) is probed, and larger (FWHM = 0.0632\AA^{-1}) when a larger section is probed (~1.6 cm). This is due to the fact that different points hit by the X-ray beam on the sample have slightly different distance between the sample and the detector.

3 nXRD Measurements and Data Analysis

During data collection, samples were scanned within the y - z laboratory plane (with the X-ray beam along the x -axis), here indicated as x - y . A picture of the beamline configuration used during the nanofocus raster scan is shown in **Figure S6**. Each scan consisted of a $40 \times 40 \mu\text{m}^2$ raster scan of 101×101 pixels, with a step size of 400 nm. A single scan resulted in 10201 2167×2070 pixel diffraction images, for a total of ~ 170 GB of data saved in the ESRF Data Format (EDF).

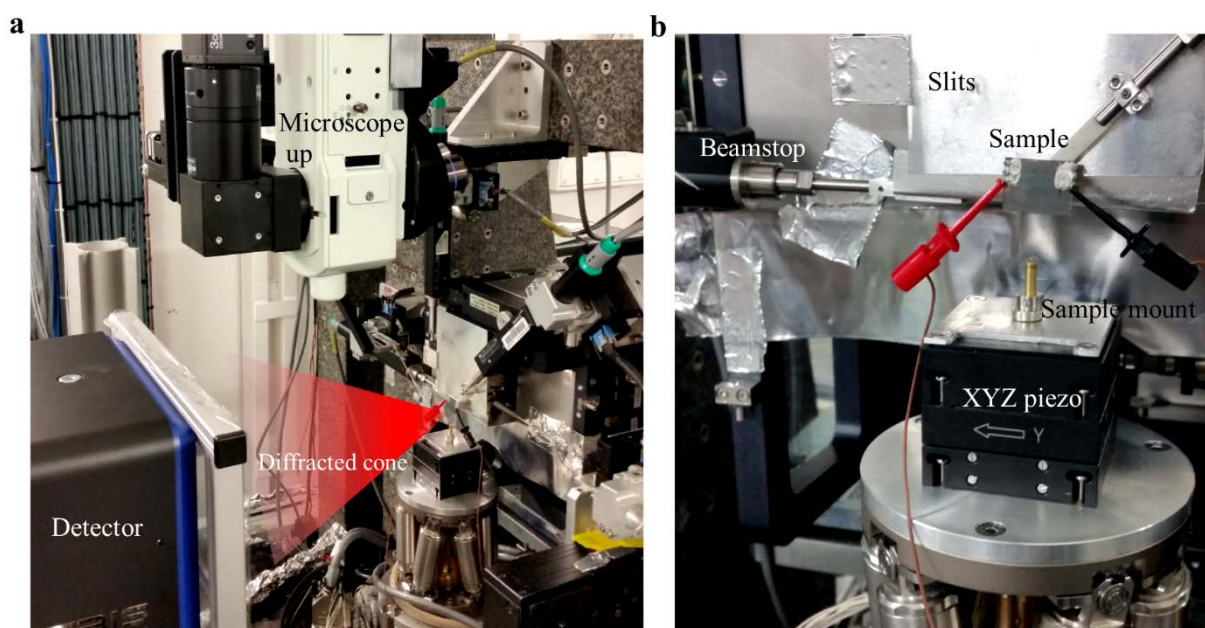


Figure S6. ID13 (ESRF) beamline setup during the raster scan. a, The photo shows the optical microscope position during the XRD measurements, the detector, and the sample positioned on the xyz piezo scanning stages. **b,** Close-up photo showing the slits, the beamstop, the sample mounted on the sample through a microscope slide, and the xyz piezo scanning stages.

Data Analysis was performed with a MATLAB^[11] application specifically designed by S. Lilliu for nXRD measurements (a snapshot of the software is shown in **Figure S7**). We used a custom NNB-XB0057 (NOVATECH, UK) notebook with a quad-core Intel® Core™ i7-4790K CPU 4GHz, 32GB of RAM, and Windows 7 (64bit).

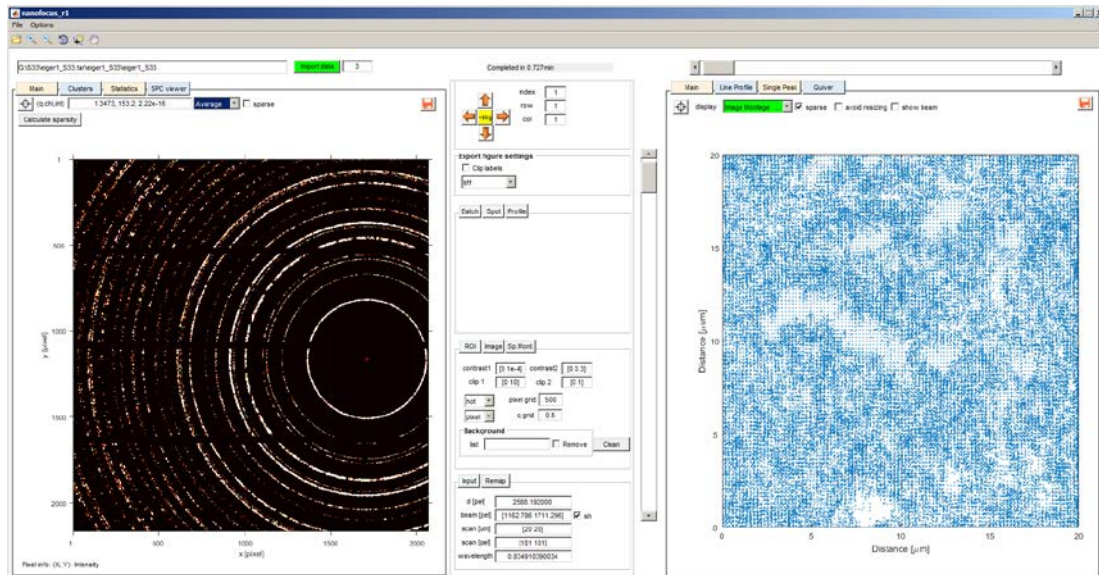


Figure S7. Snapshot of the MATLAB application for nanofocus data processing developed by S. Lilliu. This snapshot shows only two of the many available panels. The panels on the left displays the average diffraction pattern and is also used to display single diffraction patterns. The panel on the right displays the sparse pattern of the montage. The controls at the centre are used to navigate the 10201 single diffraction patterns across the scan area, to adjust the figures, to control data processing, and to input the calibration parameters, as discussed in the following paragraphs.

3.1 Spatial Resolution

The size of the focused X-ray beam was determined by scanning two crossed gold wires (diameter 50 μm). The gold wires act as a knife edge, meaning that the derivative of the absorption profiles corresponds to the profile of the X-ray beam in the horizontal (y , **Figure S8a**) and vertical (z , **Figure S8c**) directions. The absorption scans were smoothed with a moving average smoothing filter (black curves in **Figure S8a** and **c**) to remove noise due to variations in the incoming X-ray flux and vibration of the gold wires. The derivatives of the smoothed absorption scans (black squares in **Figure S8b** and **d**) reveal a well-defined beam profile. The nominal beam size of $200 \times 200 \text{ nm}$ was estimated from the full-width at half-maximum (FWHM) of the derivative of the smoothed absorption scans. Three methods were used to extract the FWHM: (i) manual assessment of the extent of the curve at half-maximum, (ii) Single Gaussian fitting (black curve in **Figure S8b** and **d**) and (iii) Double Gaussian fitting (blue curve in **Figure S8b** and **d**). The fit parameters and the beam profile data are summarised in **Table S1** and **Table S2**, respectively.

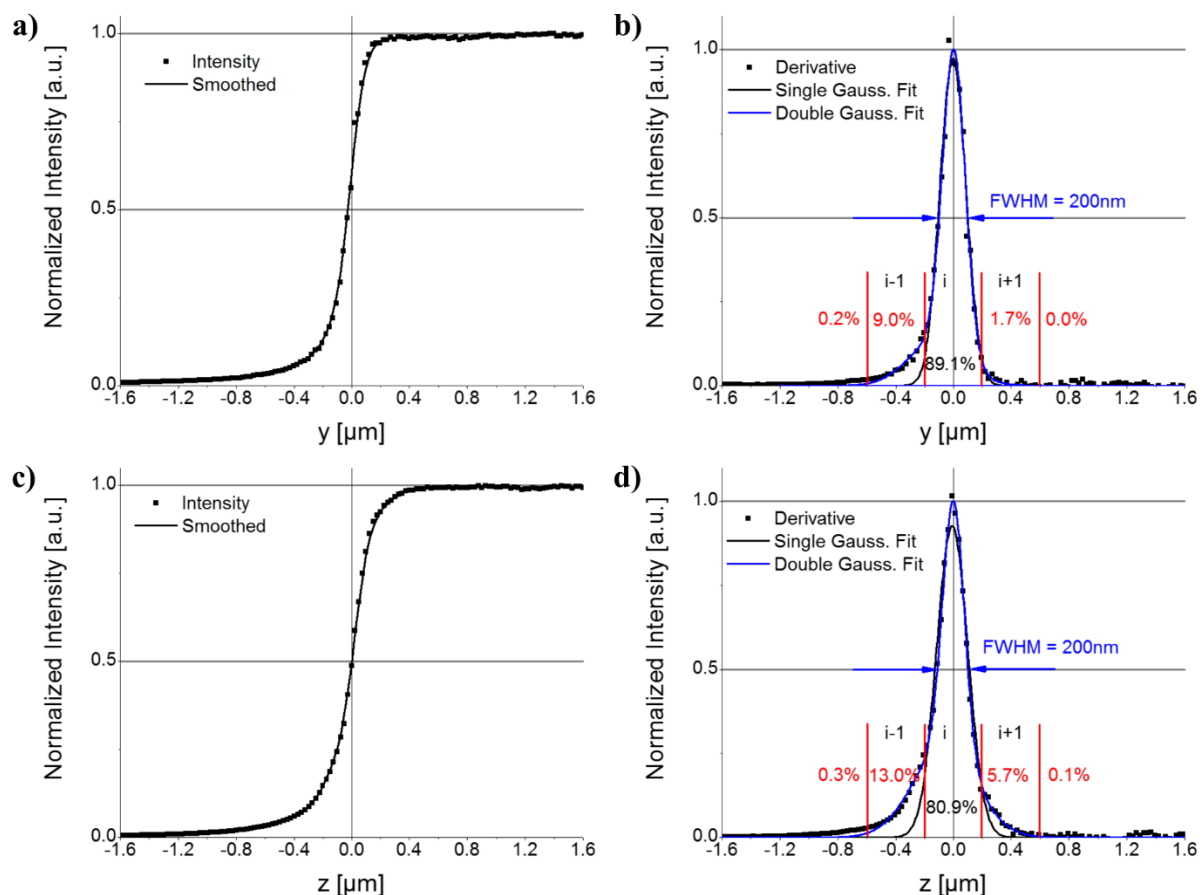


Figure S8. Beam Profiling Data. **a** and **c**, Absorption scans (black squares) and smoothed data (black lines) for the horizontal and vertical beam profiles, respectively. **b** and **d**, derivative of the smoothed absorption data (black squares), single Gaussian fits (black lines), and double Gaussian fits (blue lines). The estimated FWHM is highlighted with two blue arrows. The indices $i-1$, i , $i+1$ indicate three points of the raster scan. The percentage values indicate the percentage of the illuminated area in a specific region.

A scan step size of 400nm was used to avoid strong scattering from the beam tails. This is illustrated in **Figure S8b** and **d**. The red bars in the figure illustrate the boundaries of the probed sample region in a specific point of the scan. In the horizontal direction $\sim 89\%$ of the beam illuminates the region indicated as 'i'. The left beam tail extends for $\sim 9\%$ of the total beam intensity in the region indicated as 'i-1', while the right beam tail extends for $\sim 2\%$ of the total

beam in the region indicated as ‘i+1’. In the vertical direction ~81% of the beam illuminates the region indicated as ‘i’. The left beam tail extends for ~13% of the total beam intensity in the region indicated as ‘i-1’, while the right beam tail extends for ~6% of the total beam in the region indicated as ‘i+1’.

Table S1. Fit parameters for the Single and Double Gaussian fit of the horizontal (Y) and vertical (Z) beam profiles. For a description of the figures of merit refer to the Curve Fitting Toolbox (MATLAB).

	Single Gaussian Fit	Double Gaussian Fit
Y	General model Gauss1: $f(x) = a1 * \exp(-((x-b1)/c1)^2)$ Coefficients (with 95% confidence bounds): a1 = 0.9724 (0.9438, 1.001) b1 = -0.001223 (-0.004273, 0.001826) c1 = 0.1271 (0.1228, 0.1314) Goodness of fit: SSE: 0.1395 R-square: 0.9751 Adjusted R-square: 0.9747 RMSE: 0.02981	General model Gauss2: $f(x) = a1 * \exp(-((x-b1)/c1)^2) + a2 * \exp(-((x-b2)/c2)^2)$ Coefficients (with 95% confidence bounds): a1 = 0.8901 (0.8401, 0.94) b1 = 0.005259 (0.002089, 0.008429) c1 = 0.1103 (0.1043, 0.1164) a2 = 0.1335 (0.09024, 0.1767) b2 = -0.1164 (-0.1724, -0.06035) c2 = 0.2757 (0.2315, 0.3198) Goodness of fit: SSE: 0.07814 R-square: 0.986 Adjusted R-square: 0.9856 RMSE: 0.02253
Z	General model Gauss1: $f(x) = a1 * \exp(-((x-b1)/c1)^2)$ Coefficients (with 95% confidence bounds): a1 = 0.9259 (0.8917, 0.9601) b1 = -0.00658 (-0.01118, -0.001978) c1 = 0.1526 (0.1461, 0.1591) Goodness of fit: SSE: 0.2399 R-square: 0.9599 Adjusted R-square: 0.9594 RMSE: 0.03909	General model Gauss2: $f(x) = a1 * \exp(-((x-b1)/c1)^2) + a2 * \exp(-((x-b2)/c2)^2)$ Coefficients (with 95% confidence bounds): a1 = 0.7451 (0.6978, 0.7923) b1 = 0.002763 (-0.0005345, 0.00606) c1 = 0.1054 (0.09866, 0.1121) a2 = 0.2663 (0.2202, 0.3124) b2 = -0.06287 (-0.08207, -0.04367) c2 = 0.3076 (0.279, 0.3363) Goodness of fit: SSE: 0.07907 R-square: 0.9868 Adjusted R-square: 0.9864 RMSE: 0.02266

Table S2. Summary of the beam profile FWHM.

	Derivative	Single Gaussian Fit	Double Gaussian Fit
Y	188	212	198
Z	212	254	213

3.2 Noise Removal and Average Diffraction Pattern

Single diffraction patterns were affected by noise in the form of 1-2 pixel counts randomly distributed in the image (**Figure S9a**). Once these pixels were set to zero, single patterns consisted of highly sparse matrices^[12] (**Figure S9b**). As an example, the average sparsity or density^[12] of the 10201 diffraction images of the cold scan was $(1.9 \pm 1.8) \times 10^{-3}\%$. We found that the fastest way to read a single diffraction pattern was to create a memory-map^[11] of each EDF file and then convert the mapped file into a double matrix variable. This matrix was then converted into a sparse matrix format,^[11] in which data were represented as list of three elements: (row, column, counts or, equivalently, intensity). Representing matrices in the sparse format saved significant amounts of storage space and resulted in faster data processing. The time taken for data conversion was mainly limited by the low data speed transfer between the external hard drive and the processor (~ 50 MBps). Therefore, parallel processing did not result in any improvement. The conversion process, where the 10201 image were saved in a struct^[11] variable, took approximately 1 hour. This step compressed data $\sim 2.4 \times 10^4$ times (from ~ 170 GB to ~ 7 MB).

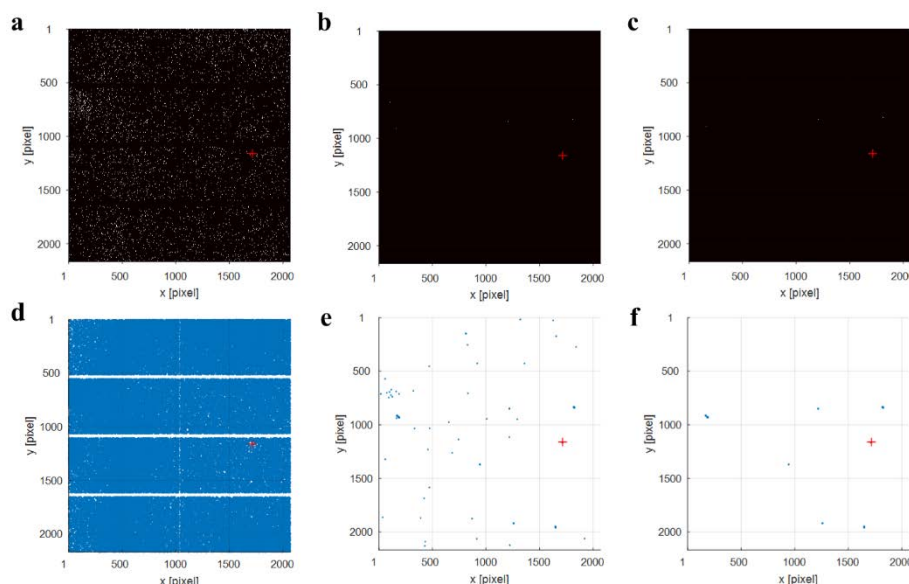


Figure S9. Noise removal from a single diffraction pattern (cold sample, first image). **a**, Diffraction pattern image as imported from the EDF file and displayed as a grayscale image. **d**, Sparsity pattern of image in **a**. **b**, Image after setting pixel with counts equal to 1 and 2 to zero. **e**, Sparsity pattern of image in **b**. **c**, Image after removing pixels surrounded by zero pixels ('clean' filtering). **f**, Sparsity pattern of **c**. The direct beam is shown as a red cross. Image coordinates are in pixels. For the grayscale images displayed here data was clipped above 10 counts, normalized between 0 and 1 and converted into grayscale format, and the contrast was adjusted between 0 and $1e-4$. The 'clean' filtering consisted in converting the sparse image into a full image and then applying the bwrmorph MATLAB function with the 'clean' option.

