

Understanding nanoscale structural distortions by utilizing x-ray nanodiffraction and clustering algorithm analysis

Joyce Christiansen-Salameh,¹ Morris Yang,¹ Geoffrey Rippey,¹ Jianheng Li,¹ Zhonghou Cai,^{2,3} Martin Holt², Guillaume Agnus,⁴ Thomas Maroutian,⁴ Philippe Lecoeur⁴, Sylvia Matzen,⁴ Roopali Kukreja¹

1- Department of Materials Science and Engineering, University of California Davis, Davis, California 95616, USA

2- Center for Nanoscale Materials, Argonne National Laboratory, Lemont, Illinois 60439, USA

3- Advanced Photon Source, Argonne National Laboratory, Lemont, Illinois 60439, USA

4- Center for Nanoscience and Nanotechnology (C2N), CNRS, Université Paris-Saclay, 91120 Palaiseau, France

Abstract

Hard x-ray nanodiffraction provides a unique nondestructive technique to quantify local strain and structural inhomogeneities at nanometer lengthscales. However, sample mosaicity and phase separation can result in complex diffraction pattern that can make it challenging to quantify nanoscale structural distortions. We have utilized k-means clustering algorithm to identify local maxima of intensity by partitioning diffraction data in a three dimensional feature space of detector coordinates and intensity. This technique is applied to x-ray nanodiffraction measurements of a patterned ferroelectric $\text{PbZr}_{0.2}\text{Ti}_{0.8}\text{O}_3$ sample. Our analysis reveals the presence of two phases in the sample with different lattice parameter. We also observe a highly heterogeneous distribution of lattice parameter with a variation of $0.02 \text{ \AA}$ within one ferroelectric domain. Our approach provides a nanoscale survey of subtle structural distortions as well as phase separation in ferroelectric domains in a patterned sample.

1. Introduction

In the last few decades, hard x-ray microscopy techniques have emerged as a novel method to characterize nanoscale heterogeneities in functional oxides [1-5]. Specifically, hard x-ray microscopy techniques have shown certain advantages over complementary electron microscopy techniques for functional oxides, as no alteration of local boundary conditions occur during sample preparation or measurement due to comparatively weaker sample/beam interaction. Hard x-ray nanodiffraction, which combines benefits of diffraction and imaging, by using a zone plate to focus down the x-ray beam to spot size of 25 nm, has been utilized to investigate strain fields and structural heterogeneities at nanoscale in a variety of functional oxides systems [6-9]. However, most of these studies have been restricted to epitaxial oxide thin films with negligible mosaicity as current analysis methods for plotting local strain maps are based on mapping out the variation of a single Bragg peak over a specific sample region [6-7]. Sample mosaicity and heterogeneities present in a sample can result in a complex diffraction pattern overlaid with zone plate diffraction pattern and can comprise of more than one local intensity maxima as x-ray beam diffracts from multiple strain states across several grains along the sample depth. In order to separate these multiple effects, for example, identifying the role of local strain and correlating it with microscopic functional properties, robust analysis methods for complex diffraction patterns obtained from real-world samples are critically needed.

In this article, we present k-means clustering algorithm based analysis of complex diffraction patterns which can extend the capabilities of x-ray nanodiffraction to highly mosaic and heterogeneous samples. Cluster analysis is based on self-consistently identifying and classifying common elements within a dataset (diffraction patterns in this case), and has been widely used for image segmentation and quantitative image analysis [10-15]. K-means is one of the clustering

algorithms which finds the centroids by iteratively alternating between two steps which include assigning data points to clusters based on the present centroids and then choosing next centroids based on the present assignment of data points to the cluster. K-means algorithm has been used in medical imaging, life sciences, astronomy and astrophysics to segment superpixels, or sets of pixels with similar attributes, from an image background [16-23]. We utilized k-means algorithm to identify local maxima of intensity by partitioning the diffraction data within a region of interest (ROI) in a 3D feature space of detector coordinates and intensity. This enabled us to ascertain centroids present in complex nanodiffraction pattern, which were then used to obtain local lattice parameters and strain.

This method is applied here to x-ray nanodiffraction measurements of a patterned $\text{Pb}(\text{Zr}_{0.20}\text{Ti}_{0.80})\text{O}_3$ (PZT) sample. PZT is an archetypical piezoelectric and ferroelectric material where the polarization state is coupled with lattice strain and can be controlled by application of electric field. Ferroelectric materials such as PZT show a thermodynamically stable and electrically switchable polarization [24 – 26] which is of interest in applications such as sensor and actuator devices for microelectromechanical systems (MEMS) [27-28] and nonvolatile memories [29]. However, at nanoscale the structural distortions can be heterogeneous, and local strain fields can play a huge role in determining the functional properties [30]. While piezoresponse force microscopy (PFM) has been utilized to manipulate and probe ferroelectric polarity in these systems, it provides relatively little information about the atomic structure of the ferroelectric domains, especially as the information obtained is limited to the surface of the sample. The structural interpretation of PFM data is further complicated by the complex interaction between the probe tip and surface of ferroelectric material [6]. Transmission electron microscopy (TEM) has also been used to map out atomic structure and ferroelectric polarization, however it is

restricted to extremely thin samples and sample preparation can be challenging and destructive [31 – 33]. In order to map out the local strain variation non-destructively, we have performed x-ray nanodiffraction studies on patterned PZT samples. Complex diffraction patterns obtained from PZT samples were then analyzed using clustering algorithms to quantify local strain in ferroelectric domains. Our analysis shows presence of two distinct phases with different lattice parameters distributed potentially across the film thickness. One of the phases shows a clear distinction in intensity and lattice parameter consistent with up and down ferroelectric domain. The second phase shows a relatively homogenous distribution of intensity, which could be due to structural response to unscreened charges at the surface or the interface.

2. Methods

X-ray nano-diffraction experiments were conducted at the Hard X-ray Nanoprobe (26-ID-C) beamline of the Advanced Photon Source and Center for Nanoscale Materials at the Argonne National Laboratory. Additional details about the beamline and nanodiffraction technique can be found here [1, 4, 5, 34]. The monochromatic incident x-ray beam (9 keV) was focused to a 25 nm spot size with a Fresnel zone plate, and an order sorting aperture and center stop blocked all but the first order beam. The focusing led to an overall effective beam divergence of 0.24° . This first-order diffracted x-ray beam was utilized to perform x-ray nanodiffraction studies in reflective geometry. Two dimensional Pixel Area Detector (PAD) with a 55 μm pixel size placed 860 cm from the sample was used to collect diffraction pattern. The detector was tuned to a 2θ -value of 38.4° to access the (002) Bragg peak.

A patterned PZT thin film sample grown using pulsed laser deposition on a $\text{La}_{0.7}\text{Sr}_{0.3}\text{MnO}_3$ (LSMO)-buffered SrTiO_3 (STO) substrate was utilized. LSMO and PZT layers were successively grown by pulsed laser deposition (KrF excimer laser, 248 nm wavelength) on a single-crystal,

TiO₂-terminated STO (001) substrate held at a temperature of 630 °C. The laser was operated at a pulse rate of 4 Hz, with a target-to-substrate distance of 5 cm and a fluence of 2.5 J cm⁻². The PZT ceramic target was sintered with stoichiometric Zr/Ti ratio and 12% excess PbO. The 30 nm thick LSMO bottom electrode was first deposited under an oxygen pressure of 120 mTorr, then replaced by nitrous oxide (N₂O), at the same pressure to grow a 100 nm thick PZT layer. After deposition, the sample was cooled down to room temperature under a static oxygen pressure of 300 Torr. A rocking curve (around the (002) PZT diffraction Bragg peak) is presented in supplementary information to show the averaged mosaicity of the as-grown PZT film. Electron beam lithography was then performed to pattern 2 µm by 2 µm islands of PZT. The ferroelectric polarization of the PZT was measured and manipulated by PFM. The as-grown polarization was pointing upwards (away from the bottom LSMO electrode). Figure 1(a) presents a PFM image of one PZT island showing a two-domain state with an up and a down domain written (by applying respectively -5V and +5V to the tip). The top of the image corresponds to the up polarization domain (light brown) and the bottom of the image corresponds to the down polarization domain (dark brown).