The average diffraction pattern of the 10201 images was calculated (in ~ 10 s) as the sum of the diffraction patterns divided by the number of diffraction images in a scan:

$$I_{av} = \frac{(\sum_{i=1}^N I_n)}{N} \quad (2)$$

where I_n is the matrix corresponding to the n -th image in a scan and N is the total number of images in the raster scan. From the average diffraction pattern shown in **Figure S10a** we observed some residual noise and broad diffused scattering from the substrate (Si/SiO₂). We

therefore reprocessed all the diffraction patterns by applying a ‘clean’ filter, which set to zero all the pixels surrounded by zeroes. The result for a single diffraction pattern is shown in Figure S9c. A comparison between the average diffraction pattern before and after the application of the filter is shown in Figure S10.

We note that this procedure is more effective than simply subtracting diffraction patterns measured from a Si/SiO₂. Reference samples are somewhat difficult to achieve for scanning nXRD. In fact, for a true reference one would have to measure the exact substrate and remount in the exact same way after sample preparation, which is practically impossible for the silicon substrates used here. For our case, we are also selecting a very limited range of q for our analysis, a q range which does not present scattering from silicon. Moreover, the broad Si/SiO₂ peak in Figure S10 is only visible in the average diffraction pattern. In the single scans we observe randomly scattered pixels around this peak, which can be easily removed with our de-noising procedure.

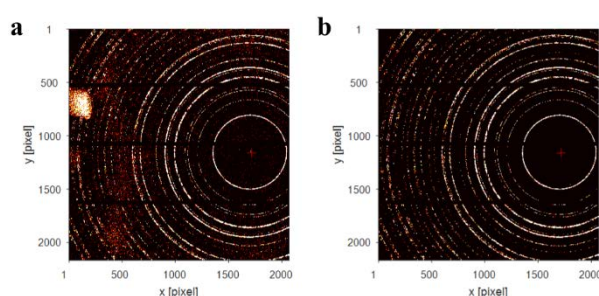


Figure S10. Average diffraction pattern (cold sample) before (a) and after (b) ‘clean’ filtering. Note that the broad scattering from the substrate is almost completely removed after filtering.

Figure S11 shows an overlay between the sparse matrix of the average diffraction pattern (in red colour) and the first diffraction pattern of the cold sample (a, in blue colour), or an average diffraction pattern performed on the first $N = 10, 100, 1000,$ and 1440 patterns (b, c, d, e, f, in blue colour, respectively), through equation (2). An animation is available at the following link: <https://youtu.be/hFTLnVBRurY>.

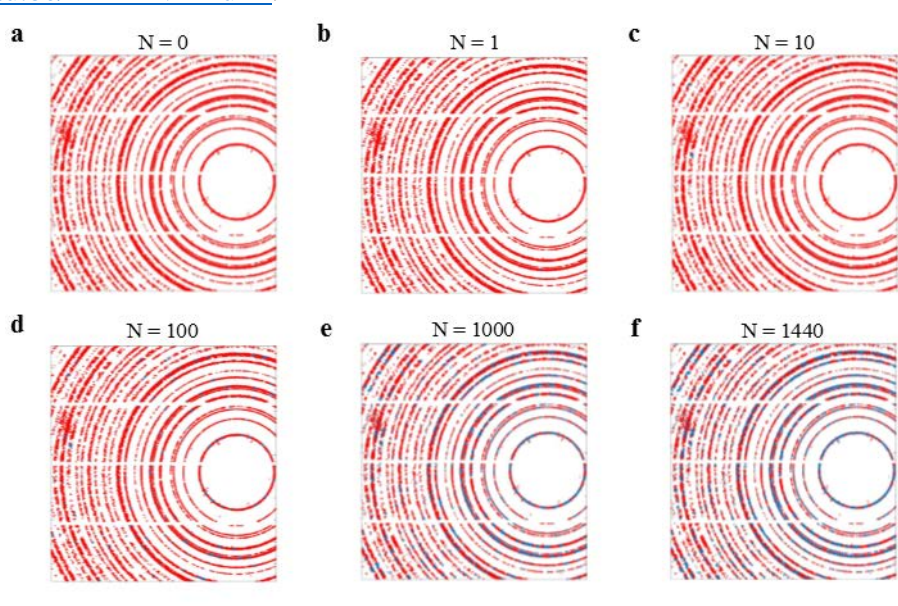


Figure S11. Overlay between the average diffraction pattern constructed using all the images in a scan ($N = 10\ 201$) and the ones constructed using N images through eq. (2). The average diffraction pattern constructed using all the images in a scan ($10\ 201$) is shown in red in the 6 images. The average diffraction pattern constructed using the first N images is shown in blue colour. Data is displayed as sparse matrices for clarity. Data employed here is from the ‘cold’ sample. A full sequence from $N = 0$ to $N = 1440$ is available at <https://youtu.be/hFTLnVBRurY>.

When recording diffraction patterns in GI-WAXS using much larger beams, a larger number of grains are illuminated, with such grains having different relative separations from the detector. Here, single grains are probed in transmission with a nanofocussed beam at a fixed grain-detector separation. As a consequence, the Debye-Scherrer rings in the nXRD average pattern are not affected by sample-size broadening and are significantly sharper than the Debye-Scherrer rings measured in GI-WAXS geometry.

3.3 Montages

Scan ‘montages’ were constructed by concatenating the 10201 patterns of a single scan in a 101×101 frame consisting of a large 218867×209070 sparse matrix, as illustrated in the MATLAB code shown in **Table S3**. The files required for plotting the montages are available as supporting files as discussed at the end of this document.

Table S3. MATLAB code illustrating how the montage matrix was built. In the first section of the code a LOOKUP matrix containing the diffraction pattern indexes arranged in a 101×101 matrix is built. In the second section the MONTAGE matrix is built. DataSize is an array containing the size of a single diffraction pattern, as [sizeRow sizeCol] = [2167 2070]. The struct PATTERN{n} contains the 10201 full diffraction patterns represented as sparse matrices. The for loop goes from the first diffracton pattern (index 1) until the last (numel(data) = 10201). The MONTAGE sparse matrix has a final size of 218867 rows and 209070 columns.

```
% Lookup table
% [1 2 3 4 ... 101;
% [102 103 104 ... 202;
% [...
% [10101 10102 10203 ... 10201]
count = 0;
for i = 1 : 101
    for j = 1 : 101
        count = count+1;
        LOOKUP(i,j) = count ;
    end
end

% Montage
DataSize = size(PATTERN{1}); S1 = DataSize(1); S2 = DataSize(2);

for n = 1 : numel(data)
    [row col] = find(LOOKUP == n);
    MONTAGE(row*S1-S1+1: row*S1, col*S2-S2+1: col*S2) = PATTERN{n};
end
```

Spatial coordinates were assigned to the centre of each diffraction pattern according to the raster scan. The concept is illustrated in **Figure S12** for 25 spatially adjacent diffraction patterns (cold sample). With a $40 \times 40 \mu\text{m}^2$ scan area, the 101×101 scan points split the scan area into $0.4 \times 0.4 \mu\text{m}^2$ tiles. To avoid negative values, we centred the first scan point (1,1) or tile at (0.2, 0.2) μm . Therefore, the last point of the scan is located at (40.2, 40.2) μm .

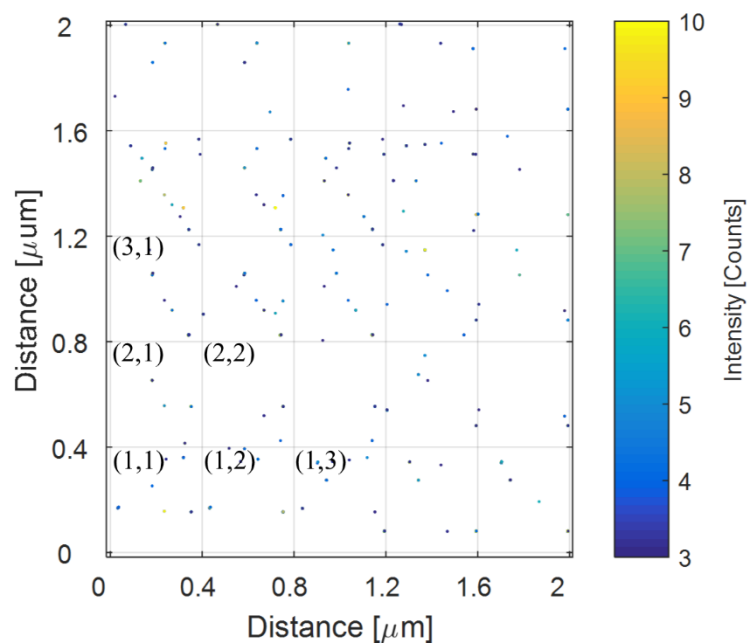


Figure S12. Montage: concept. Montage of the diffraction patterns from the cold sample scan including patterns with scan column and row index (indicated in figure) ranging from 1 to 5. The coordinates correspond to the relative point in space where the scan was performed. Single tiles, here indexed as (i,j) correspond to a single diffraction pattern, where i is the row scan index and j is the column scan index. NOTE: (i, j) are not Miller indices.

3.4 Remapping: from Pixel to Scattering Vector Coordinates

A diffracting sample projects a cone of rays (Debye-Scherrer rings^[7]) onto the detector. The projection of this cone onto a flat detector perpendicular to the direct beam is a circle. However, for various reasons, the detector might not be perfectly perpendicular to the direct beam. As a result, the rings appear on the acquired image as tilted ellipses, rather than circles.^[3] The ‘calibration’ step consisted in extracting the parameters required for correcting this distortion and converting the row and column image coordinates into scattering vector coordinates.^[3] The geometry employed in the calibration is described in ref.^[13]. Diffraction patterns from the α -Al₂O₃ were measured before the experiments. With the *pyFAI-calib* library^[14-16] we extracted the distance between the sample and the detector, 3 rotation angles (rot1, rot2, rot3), and the points of normal incidence (PONI),^[13] which were used in the conversion of the diffraction image coordinates from rows and columns to scattering vector polar coordinates, as discussed in the next section. The extracted parameters are reported in **Table S4**.

Table S4. Calibration Parameters extracted with the pyFAI-calib library^[17].

	in pixels	in meters	in radians
sample-detector distance	2588.760886	0.194157066	-
PONI (row)	1153.703231	0.086527742	-
PONI (column)	1742.352078	0.130676406	-
rot1	-	-	0.012397374
rot2	-	-	0.003393846
rot3	-	-	0

These parameters were used for converting row and column image coordinates into scattering vector coordinates in the following way (‘remapping’ step). We first calculated the detector rotation matrix \mathbf{R}

$$\mathbf{R}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \rho_1 & -\sin \rho_1 \\ 0 & \sin \rho_1 & \cos \rho_1 \end{bmatrix} \quad (3)$$

$$\mathbf{R}_2 = \begin{bmatrix} \cos \rho_2 & 0 & \sin \rho_2 \\ 0 & 1 & 0 \\ -\sin \rho_2 & 0 & \cos \rho_2 \end{bmatrix} \quad (4)$$

$$\mathbf{R} = \mathbf{R}_2 \mathbf{R}_1 \quad (5)$$

where $\rho_1 = -\text{rot1}$, $\rho_2 = -\text{rot2}$, rot1 and rot2 are the detector rotation angles extracted from the calibration (see Table S4). We then calculated:

$$p = \begin{bmatrix} c - \text{PONI}c \\ r - \text{PONI}r \\ L \end{bmatrix} \quad (6)$$

where c and r are the column and row coordinates of a non-zero pixel in the diffraction pattern, PONI c is the column coordinate of the point of normal incidence,^[15] and PONI r its row coordinate. The point p was rotated through the detector rotation matrix:

$$t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \mathbf{R}p \quad (7)$$

The Bragg angle in radians 2θ is:^[7, 13]

$$2\theta = \text{atan2}\left(\sqrt{t_1^2 + t_2^2}, t_3\right) \quad (8)$$

The norm of the scattering vector q is:^[7]

$$q = \frac{4\pi}{\lambda} \sin \frac{2\theta}{2} \quad (9)$$

And the azimuth in radians is:

$$\chi = \begin{cases} \chi' & \text{for } \chi' > 0 \\ \chi' + 2\pi & \text{for } \chi' < 0 \end{cases} \quad (10)$$

For easier data processing and visualization we used χ (instead of χ') in degrees (from 0° to 360°). Therefore, the ‘remapping’ step remapped a non-zero (row, column, intensity) triplet into a $(q, \chi, \text{intensity})$ triplet. In **Figure S13** we show the q, χ coordinates, or simply polar coordinates (distance and azimuth), for a diffraction spot.

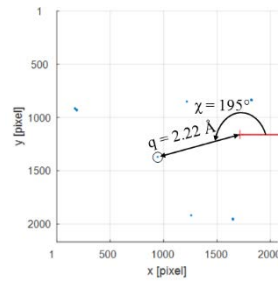


Figure S13. Representation of the scattering vector in polar coordinates (q, χ) .

The extraction of azimuthally integrated line profiles required two further steps: rebinning and averaging. The intensities of the $(q, \chi, \text{intensity})$ triplet were average-binned into a sparse matrix having a regular step size for q and χ ($\Delta q = 1 \times 10^{-3} \text{ \AA}^{-1}$ and $\Delta \chi = 0.5^\circ$). The choice of the step size had to take into account the q resolution. In fact, the size of pixels close to the direct beam ($q \approx 0 \text{ \AA}^{-1}$) is $\Delta q \approx 3 \times 10^{-3} \text{ \AA}^{-1}$, while the size of pixels close to the left edge of the diffraction pattern ($q \approx 4.3 \text{ \AA}^{-1}$) is $\Delta q \approx 2 \times 10^{-3} \text{ \AA}^{-1}$. Therefore, we can set the resolution to $\Delta q \approx (2.5 \pm 0.5) \times 10^{-3} \text{ \AA}^{-1}$. As a consequence, a step size of $\Delta q < 1 \times 10^{-3} \text{ \AA}^{-1}$ would result in oversampling and noisy data.

3.5 Azimuthally Integrated Line Profiles

In the binning step we collected all the elements from the triplet that fell into the (m,n)-th bin, i.e with q coordinates between $m\Delta q$ and $(m+1)\Delta q$ and χ coordinates between $n\Delta\chi$ and $(n+1)\Delta\chi$, and calculated the average of the intensity of these elements. The obtained ‘polar’ matrix $I(q,\chi)$ was then used for the calculation of the ‘average azimuthally integrated line profiles’:

$$i(q) = \sum_{\chi=0^{\circ}}^{360^{\circ}} \frac{I(q,\chi)}{N_q} \quad (11)$$

where, here, the azimuthal integration was performed from 0° to 360° , and N_q is the number of elements with the same q . In practice, the line profile was simply calculated as $\text{mean}(I)$ in MATLAB. Line profiles extracted from the average diffraction pattern of the cold, medium, and hot sample are shown in **Figure S14**.

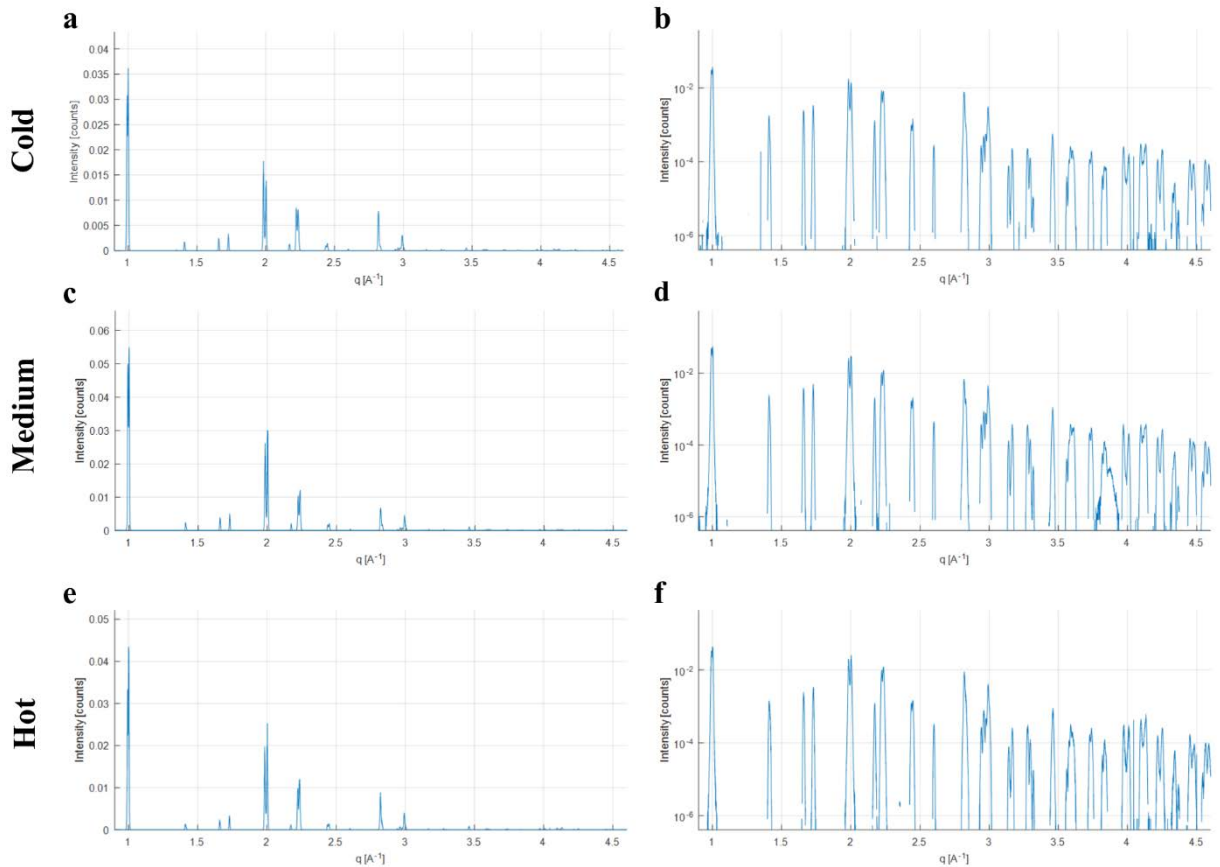


Figure S14. Azimuthally integrated line profiles in linear and logarithmic scale from the average diffraction patterns for the cold (a,b), medium (c,d), and hot sample (e,f).

With our MATLAB application it is possible to inspect single diffraction patterns in real time by simply selecting the image index, the column and row of the raster scan, or by simply pressing up/down or left/right arrows. The line profile extraction, which includes ‘remapping’ and ‘binning’ is performed in real time. The processing time for these two steps is ~ 1.2 ms per image (average value calculated on the 10201 images of the cold sample). Note that the processing speed is inversely proportional to the density of the sparse matrices.

Parallel processing or GPU processing^[11] was not an option for highly sparse dataset, as the time taken to dispatch/collect the job or to generate a GPU variable and retrieve results after processing was longer than the actual processing time.

3.5.1 Refinement

The line profile extracted from the average diffraction pattern of the three samples were converted into 2θ coordinates. The MAPI crystallographic file used for the Rietveld refinement^[18] was obtained from ref.^[19] The pattern was fitted with Pseudo-Voigt functions.^[5] Results are shown in **Figure S15**, **Table S5**, **Table S6**, **Table S7**, and **Table S8**.

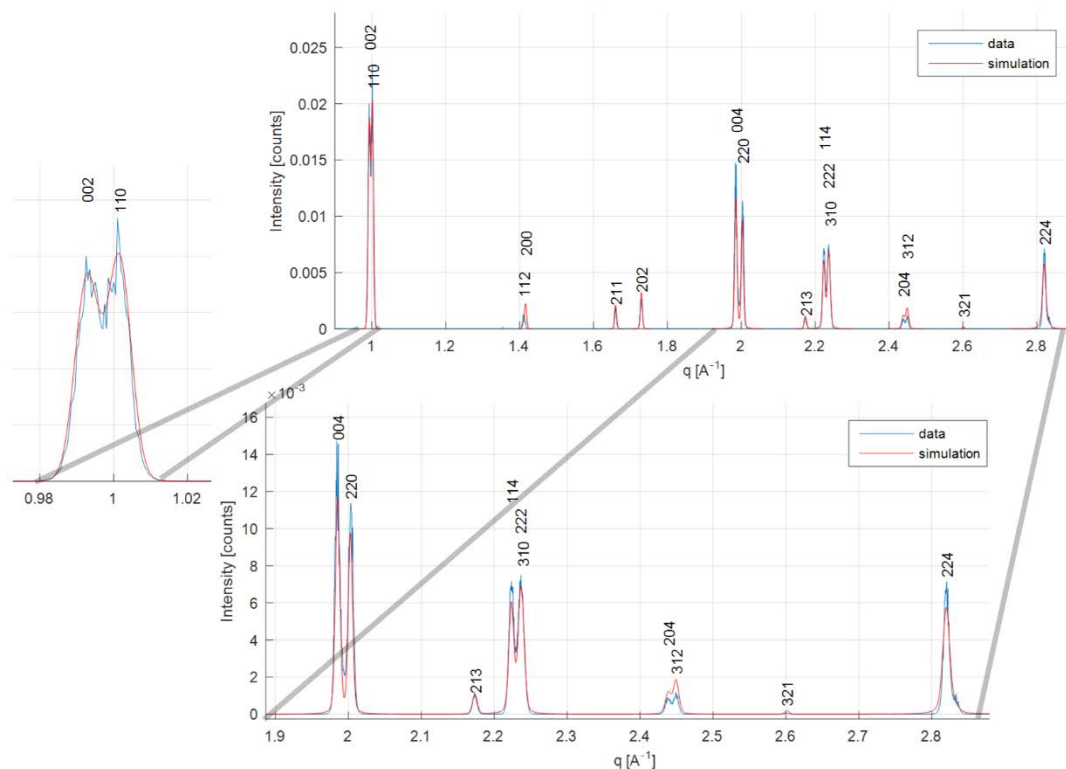


Figure S15. Comparison between measurements and simulated diffraction pattern from the refined tetragonal cell (cold sample).

Table S5. Cold sample, first 40 reflections from the simulated diffraction pattern obtained from the refined crystal.

q [\AA^{-1}]	relative intensity	h	k	l
0.992	97.331	0	0	2
1.001	100.000	1	1	0
1.409	9.329	1	1	2
1.416	8.647	2	0	0
1.659	16.267	2	1	1
1.729	16.372	2	0	2
1.984	62.237	0	0	4
2.002	66.359	2	2	0
2.172	8.169	2	1	3
2.222	50.505	1	1	4

q [\AA^{-1}]	relative intensity	h	k	l
2.234	28.760	2	2	2
2.238	40.464	3	1	0
2.437	8.509	2	0	4
2.448	13.966	3	1	2
2.600	1.651	3	2	1
2.819	53.241	2	2	4
2.831	17.564	4	0	0
2.942	2.798	2	1	5
2.954	1.073	3	2	3
2.960	8.060	4	1	1

q [\AA^{-1}]	relative intensity	h	k	l
2.976	3.942	0	0	6
2.991	43.671	3	1	4
3.000	8.756	4	0	2
3.003	2.321	3	3	0
3.140	1.188	1	1	6
3.162	0.058	3	3	2
3.165	3.639	4	2	0
3.276	5.433	4	1	3
3.296	0.710	2	0	6
3.317	0.078	4	2	2

q [\AA^{-1}]	relative intensity	h	k	l
3.457	18.893	4	0	4
3.559	0.492	3	2	5
3.573	1.816	4	3	1
3.587	6.088	2	2	6
3.599	3.731	3	3	4
3.609	4.249	5	1	0
3.724	3.630	3	1	6
3.736	6.068	4	2	4
3.743	1.187	5	1	2
3.816	0.792	2	1	7

Table S6. Medium sample, first 40 reflections from the simulated diffraction pattern obtained from the refined crystal.

q [Å ⁻¹]	relative intensity	h	k	l
0.992	97.295	0	0	2
1.001	100.000	1	1	0
1.409	9.326	1	1	2
1.416	8.647	2	0	0
1.659	16.266	2	1	1
1.729	16.368	2	0	2
1.984	62.211	0	0	4
2.002	66.358	2	2	0
2.173	8.167	2	1	3
2.223	50.485	1	1	4

q [Å ⁻¹]	relative intensity	h	k	l
2.234	28.756	2	2	2
2.238	40.463	3	1	0
2.438	8.505	2	0	4
2.448	13.964	3	1	2
2.600	1.651	3	2	1
2.819	53.224	2	2	4
2.831	17.563	4	0	0
2.942	2.797	2	1	5
2.954	1.072	3	2	3
2.960	8.059	4	1	1

q [Å ⁻¹]	relative intensity	h	k	l
2.977	3.940	0	0	6
2.991	43.659	3	1	4
3.000	8.755	4	0	2
3.003	2.321	3	3	0
3.140	1.187	1	1	6
3.163	0.058	3	3	2
3.165	3.639	4	2	0
3.276	5.432	4	1	3
3.296	0.709	2	0	6
3.317	0.078	4	2	2

q [Å ⁻¹]	relative intensity	h	k	l
3.457	18.888	4	0	4
3.559	0.492	3	2	5
3.574	1.816	4	3	1
3.587	6.085	2	2	6
3.599	3.730	3	3	4
3.609	4.249	5	1	0
3.724	3.628	3	1	6
3.736	6.066	4	2	4
3.743	1.187	5	1	2
3.816	0.792	2	1	7

Table S7. Hot sample, first 40 reflections from the simulated diffraction pattern obtained from the refined crystal.

q [Å ⁻¹]	relative intensity	h	k	l
0.993	97.550	0	0	2
1.001	100.000	1	1	0
1.410	9.336	1	1	2
1.416	8.646	2	0	0
1.659	16.268	2	1	1
1.729	16.376	2	0	2
1.986	62.362	0	0	4
2.002	66.354	2	2	0
2.173	8.174	2	1	3
2.224	50.575	1	1	4

q [Å ⁻¹]	relative intensity	h	k	l
2.235	28.767	2	2	2
2.239	40.462	3	1	0
2.439	8.517	2	0	4
2.449	13.967	3	1	2
2.600	1.651	3	2	1
2.820	53.274	2	2	4
2.832	17.562	4	0	0
2.944	2.801	2	1	5
2.955	1.073	3	2	3
2.961	8.059	4	1	1

q [Å ⁻¹]	relative intensity	h	k	l
2.979	3.949	0	0	6
2.992	43.695	3	1	4
3.001	8.756	4	0	2
3.003	2.321	3	3	0
3.142	1.190	1	1	6
3.163	0.058	3	3	2
3.166	3.639	4	2	0
3.277	5.434	4	1	3
3.298	0.710	2	0	6
3.318	0.078	4	2	2

q [Å ⁻¹]	relative intensity	h	k	l
3.458	18.898	4	0	4
3.560	0.493	3	2	5
3.574	1.816	4	3	1
3.589	6.094	2	2	6
3.600	3.732	3	3	4
3.610	4.248	5	1	0
3.726	3.633	3	1	6
3.737	6.069	4	2	4
3.744	1.187	5	1	2
3.819	0.793	2	1	7

Table S8. Refined Lattice Constants Parameters.

	a, b	c
Cold	8.877	12.666
Medium	8.877	12.665
Hot	8.876	12.657

3.6 Circular Region of Interest Around the Main Perovskite Double Peak

To process data within a certain q range we defined circular regions of interest (ROIs). **Figure S16** shows an example of ROIs including data from the (002) and the (110) reflections, from the average pattern. Here we used the average pattern just to display the concept of circular ROI, as single diffraction patterns would only display a few scattered diffraction spots. We first generated a binary mask by defining an annulus in a 2167×2070 (which is the same size of a single diffraction image) null matrix with a large radius $R = 355$ pixels and a small radius $r = 336$ pixels and assigning 1 to the pixels inside it. The binary mask was converted into a sparse matrix and multiplied point by point to the (i,j) -th diffraction pattern of the raster scan. This allowed us to consider only data from the (002) and (110) reflections.

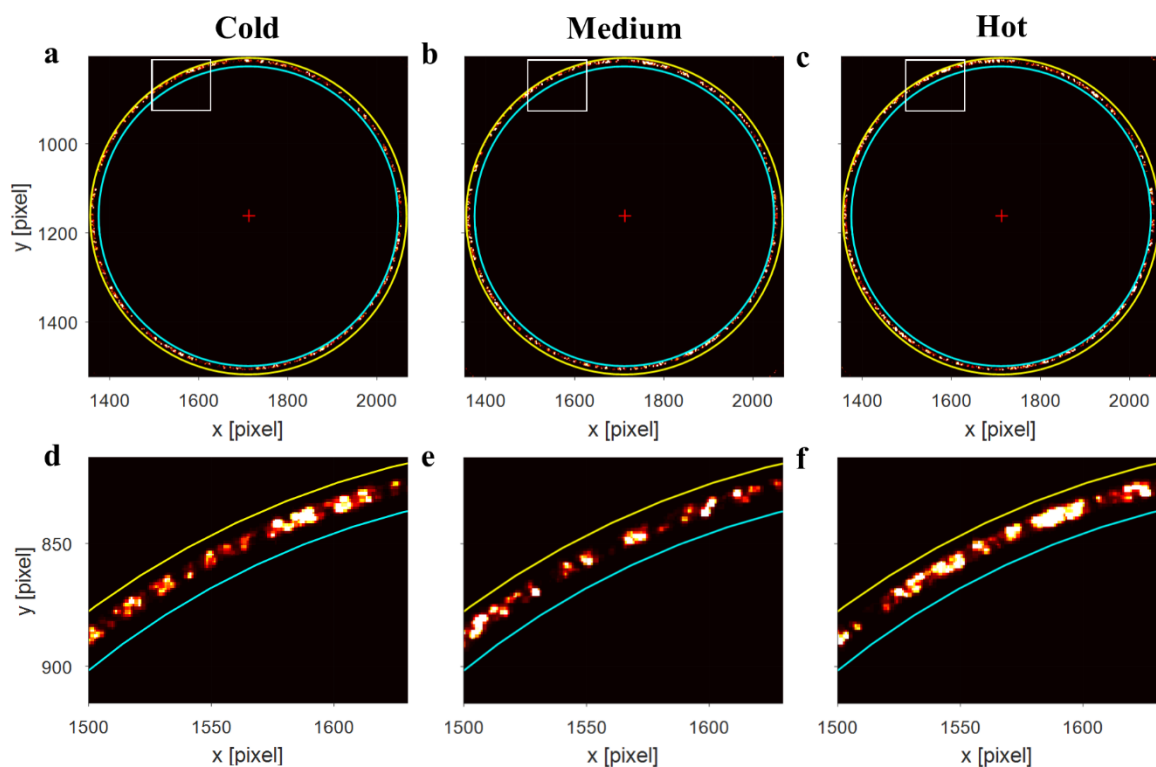


Figure S16. Circular regions of interest. Circular region of interest including the (002) and (110) reflections shown on the average diffraction pattern for the cold (a), medium (b), and hot sample (c). The selected region of interest included data from 336 to 355 pixels from the centre of the diffraction pattern, which is indicated with a red cross. For this images the column limit was set to [1353 2069] and the row limit was set to [804 1523]. For visualization purposes only data was clipped in the [0 10] counts range and the contrast was adjusted between 0 and 10^{-2} to highlight the most intense diffraction spots or diffraction spot clusters. The zooms (d, e, f) are taken the row limit set to [1500 1630] and the column limit set to [815 915] and the other parameters unchanged.

The analysis followed two conceptual steps: (i) diffraction spots clustering in the single diffraction patterns and extraction of their features including the peak position, (ii) spots clustering on the full dataset based on the spatial distance in the raster scan and on the spot orientation (χ).

3.7 First Clustering Step

We now discuss the first step. The following steps were repeated for each of the 10201 diffraction patterns in a scan. The row, column and intensity vectors were extracted from the sparse matrix representing the ROI described above in a single diffraction pattern. Rows and columns were remapped (see equations above) into q and χ coordinates. Row and column vectors defining the pixel positions in ROI of the diffraction pattern were clustered based on the pairwise Euclidean distance between the pixels.^[11] We used a threshold of 0.8 for cutting the hierarchical tree.^[11] As an example, the ROI of the first diffraction pattern of the scan displayed only one diffraction spot, i.e. only one cluster of 83 pixels. This is shown in **Figure S17a** in pixel coordinates.