3. X-ray nanodiffraction measurements of micro-patterned PZT thin film

Nanodiffraction maps of PZT sample were collected by raster scanning the x-ray beam and acquiring a diffraction pattern for each sample position. Figure 1(b) shows a nanodiffraction map of the PZT sample, where the integrated intensity of the (002) PZT Bragg peak is plotted as a function of sample position. The intensity difference between the top and the bottom of the image is due to oppositely polarized domains, which results in a 30% variation of the diffracted intensity due to lack of centrosymmetry of the tetragonal perovskite unit cell [6, 35-36]. This intensity variation corresponds to the up and down domain seen in the PFM image (Figure 1(a)) for the same sample region. However, while the PFM image shows only slight nonuniformity in overall

up and down written domains, the nanodiffraction map reveals a significant variation in the diffracted intensity highlighting potential inhomogeneities within a single domain.

Figure 1(c) shows a zoomed-in version of the nanodiffraction map which was used for further data analysis. Figure 1(d) and (e) show detector images of the diffraction pattern for different areas on the sample. The total intensity of such a diffraction image was used to plot the nanodiffraction map shown in Figure 1(b) and (c). Here we note that the concentric ring pattern or the annulus observed in the diffraction pattern arises from the Fresnel zone plate and center stop which is used to focus the x-rays and blocks all but the first order beam. The spread of intensities across the annulus in the diffraction pattern is due to the presence of several mosaic blocks in the sample. The contribution from each mosaic block is difficult to distinguish using simple descriptive statistics (e.g. the center of mass of the intensity) which have been previously used for nanodiffraction studies of epitaxial films [6-7] to characterize the local variation of strain. While the total intensity of the diffraction pattern can be used to differentiate between the two domains, the complex diffraction pattern makes it challenging to accurately quantify the lattice parameter of the up and down domains as well as its local variation within a single domain. Supplementary information section 1 provides such an analysis where typical centroid calculation method for nanodiffraction datasets for epitaxial samples [1, 6, 7] is applied to the PZT sample. However, it can be seen from figure S2(b) that no difference in lattice parameter for up and down domain is observed, highlighting that such a method can easily miss the nuances of real-world mosaic samples. Failure of centroid calculation method for a mosaic sample can be due to averaging of the peak position if multiple peaks are present or completely missing the intensity maxima within region of interest (ROI) typically used for nanodiffraction analysis or dependency on the overlap of the chosen ROI with the annulus.

4. K-means cluster analysis

K-means algorithm [16-23] was utilized to negate these effects, and partition the diffraction data within each ROI in 3D feature space of detector coordinates (horizontal axis - θ , vertical axis - χ) and intensity, to isolate high intensity regions that are uniquely distributed in each diffraction pattern. Multiple ROIs were selected to cover the diffraction pattern in different ways, so that the new approach is independent of the ROI selected (see supplementary Figure S3).

Figure 2 outlines the process of applying the k-means algorithm (available in MATLAB) [37, 38] to discern a local maximum within one such ROI. Figure 2(a) shows the ROI along with the three initial guesses for centroids which were used to encourage the clustering of high and low intensity data points. Figure 2(b) shows the resulting three clusters obtained upon convergence of the algorithm based on iteratively minimizing the Euclidean distance between cluster points and cluster centroids in the 3D feature space of detector 2θ , χ and intensity. Here we stress that the high intensity data points separated into cluster 3 (red distribution) is not equivalent to setting an intensity threshold for the diffraction pattern, as the cluster is determined both by spatial (detector coordinates) and intensity metrics. This allows us to locate centroid for clusters taking into account both the distance between the pixels and their intensity. In order to further refine the high intensity dataset, cluster 3 is further partitioned into two additional clusters using the same algorithm. Figure 2(c) shows the initial guess for centroids, again to encourage clustering of high and low intensity data points. Figure 2(d) shows the resulting two clusters (3-1 and 3-2), where the centroid of cluster 3-2 marks the local maximum of the ROI. This two-stage implementation of k-means first separates a superpixel (or a cluster) of general high intensity signal from the background of the diffraction pattern, and then in the second step, calculates a centroid using the high intensity points that maintain a spatial relationship within this superpixel. This two-step method enables

identification of the centroid of the highest intensity cluster which would be associated with the Bragg peak following diffraction from the PZT sample.