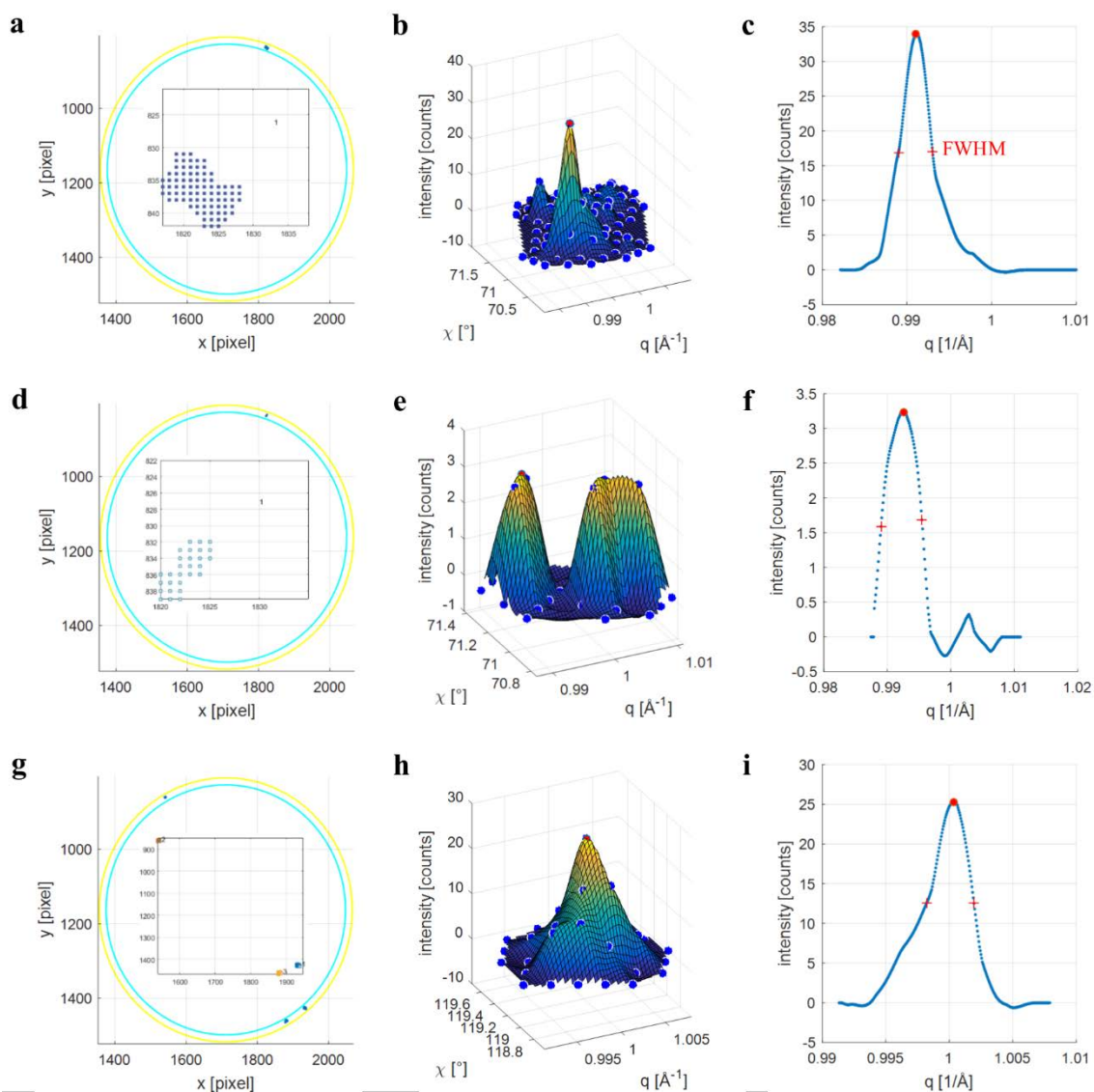


Figure S17. Examples of diffraction spot clustering (cold sample). **a**, sparse pattern in ROI with inset showing the clustered diffraction spot. In this case it is trivial as there is only one diffraction spot. **b**, intensity vs q and χ showing the original data (blue spots), the spline fitted data (mesh), and the peak position (red spot). **c**, line profile passing by the peak and the origin of the diffraction pattern, showing the peak position (intensity vs χ), the left hand side width at half maximum (LHS), and the right hand width at half maximum (RHS). The full width at half maximum (FWHM) is equal to LHS + RHS. **d**, **e**, **f** show similar plots for image index 3. **g**, **h**, **i** show similar plots for image index 1718.

The cluster represented as a set of q , χ and intensity coordinates (inset in Figure S17a) was fitted with a cubic spline interpolant using a mesh with a regular q , χ step size (Figure S17b). The maximum intensity was extracted and the corresponding q , χ were labelled as the peak position (q_p , χ_p). A line profile (Figure S17c) passing by the peak position and the origin of the diffraction pattern was extracted and used for the extraction of the left hand side width at half maximum (LHS), right hand width at half maximum (RHS), and the full width at half maximum (FWHM = LHS + RHS). This was the simplest case.

Since the ROI used for this analysis includes two reflections (002) and (110), we also observed diffraction spots with double peaks. Figure S17d (image index 3) shows a diffraction spot displaying both reflections in the ROI of the third diffraction pattern.

The most complicated case was when the ROI of a single diffraction pattern displayed multiple diffraction spots. This is shown in Figure S17g (image index 1718). The three diffraction spots were clustered as shown in the inset. Figure S17h and i show the previously described graphs for the second cluster.

3.7.1 Preliminary Quiver Plot

Once this procedure was iterated for the 10201 diffraction patterns, we generated a quiver^[11] (or velocity) plot representing the χ coordinate of each detected peak in spatial coordinates. As for the montage, the first diffraction pattern, which displayed only one diffraction spot, resulted in an arrow with origin at $(0.2, 0.2) \mu\text{m}$, an angle corresponding to χ , and a colour corresponding to a one to one correspondence between the angles from 0° to 360° and a jet colour map.^[11] The above mentioned ROI of image index 1718, which was made of 3 diffraction spots in the ROI, resulted in three arrows with angles equal to 310.1° , 119.21° , 299.81° . A quiver plot was employed as it allowed visualizing multiple peaks on the same spatial position. This is illustrated in **Figure S18**.

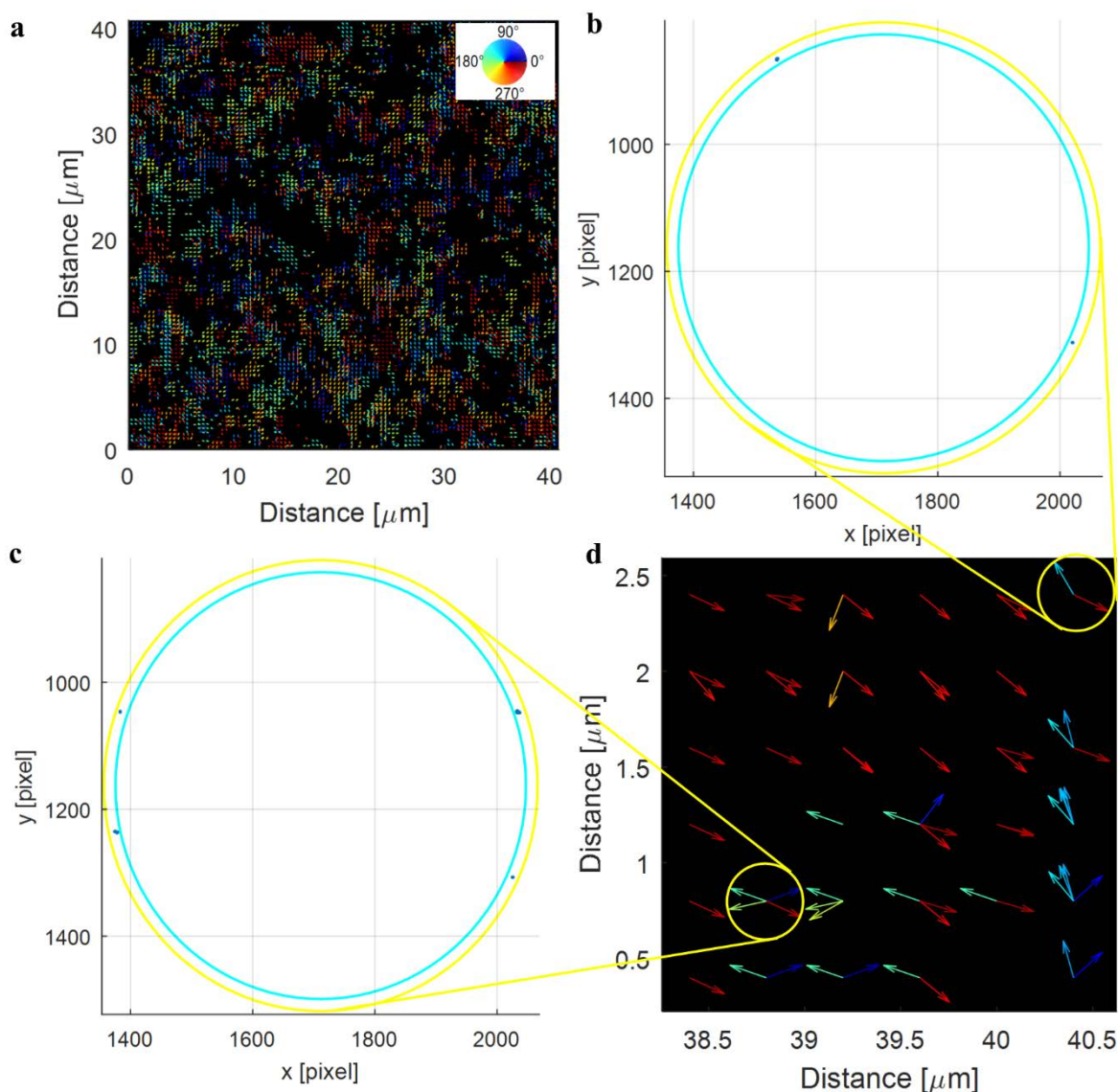


Figure S18. Quiver plot (cold sample). Representation of the χ coordinate from the peak extracted from diffraction spots in each image. **a**, quiver plot. **d**, zoom in the bottom right corner. **b**, and **c**, sparse pattern of the diffraction patterns (ROI) corresponding to two positions in the quiver point.

3.8 Second Clustering Step

We now discuss the second step. We performed a second level of clustering on (x, y, χ) , where χ is the χ coordinate of diffraction spot and x, y are the spatial coordinates of the point where the diffraction pattern containing the diffraction spot was taken. The purpose of the second level of clustering was to identify crystallites or grains based on their crystallographic orientation. The method used for this clustering is similar to the previous, with the difference that this one is performed in three dimensions and on different coordinates. **Figure S19** shows results for the first two largest detected clusters, along with maps of the q, χ coordinates of the peak extracted from the diffraction spots within the clustered grain.

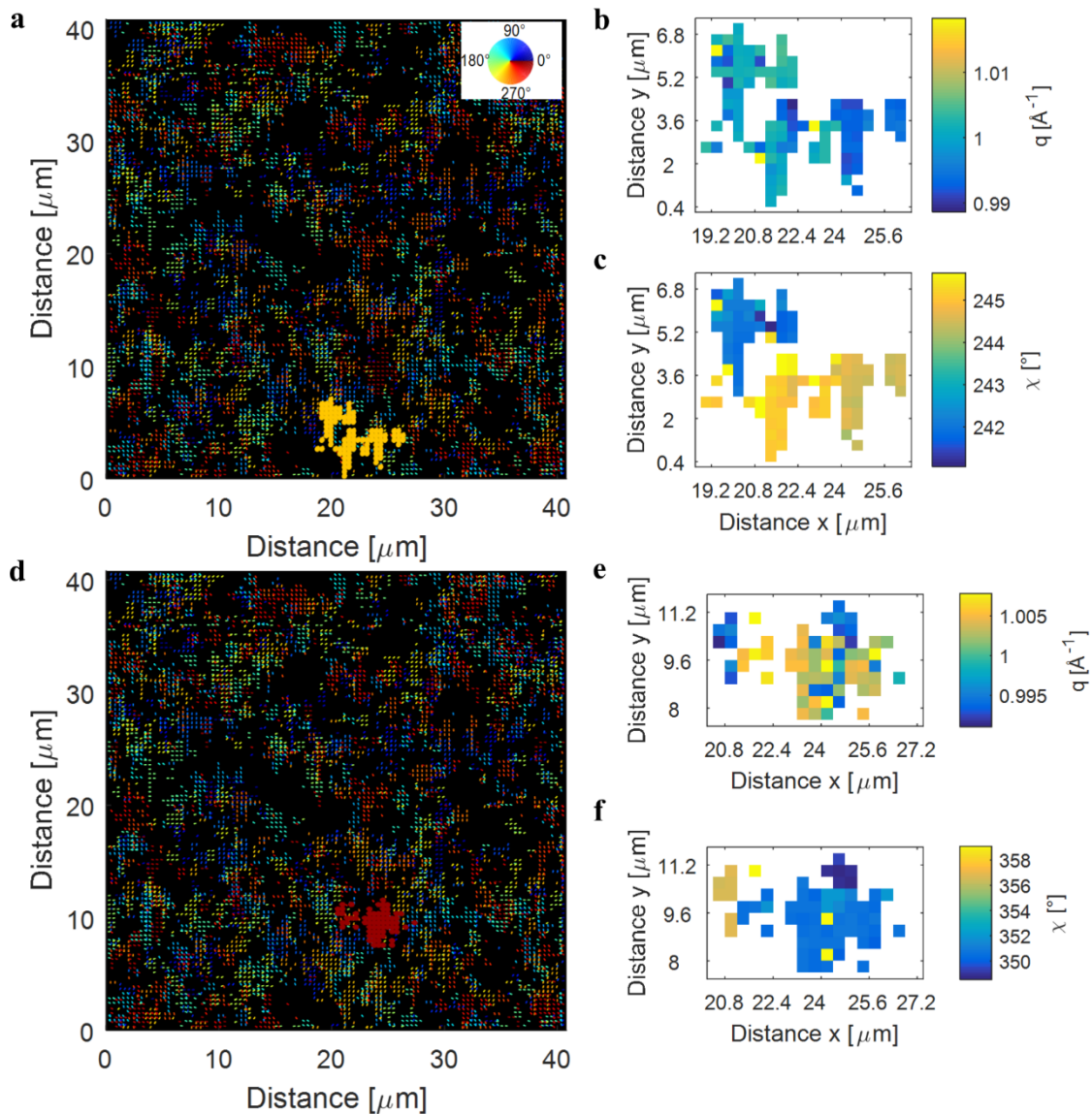


Figure S19. Second level of clustering: grains detection (cold sample). **a** and **d** show the first two largest clusters (grains) overlapped to the quiver plot. **b**, **e** show a map of the q coordinates of the peaks extracted from the diffraction spots belonging to the two largest grains. **c**, **f**, show a map of the χ coordinate.

In Figure S19b and c we can observe two regions corresponding to different q and χ values. The region on top clearly belongs to the (002) reflection, while the one on the bottom belongs to the (110) reflection. An accurate analysis should take into account that (002) and (110) are different peaks. Peaks segregation had to be implemented in simple and robust way. We could not perform this step by defining a finer ROI separating the (002) from (110) reflections because the ROI was defined on the raw pattern, which presented distortion due to detector misalignment (non-circular pattern).

Figure S20 shows a plot of q vs. χ for the peak extracted from all the diffraction spots in the raster scan. These values were obtained after applying the distortion correction equations described above.

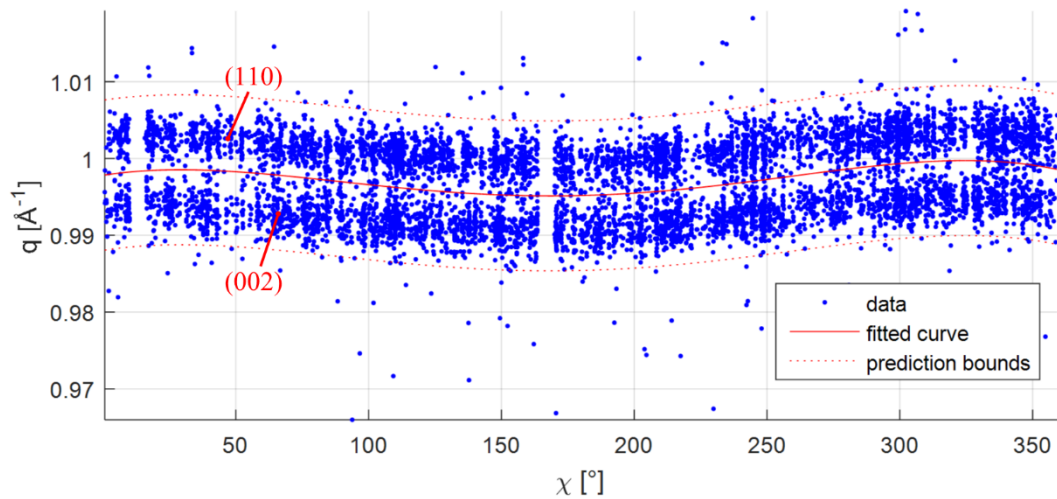


Figure S20. Residual distortion. Plot of q vs χ for the peak extracted from all the diffraction spots in the raster scan (cold sample).

The expected q vs. χ curve should have been flat with some fluctuations. However, we found that the graph was best fitted by a 4th order polynomial (red curve), which confirmed that the distortion caused by the rotated detector was not completely removed. This was so far the best distortion correction we could obtain. We observed that the fitting curve could be used to discriminate between the (002) and the (110) reflections.

Figure S21 shows the diffraction spot extracted from the 113th diffraction pattern (cold sample). The green line in Figure S21a is a plot of the 4th degree polynomial (shifted up for clarity), which passes between the two diffraction peaks. If we opt for the detection of the peak below this line, i.e. (002), we find $q = 0.9915 \text{ \AA}^{-1}$ and $\chi = 228.14^\circ$ as the detected peak. The line profile passing by $\chi = 228.14^\circ$ is shown in Figure S18b. If we opt for the detection of peaks above this line, i.e. (110), we find $q = 1.0003 \text{ \AA}^{-1}$ and $\chi = 227.51^\circ$ as the detected peak. The line profile passing by $\chi = 227.51^\circ$ is shown in Figure S21c. For both line profiles we fit the detected peak with a double Gaussian. Double Gaussians were used to take into account peak asymmetries, approximate quite well pseudo-Voigt profiles, and are easier and more robust to fit than pseudo-Voigt functions. Finally, we repeated the steps illustrated here twice for the ROIs corresponding to the 10201 images in a scan. The first time we focused on the (002) reflection, and the second time we focused on the (110) reflection.

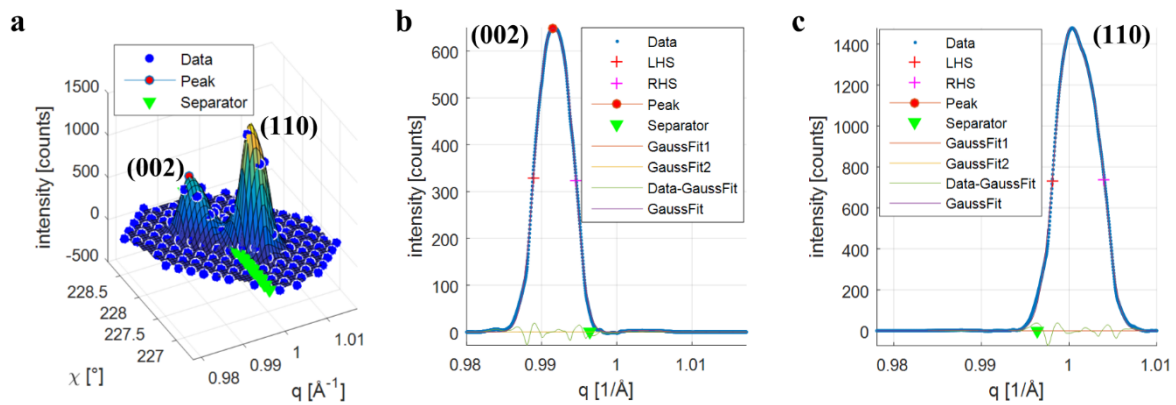


Figure S21. (002) and (110) discrimination (diffraction pattern 114, cluster 3, cold sample). **a**, Diffraction spot showing the (002) and the (110) reflections. The 4th degree polynomial passing by the two peaks is shown in green (Separator). Data is shown as blue ‘o’ and the mesh represents the 3D cubic spline interpolant. **b**, (002) peak detection. **c**, (110) peak detection. In b and c we also show the Gaussian fits and the error.

3.9 Maps

In **Figure S22**, **Figure S23**, and **Figure S24** we show maps of the number of pixels in a diffraction spot as discussed in the first clustering step (a, b), the q position of the peak (c, d), the corresponding lattice constant (e, f), the χ position of the peak (g, h), the peak intensity (i, j), and the FWHM (k, l), for the (002) and (110) largest grains in the cold, medium, and hot sample respectively.

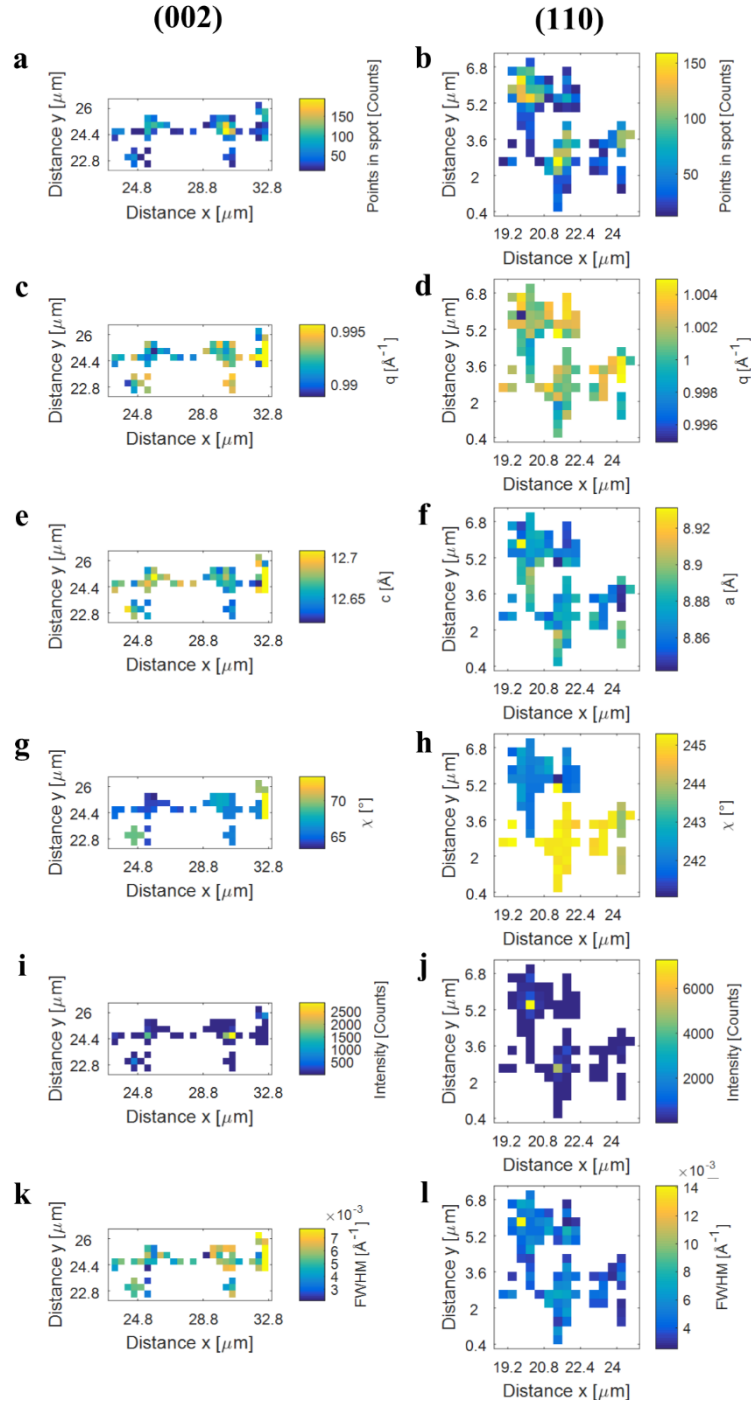


Figure S22. Spatial maps of diffraction spot peak properties in the largest (002) and (110) grains (cold sample). **a, b**, number of non-zero pixels in the diffraction spot detected in the first level of clustering. **c, d**, diffraction spot q_p peak position. **e, f**, lattice constant calculated from q_p . **g, h**, diffraction spot χ_p peak position. **i, j**, diffraction spot peak intensity. **k, l**, FWHM extracted from the line profile passing by χ_p and q_p and the origin of the diffraction pattern.

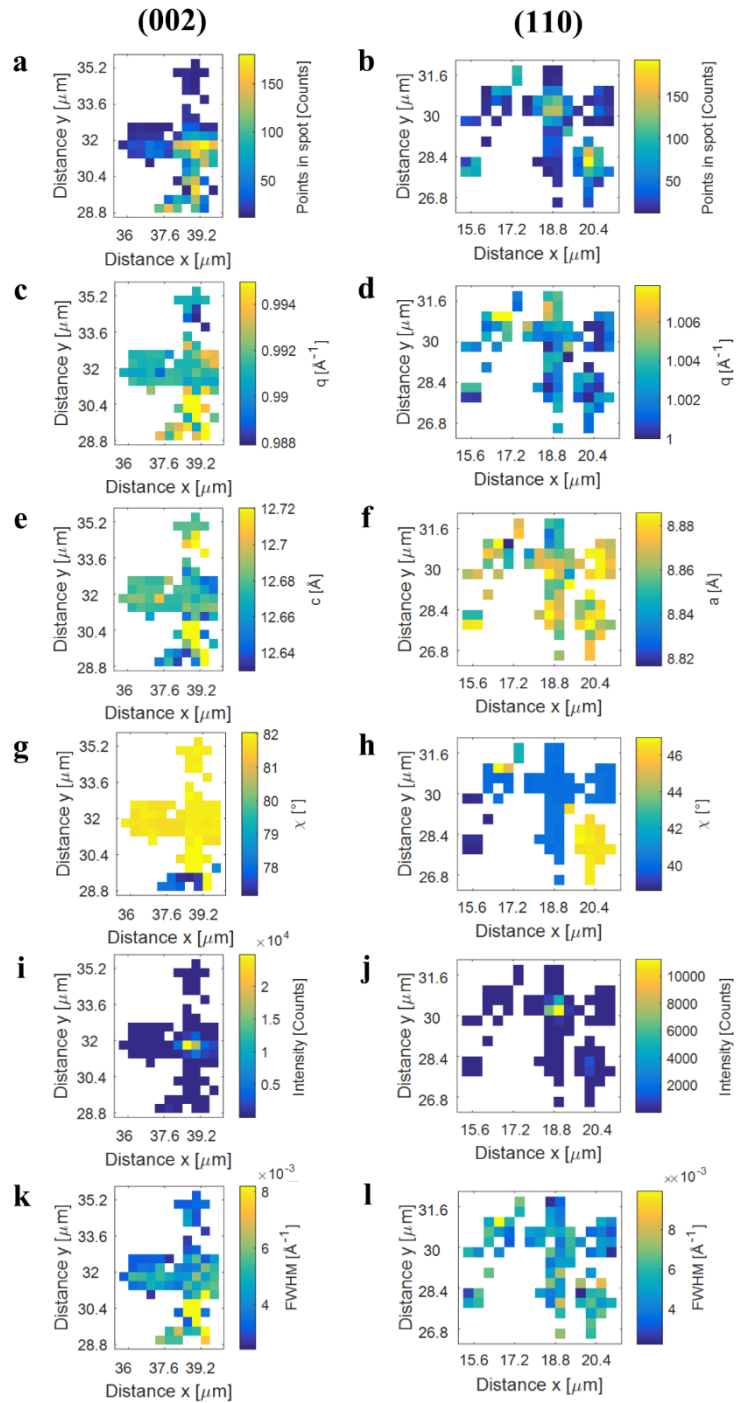


Figure S23. Spatial maps of diffraction spot peak properties in the largest (002) and (110) grains (medium sample). **a, b**, number of non-zero pixels in the diffraction spot detected in the first level of clustering. **c, d**, diffraction spot q_p peak position. **e, f**, lattice constant calculated from q_p . **g, h**, diffraction spot χ_p peak position. **i, j**, diffraction spot peak intensity. **k, l**, FWHM extracted from the line profile passing by χ_p and q_p and the origin of the diffraction pattern.

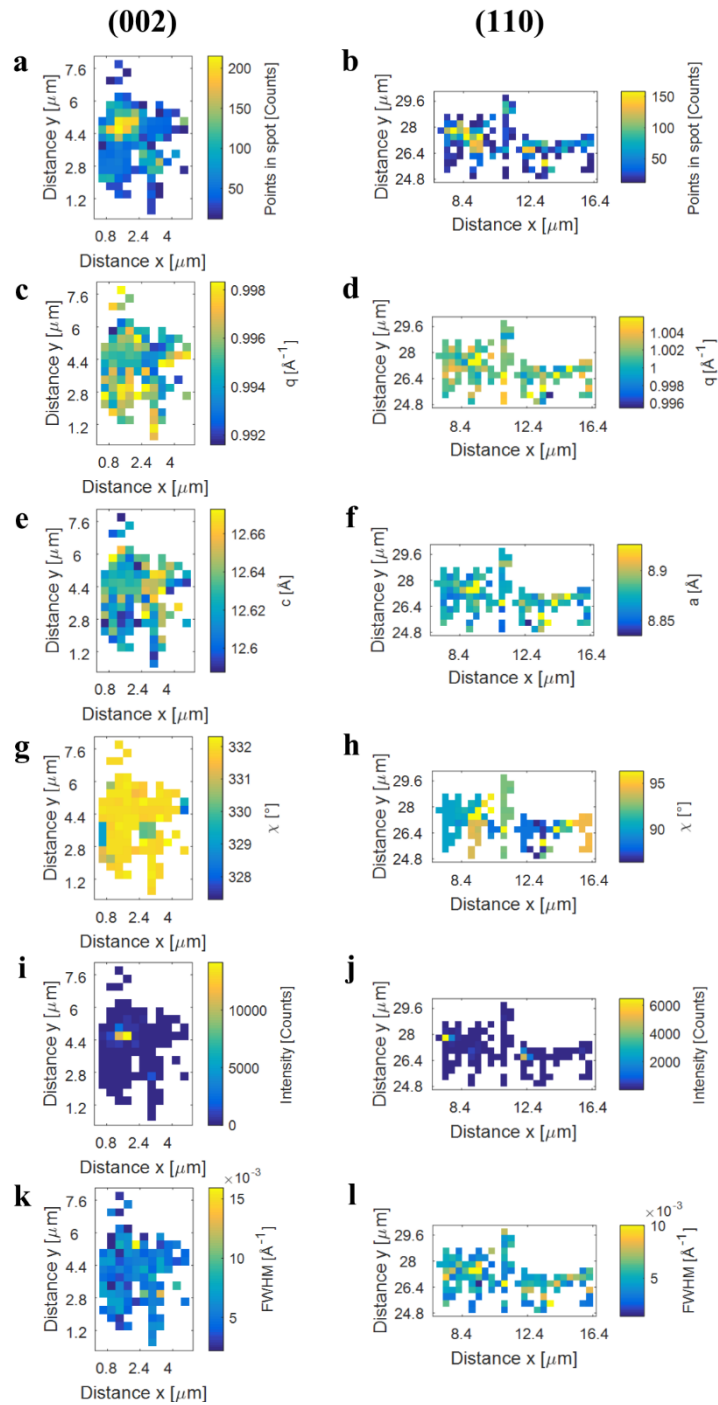


Figure S24. Spatial maps of diffraction spot peak properties in the largest (002) and (110) grains (hot sample). **a, b**, number of non-zero pixels in the diffraction spot detected in the first level of clustering. **c, d**, diffraction spot q_p peak position. **e, f**, lattice constant calculated from q_p . **g, h**, diffraction spot χ_p peak position. **i, j**, diffraction spot peak intensity. **k, l**, FWHM extracted from the line profile passing by χ_p and q_p and the origin of the diffraction pattern. Note that **l** is blank because the identified diffraction spots did not contain enough pixels for a good estimate of the FWHM.

3.10 Quiver Plots Magnifications

The following are magnifications of the quiver plots shown in the manuscript.

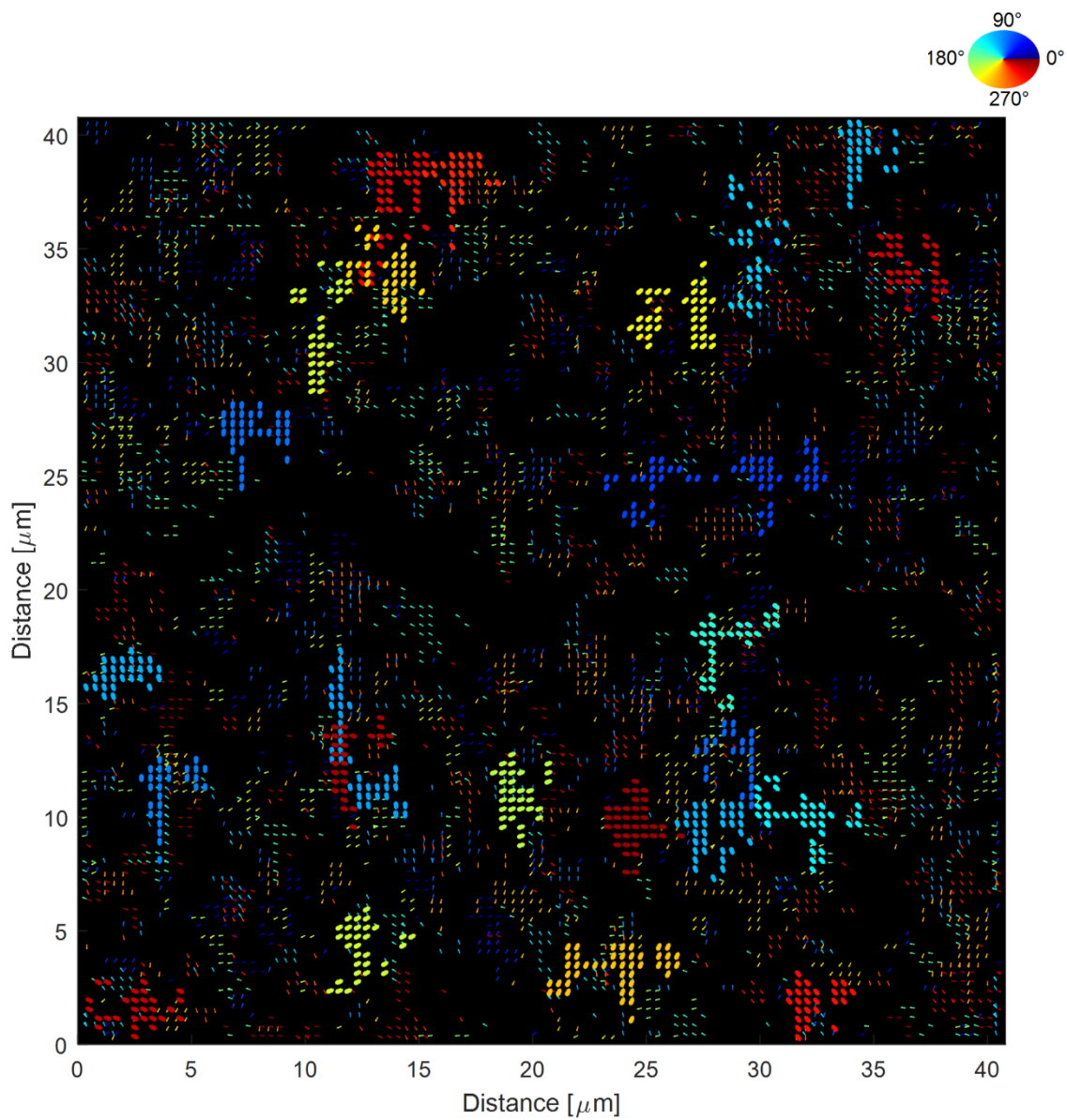


Figure S25. (002) Cold quiver magnification.

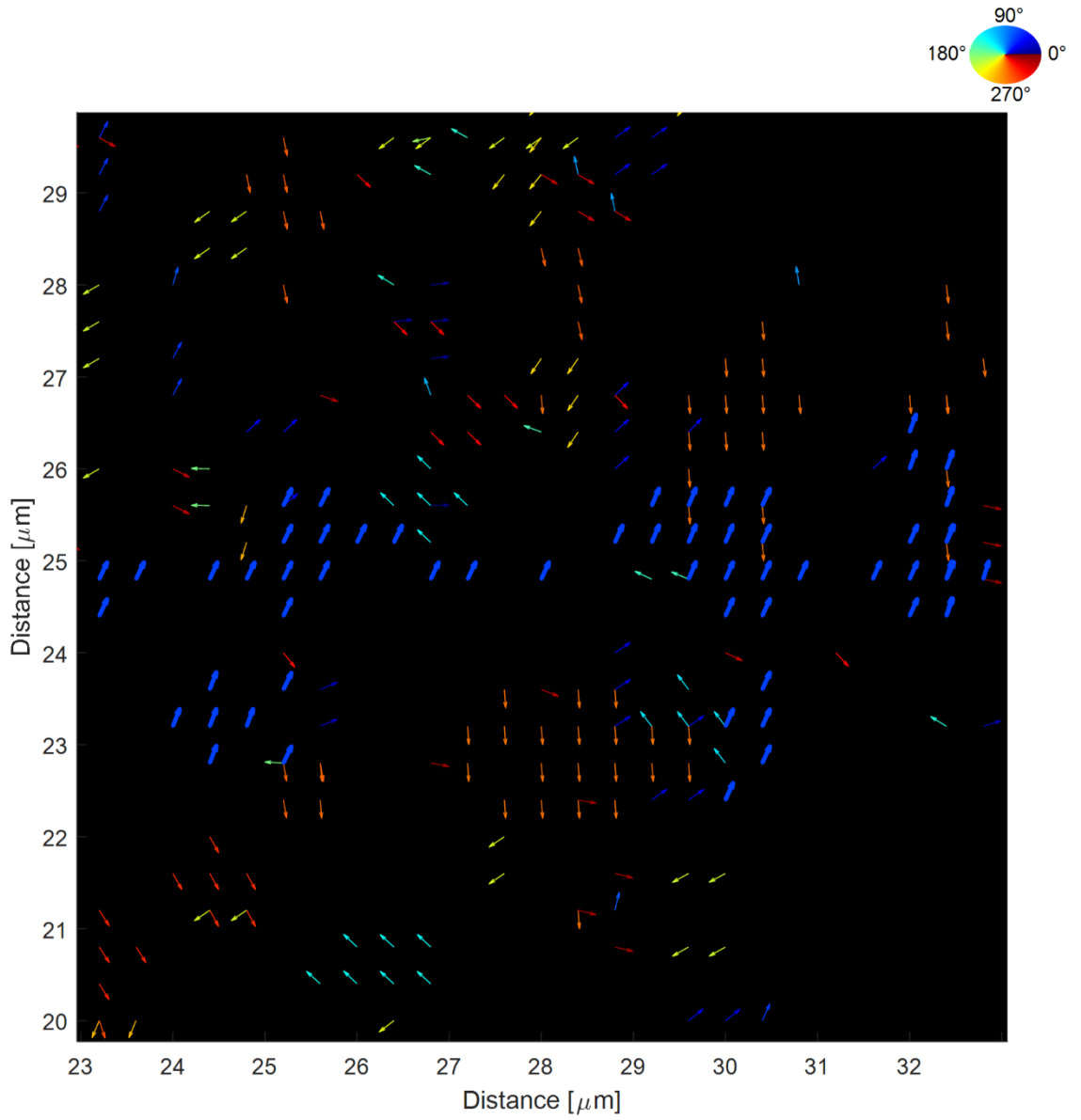


Figure S26. (002) Cold quiver magnification around the largest grain (blue thick arrows).

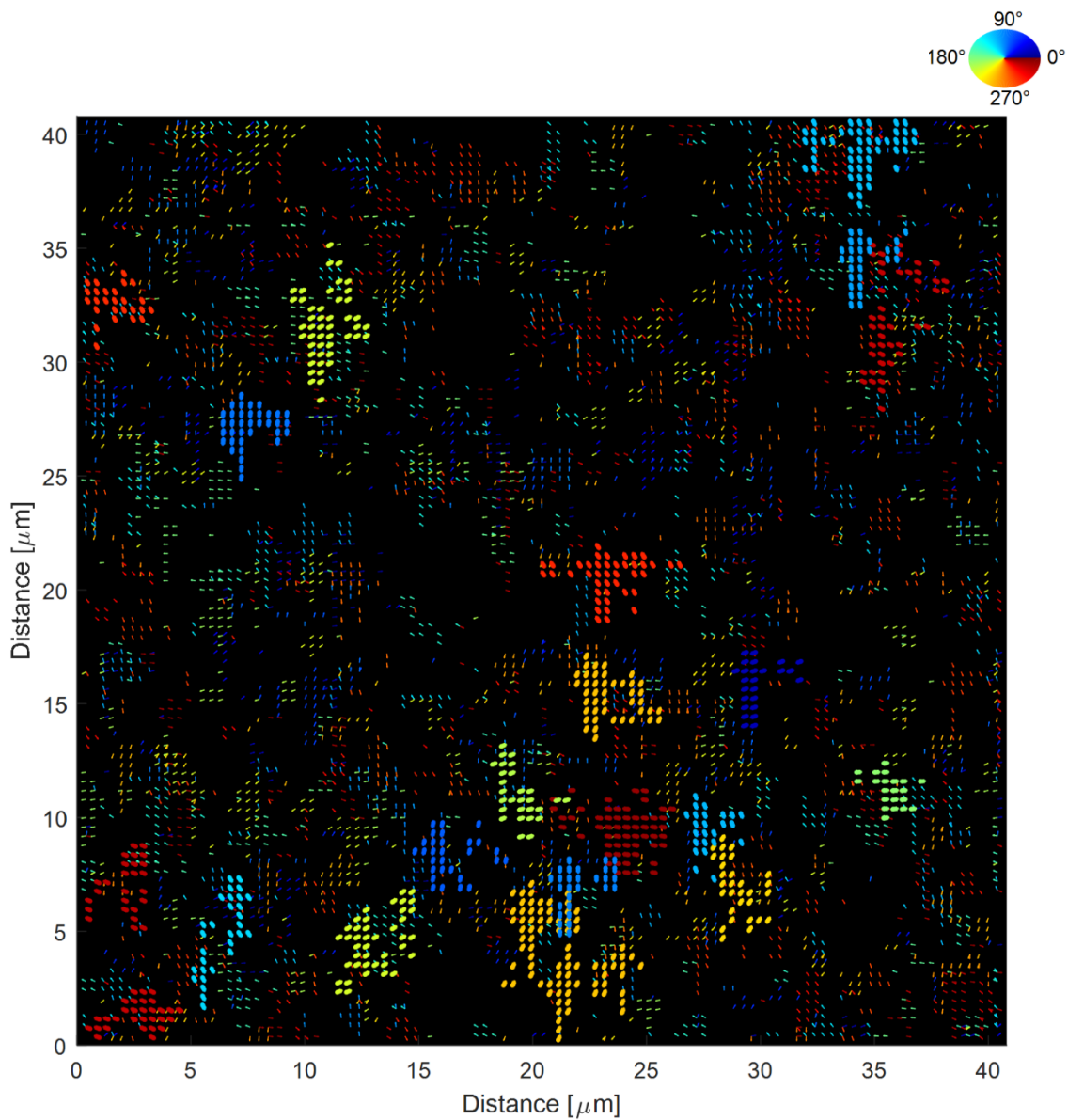


Figure S27. (110) Cold quiver magnification.

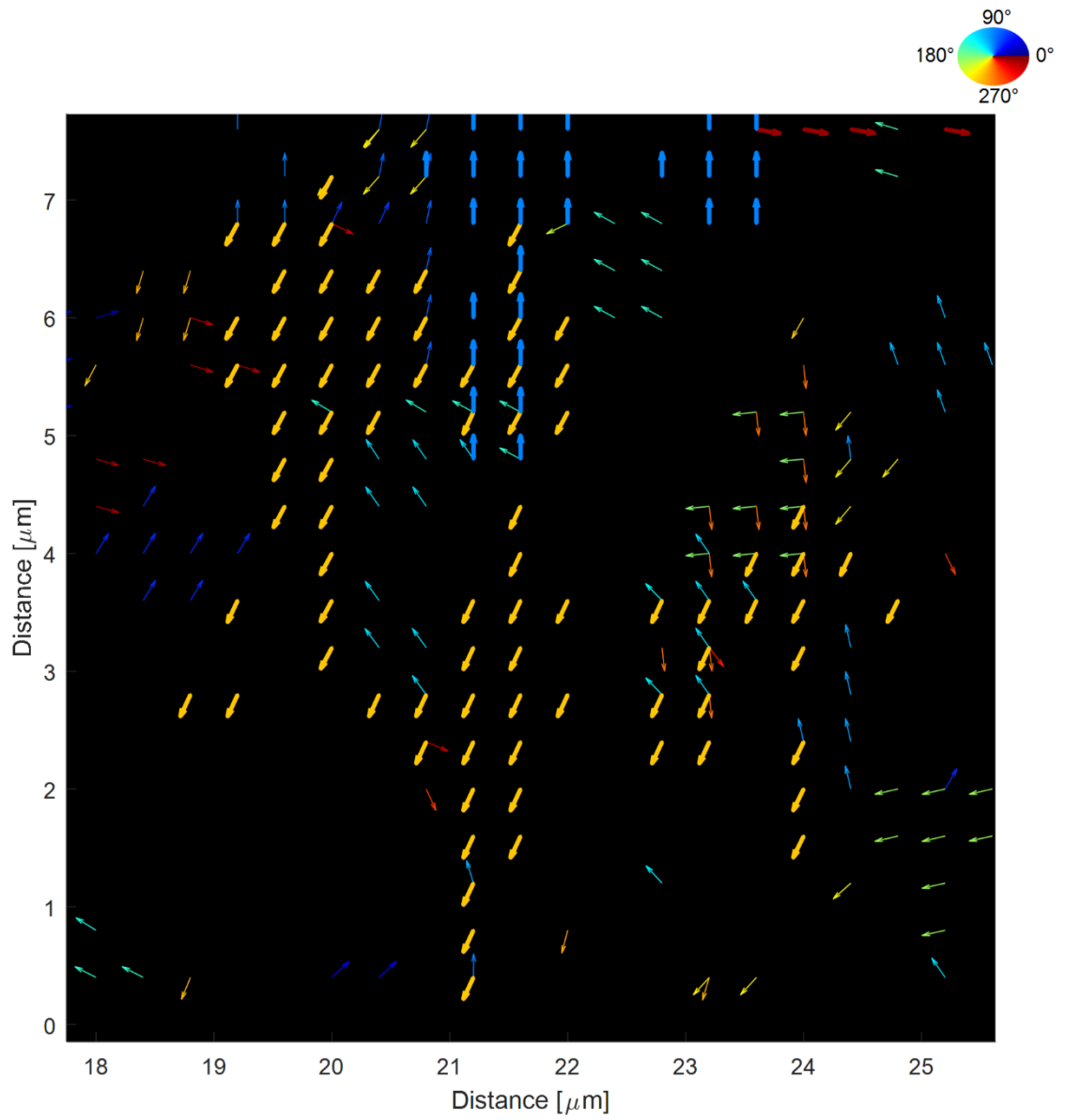


Figure S28. (110) Cold quiver magnification around the largest grain (yellow thick arrows).

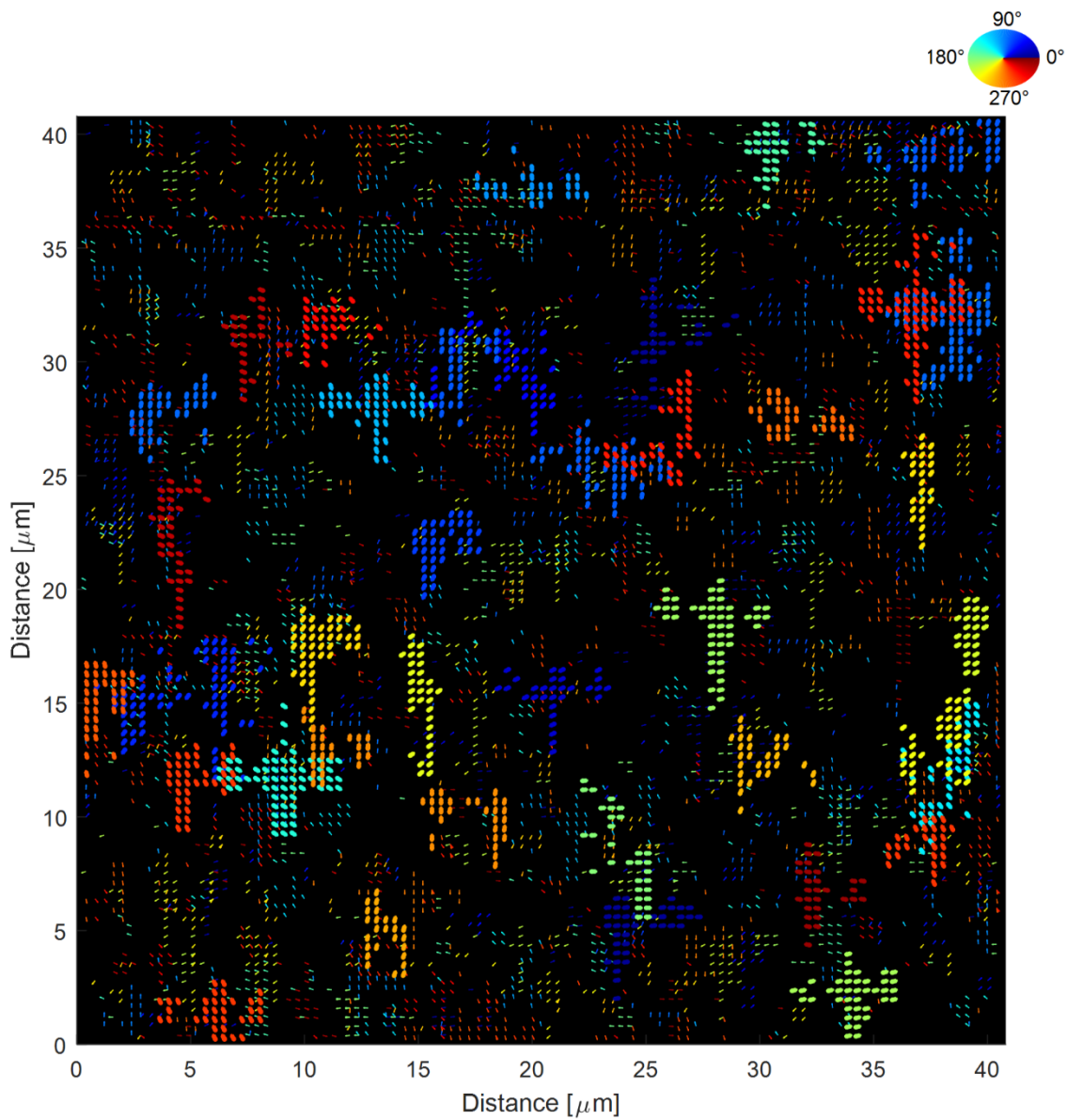


Figure S29. (002) Medium quiver magnification.

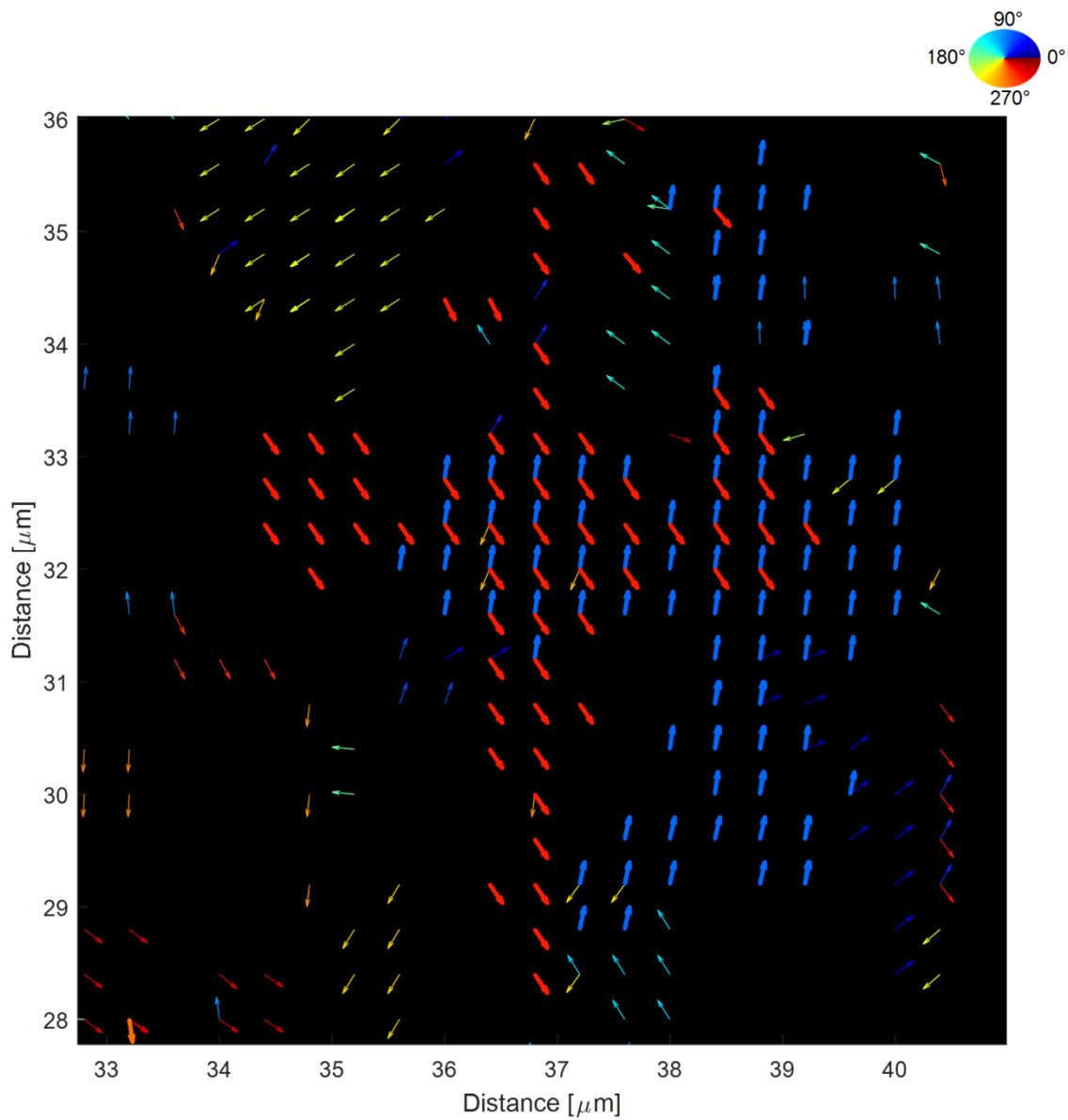


Figure S30. (002) Medium quiver magnification around the largest grain (Jordy blue thick arrows).

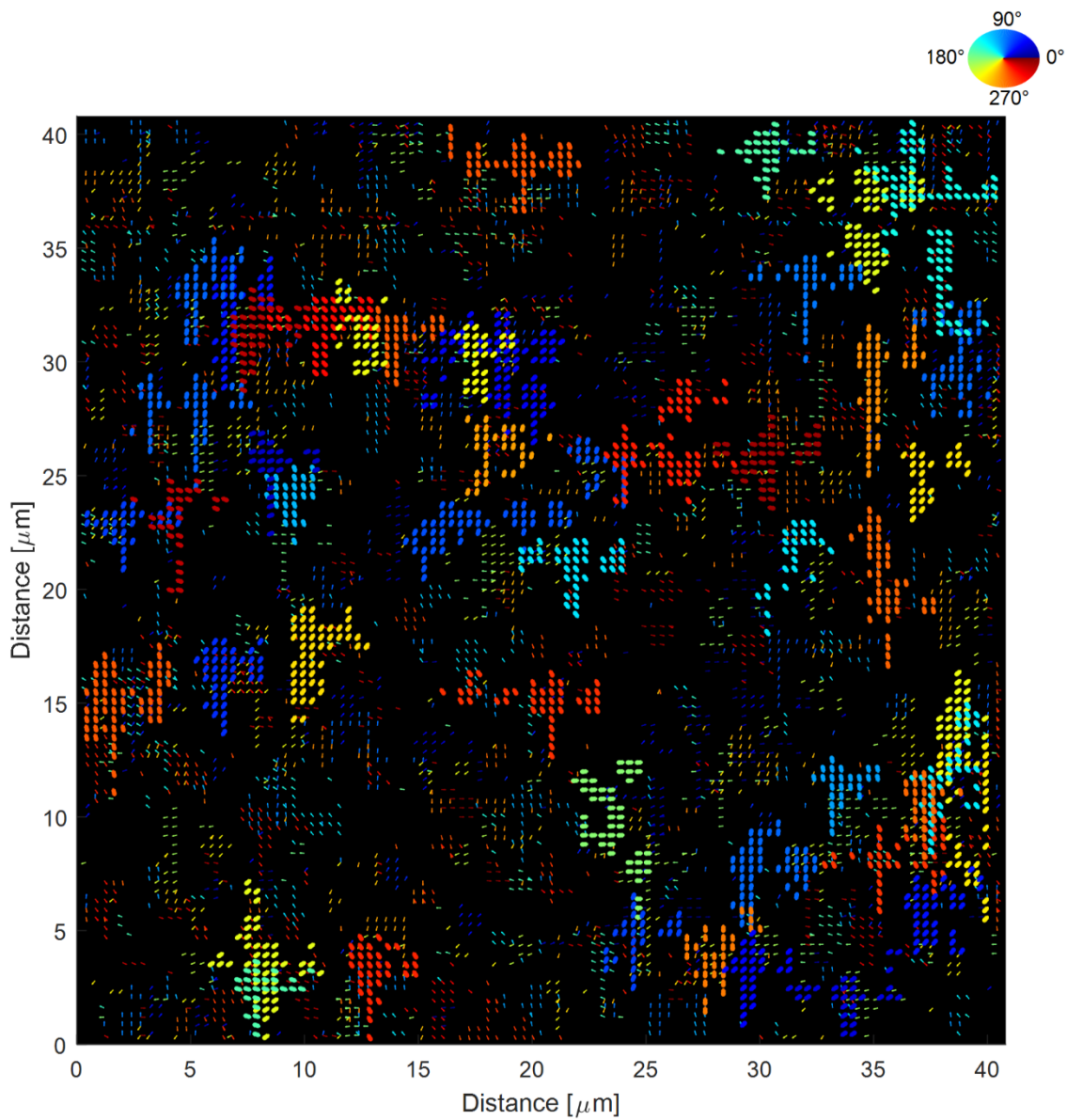


Figure S31. (110) Medium quiver magnification.

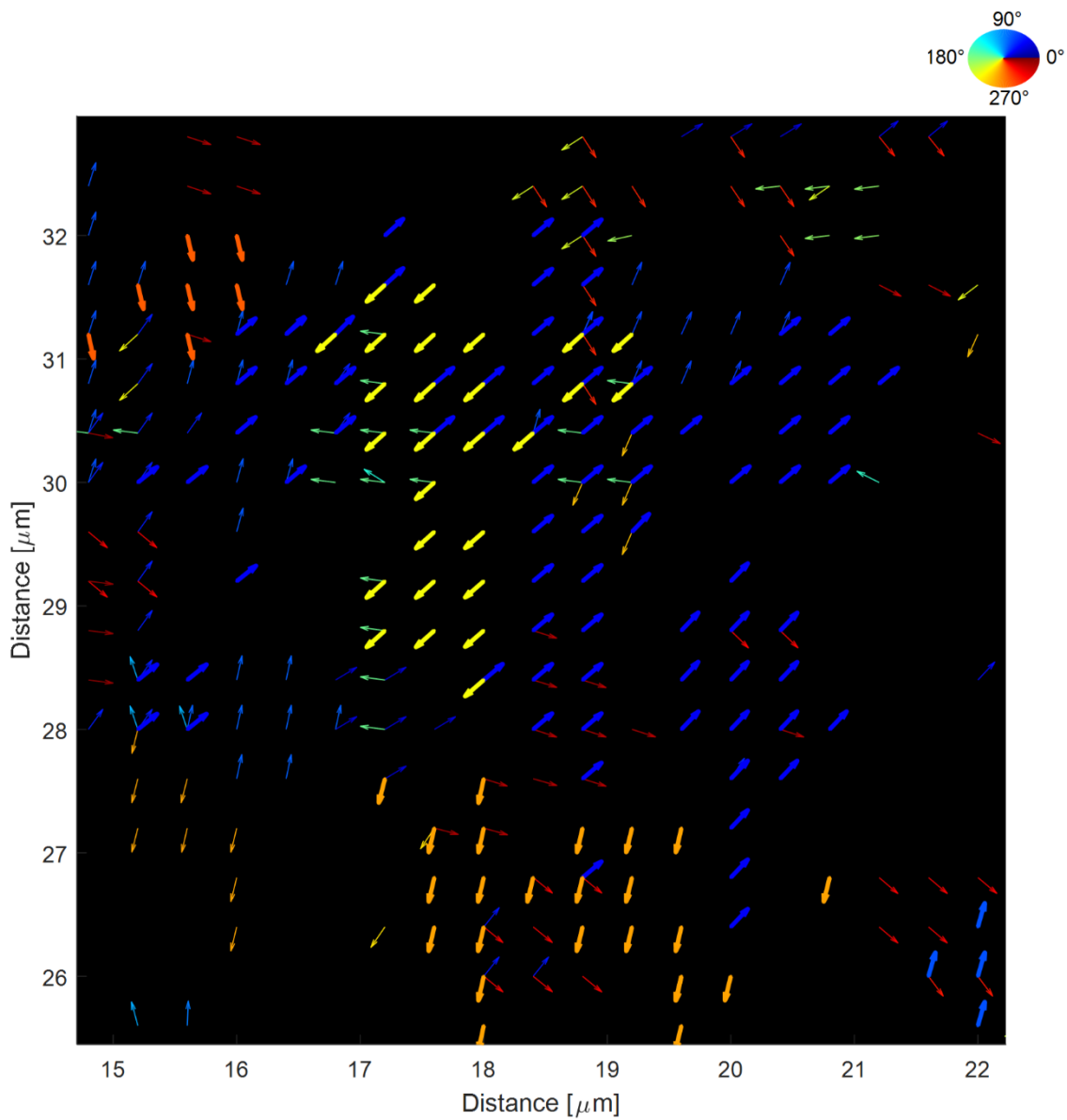


Figure S32. (110) Medium quiver magnification around the largest grain (blue thick arrows).

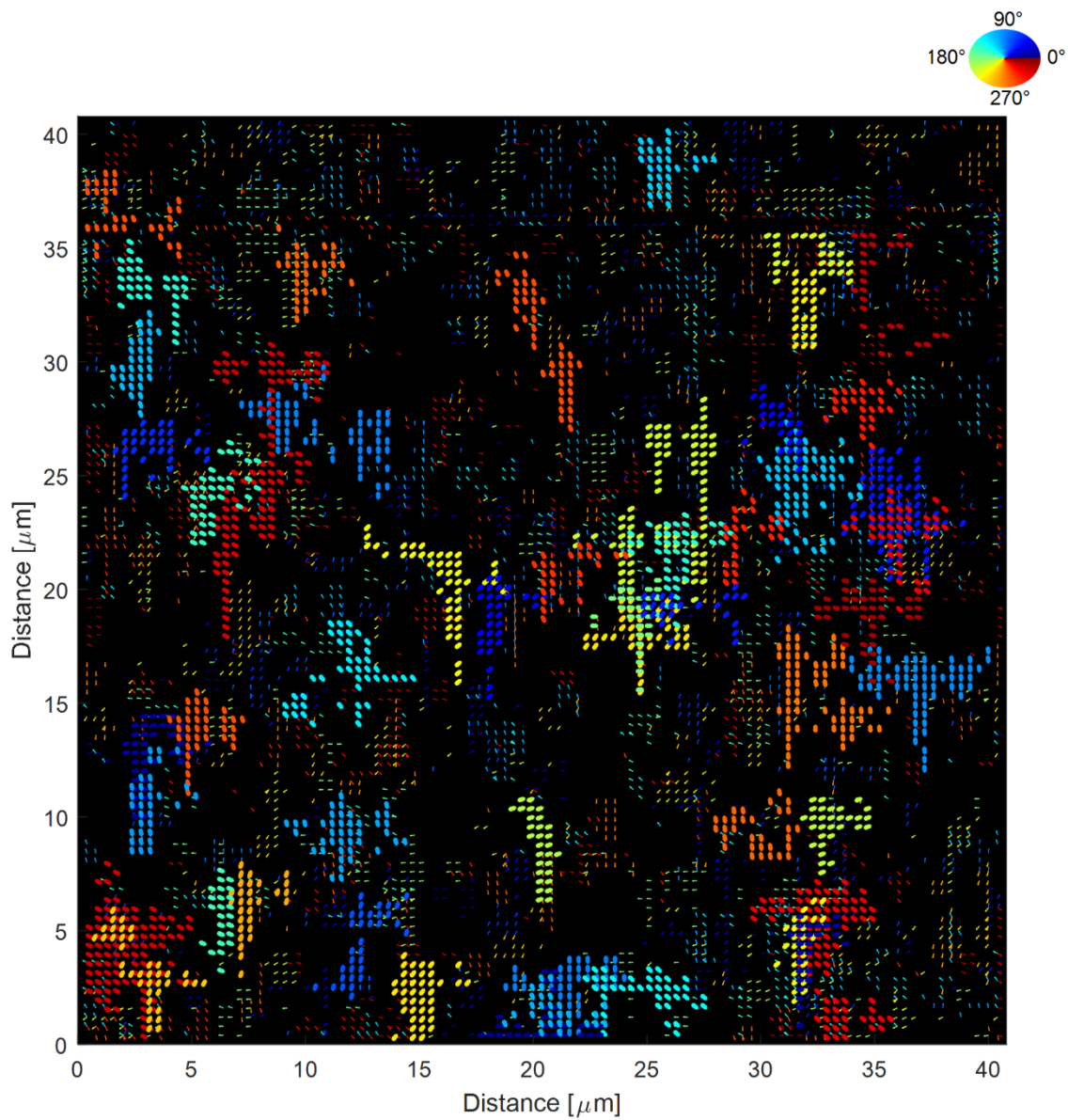


Figure S33. (002) Hot quiver magnification.

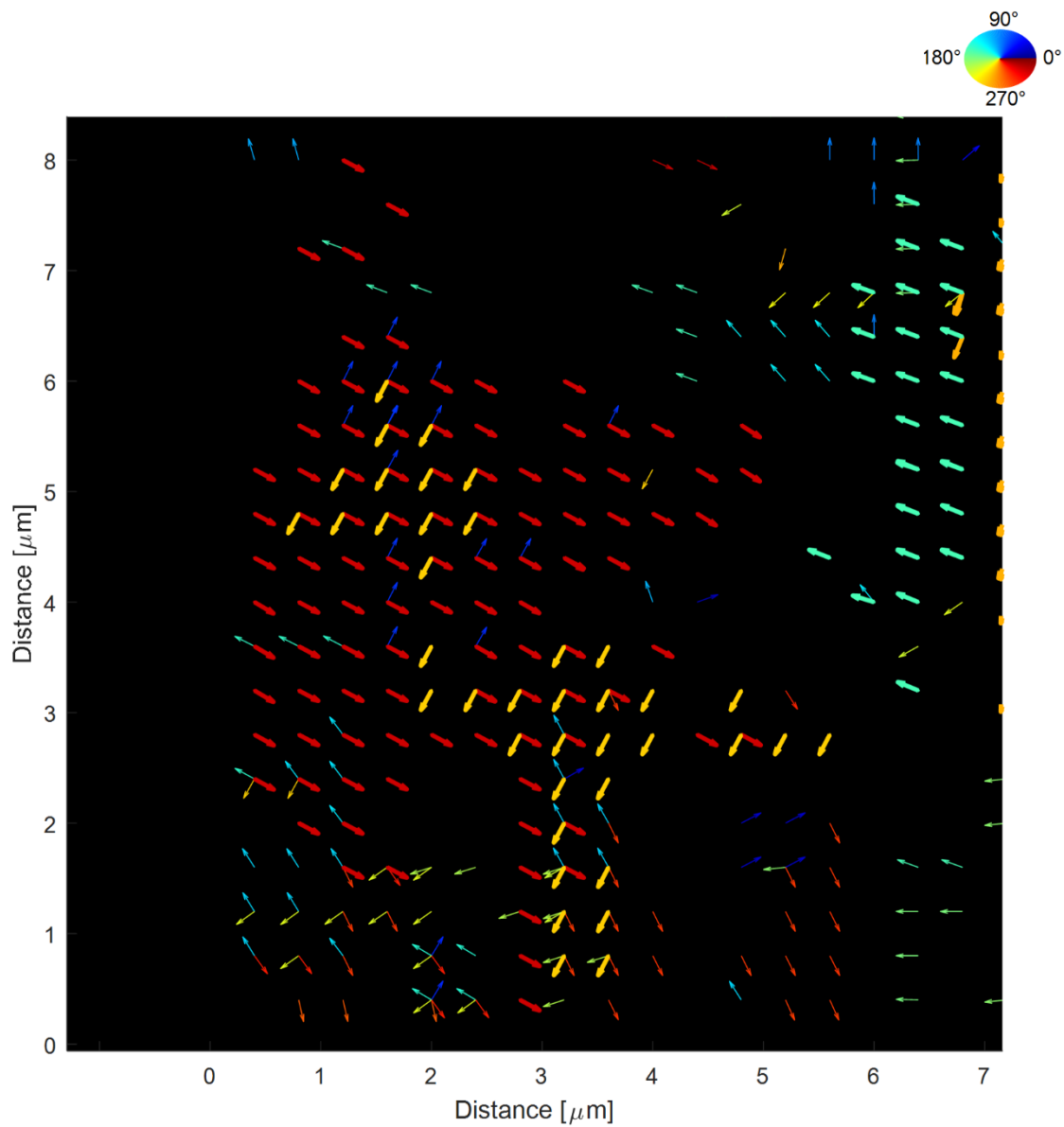


Figure S34. (002) Hot quiver magnification around the largest grain (red thick arrows).

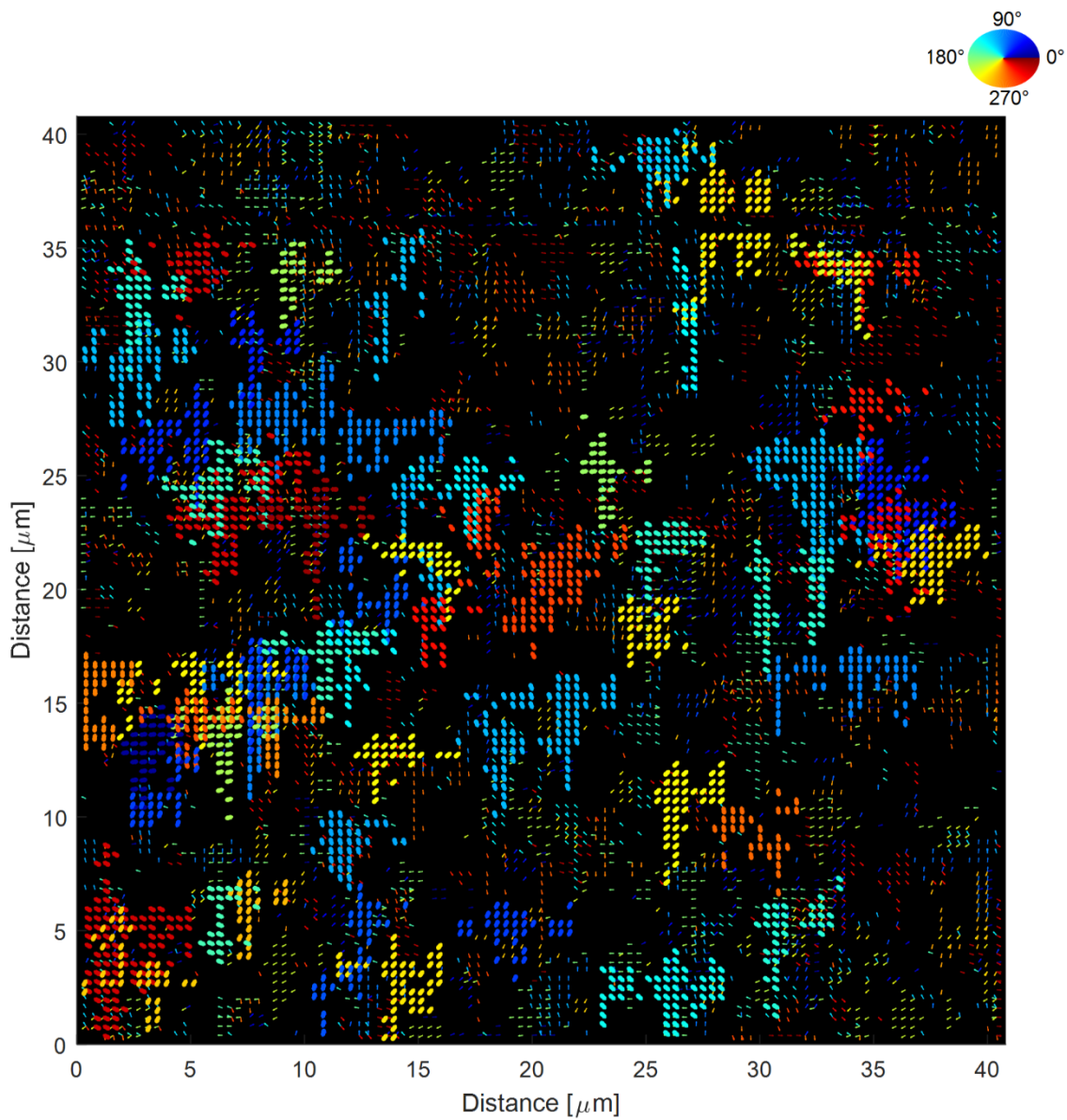


Figure S35. (110) Hot quiver magnification.

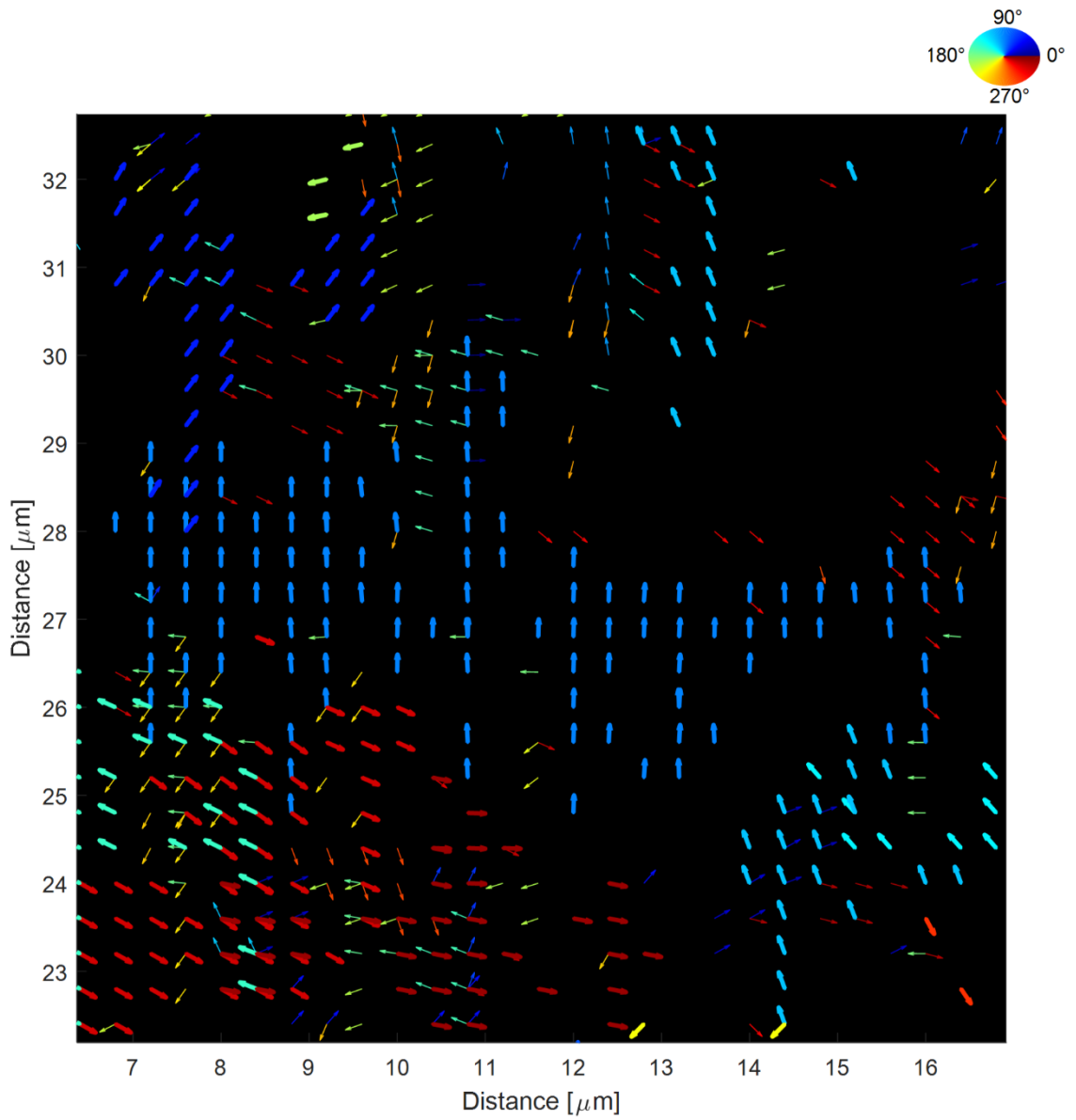


Figure S36. (002) Hot quiver magnification around the largest grain (azure thick arrows).

4 Raw Data Instructions

4.1 nXRD

We provide six *.mat files named cold_002, cold_110, medium_002, medium_110, and hot_002, hot_110, which contain the data processed by our software and discussed in this work. The raw data in *.edf format (510 GB) can be requested either to the ID13 beamline or by sending an email to samuele_lilliu@hotmail.it. The *.mat files can be opened using MATLAB by simply dragging and dropping them to the Workspace. The data is accessible from the variable 'nongraphicHandles'.

The data structures listed in **Table S9** are common to the '002' and '110' files (e.g. would be the same in both cold_002 and cold_110) are generated before the first clustering step discussed above.

Table S9. Data structures common in both 002 and 110 files.

Struct Name	Type	Description
nongraphicHandles.data{n}	Sparse matrix	Single full diffraction pattern. n can be any integer number between 1 and 10201.
nongraphicHandles.montage.ImageL	Sparse matrix	Montage (as shown in Figure 2)
nongraphicHandles.image.lav	Sparse matrix	Average diffraction pattern (as shown in Figure 2).
nongraphicHandles.image.ImageSum	Sparse matrix	101 × 101 image where the <i>n</i> -th pixel corresponds to the average intensity of the <i>n</i> -th full diffraction pattern in a scan. NOTE: 'Sum' is just a name.

Sparse patterns can be visualized through the command ‘spy’. Unfortunately, the spy command does not render intensities with different colors. A simple workaround can be found at the beginning of the code in **Table S10**, which can be used for plotting the montage.

Table S10. Simple improvement to the ‘spy’ command and code for plotting the montage.

```
I0 = nongraphicHandles.montage.ImageL;
[x,y, I] = find(I0);
scatter(y, x, 10, I, '.');
set(gca, 'CLim', [3 10]);
xlim([min(y)-500 max(y)+500]);
ylim([min(x)-500 max(x)+500]);
c = colorbar;

ax = gca;
set(ax, 'Visible', 'on');
set(ax, 'layer', 'top');
set(ax, 'GridLineStyle', '-');
set(ax, 'Xgrid', 'on');
set(ax, 'Ygrid', 'on');

S = size(I0);
xgap = 10; ygap = 10;

xgap_pel = (S(2)/nongraphicHandles.input.scansize_um(2))*xgap;
xTick = [1 round(xgap_pel:xgap_pel:S(2))];
xTickLabel = [0 xgap:xgap:nongraphicHandles.input.scansize_um(2)];
set(ax, 'xTick', xTick);
set(ax, 'xTickLabel', xTickLabel);
xlabel('Distance [\mum]');

ygap_pel = (S(1)/nongraphicHandles.input.scansize_um(1))*ygap;
yTick = [1 round(ygap_pel:ygap_pel:S(1))];
yTickLabel = fliplr([0 ygap:ygap:nongraphicHandles.input.scansize_um(1)]);
set(ax, 'yTick', yTick);
set(ax, 'yTickLabel', yTickLabel);
ylabel('Distance [\muum]');
```

Images having a ‘reasonable’ size (i.e. data{index} and image.Iav) such as the single full diffraction patterns or the average diffraction pattern can be either visualized with the ‘spy’ command or as regular images after converting them into full matrices with the ‘full’ command. The code in **Table S11** can be used for plotting single diffraction patterns (nongraphicHandles.data{n}) or average diffraction patterns (nongraphicHandles.image.Iav).

Table S11. Rendering sparse matrices into images. The following code can also be used with nongraphicHandles.data{index} or nongraphicHandles.image.ImageSum, but not with nongraphicHandles.montage.ImageL.

```
I0 = nongraphicHandles.image.Iav;
I1 = full(I0);
I2 = mat2gray(I1, [0 10]);
I3 = imadjust(I2, [0 1]*1e-4);
imshow(I3);
colormap hot

xgap = 500;
xTick = [1 (1:floor(nongraphicHandles.image.size(2)/xgap))*xgap];
xTickLabel = xTick;
xlabel('x [pixel]');
ygap = xgap;
yTick = xTick;
yTickLabel = xTickLabel;
ylabel('y [pixel]');

ax = gca;
set(ax, 'Visible', 'on');
set(ax, 'layer', 'top');
set(ax, 'GridLineStyle', '-');
set(ax, 'Xgrid', 'on');
set(ax, 'Ygrid', 'on');

set(ax, 'YTick', yTick);
set(ax, 'YTickLabel', yTickLabel);
ylabel(yLabel);

set(ax, 'XTick', xTick);
set(ax, 'XTickLabel', xTickLabel);
xlabel(xLabel);

hold on;
plot(nongraphicHandles.input.beam(2), nongraphicHandles.input.beam(1), 'r+'); %
direct beam
```

The surface coverage reported in the manuscript is calculated with the code in **Table S12**.

Table S12. Calculation of the surface coverage.

```
100*(nnz(nongraphicHandles.montage.ImageSum)...
/ numel(nongraphicHandles.montage.ImageSum))
```

The following data is obtained after the second clustering step and is different between the 002 and the 110 files.

Table S13. Data structures different for 002 and 110 files.

Struct Name	Type	Description
nongraphicHandles.montage.quiver	cell	Quiver data
nongraphicHandles.montage.quiver{1}.data(:,4)	array	Number of 002 or 110 diffraction spots per diffraction image.

The code in **Table S14** can be used for plotting the quiver plot.

Table S14. MATLAB code for plotting the quiver plot.

```

popup_value = 1;

figure; ax = gca; hold on;
set(ax, 'Color', [0 0 0]);

ColorMap = [linspace(0, 360, 256)' jet(256)];

CLM = [nongraphicHandles.montage.quiver{popup_value}.data(:,2)...
       nongraphicHandles.montage.quiver{popup_value}.data(:,3)...
       nongraphicHandles.montage.quiver{popup_value}.data(:,7)]; % col, row,
angle

quiverGap =
nongraphicHandles.input.scansize_um./nongraphicHandles.input.scansize_pel;

for i = 1:numel(ColorMap(:,1))-1
    index = find(CLM(:,3) >= ColorMap(i,1) & CLM(:,3) < ColorMap(i+1,1));
    [u, v] = pol2cart(deg2rad(CLM(index,3)), min(quiverGap)/2);
    quiver(CLM(index,1)*quiverGap(2), (nongraphicHandles.input.scansize_pel(1)-
CLM(index,2)+1)*quiverGap(1),...
           u, v, ...
           0,... % resize set to zero
           'Color', ColorMap(i,2:4), 'MarkerSize', 2,...
           'MarkerFaceColor', [0 0 0]);
end

set(ax, 'XLim', [0 nongraphicHandles.input.scansize_um(2)+quiverGap(2)])
set(ax, 'YLim', [0 nongraphicHandles.input.scansize_um(1)+quiverGap(1)])

set(ax, 'DataAspectRatio', [1 1 1])
xlabel('Distance [\mum]); ylabel('Distance [\mum]);

```

Statistics on the degree of overlap between different grains that exist at the same spatial position, can be performed with the code in **Table S15**.

Table S15. MATLAB code for plotting the quiver plot.

```
data = nongraphicHandles.montage.quiver{1}.data(:,4);

edges = 0.5:9.5;
x = 1:9;
counts = histcounts(data, edges);
figure, bar(x, counts);
ax = gca; ax.YScale = 'log';
xlabel('Diffraction spots per diffraction pattern');
ylabel('Counts');

mean(data)
std(data)
```

4.2 Solar Cell Measurements

JV sweeps are available as supporting files. Note that the external pixels (1 and 6) have been excluded from the statistics as they consistently perform poorly due to the proximity to the substrate edge. A faulty pixel (or a bad measurement) was found in the hot backward measurements and removed from the statistics as well.

5 References

- [1] S. Lilliu, J. Griffin, A. T. Barrows, M. Alsari, B. Curzadd, T. G. Dane, O. Bikondoa, J. E. Macdonald, and D. G. Lidzey, *CrystEngComm* **2016**, DOI: 10.1039/C6CE00965D,
- [2] S. Lilliu, M. Alsari, O. Bikondoa, J. E. Macdonald, and M. S. Dahlem, *Sci. Rep.* **2014**, 5,
- [3] S. Lilliu and T. Dane, *arXiv:1511.06224 [cond-mat.soft]* **2015**, -,
- [4] S. Lilliu, T. Agostinelli, E. Pires, M. Hampton, J. Nelson, and J. E. Macdonald, *Macromolecules* **2011**, 44, 2725.
- [5] M. Birkholz, *Thin Film Analysis by X-Ray Scattering*. John Wiley & Sons, **2006**.
- [6] The_Center_for_X-Ray_Optics, X-Ray Interactions with Matter, http://henke.lbl.gov/optical_constants/getdb2.html, accessed: 04, 2016.
- [7] J. Als-Nielsen and D. McMorrow, *Elements of modern X-ray physics*. John Wiley & Sons, **2011**.
- [8] D.-M. Smilgies, *J. Appl. Crystallogr.* **2009**, 42,
- [9] M. Saliba, K. W. Tan, H. Sai, D. T. Moore, T. Scott, W. Zhang, L. A. Estroff, U. Wiesner, and H. J. Snaith, *The Journal of Physical Chemistry C* **2014**, 118,
- [10] D. T. Moore, H. Sai, K. W. Tan, D.-M. Smilgies, W. Zhang, H. J. Snaith, U. Wiesner, and L. A. Estroff, *J. Am. Chem. Soc.* **2015**, 137, 2350.
- [11] MathWorks, MATLAB Documentation, <http://uk.mathworks.com/help/MATLAB/getting-started-with-MATLAB.html>, accessed: 04, 2016.
- [12] A. Quarteroni, R. Sacco, and F. Saleri, *Numerical mathematics*. Springer Science & Business Media, **2010**.
- [13] P. Boesecke, *J. Appl. Crystallogr.* **2007**, 40, s423.
- [14] G. Ashiotis, A. Deschildre, Z. Nawaz, J. P. Wright, D. Karkoulis, F. E. Picca, and J. Kieffer, *J. Appl. Crystallogr.* **2015**, 48, 510.
- [15] pyFAI, Fast Azimuthal Integration using Python, <http://pyfai.readthedocs.org/en/latest/index.html>, accessed: 04, 2016.
- [16] K. Jérôme and K. Dimitrios, *JPCS* **2013**, 425, 202012.
- [17] pyFAI, Fast Azimuthal Integration using Python, <http://pyfai.readthedocs.org/en/latest/index.html>, accessed:
- [18] H. Rietveld, *J. Appl. Crystallogr.* **1969**, 2, 65.
- [19] C. C. Stoumpos, C. D. Malliakas, and M. G. Kanatzidis, *Inorg. Chem.* **2013**, 52, 9019.

NOTE on Figure 1a-b of the manuscript. Microscope part, credits: muhtesemozcinar, <http://www.blendswap.com/blends/view/80149>