This method is then applied for all the ROIs to generate a set of twelve centroids (for twelve ROIs, see supplementary Figure S3) denoting local maxima of a diffraction pattern. It can be seen that the centroids are not randomly distributed over the diffraction pattern, but cluster around two higher intensity regions. In order to obtain the final centroid position of these two higher intensity regions, the twelve centroids were grouped based on the horizontal distance (along 2θ) using k-means algorithm. Only horizontal distance along 2θ direction was used, as it is directly related to d-spacing via Bragg's law, thus allowing calculation of the lattice parameter. Two final centroid positions (only x-value corresponding to a specific 2θ) are obtained corresponding to the local maxima of two higher intensity regions, and are shown in Figure 3 by red and blue dashed lines. The centroids for ROIs which were below a certain intensity threshold were neglected in this calculation, as they do not contain local maxima i.e. the gray dots in the Figure 3.

The analysis described in Figure 2 and 3 was applied for every diffraction pattern obtained for each sample position shown in Figure 1(c). In most of the diffraction patterns (>80%) presence of two final centroids was observed, highlighting the importance of this approach. Figure 4(a) plots a histogram of both centroids for every diffraction pattern obtained for each sample position. The x-axis represents the position of the centroid (along 2θ) in the pixel-value away from the center of annulus (which is taken as reference and marked as 0 pixel), and y-axis represents the number of centroids. The histogram clearly shows a bimodal distribution, labeled 1 and 2, indicating that the centroid position i.e. intensity maxima were primarily localized near two regions on opposite side of annulus. The bimodal distribution for the intensity maxima, clearly point towards two different values of d-spacing and lattice parameter indicating presence of two phases. Here we stress that

two values of lattice parameters were observed for almost all of the measured sample position (>80%), implying that two phases are present across the entire sample. These two phases have slightly different lattice parameter resulting in two different 2θ -values within a diffraction pattern which was identified by using k-means clustering algorithm.

For these two phases, diffracted intensity was separately plotted in Figure 4(b) and 4(c) using the distribution 1 and 2 respectively. The intensity map for the distribution 1 aligns with the total intensity map shown in Figure 1(c) and polarization shown by PFM image in Figure 1(a), indicating that this is indeed the ferroelectric phase with two oppositely polarized domains. These domains produce Bragg reflections at the same angle but with opposite sign of crystallographic indices i.e. (002) and (00 $\bar{2}$), and with different intensities [6]. It is interesting to note that no such demarcation between the two domains is observed for the intensity map for the distribution 2. Figure 4(d) and 4(e) show the lattice parameter map for the two phases calculated using the obtained centroid values for distribution 1 and 2 respectively. For distribution 1, the lattice parameters for the two oppositely polarized domain are $4.157 \pm 0.004 \text{ \AA}$ and $4.153 \pm 0.004 \text{ \AA}$. The lattice parameter variation observed within each domain is approximately $0.02 \text{ \AA}$. For the phase associated with distribution 1, we thus observe higher diffracted intensity and lower lattice parameter for polarization down domain compared to polarization up domain, which is consistent with the literature [6]. Surprisingly, a small variation in the lattice parameter is observed for the phase associated with the distribution 2, in spite of no clear demarcation in the intensity. The lattice parameters in this case are $4.132 \pm 0.004 \text{ \AA}$ and $4.129 \pm 0.004 \text{ \AA}$ and the lattice parameter variation observed within each domain is approximately $0.02 \text{ \AA}$. It is interesting to note that while for phase 1 the top (bottom) part of the image has higher (lower) lattice parameter, for phase 2 this contrast is reversed and the top (bottom) part of the image has lower (higher) lattice parameter.

Additionally, the lattice parameter distribution in both phases show a variation of 0.02 Å within one domain pointing towards non-homogenous nature of these domains.

5. Discussion

K-means clustering based analysis algorithm enables us to identify two phases present in a patterned PZT sample. This identification of two phases in a mosaic and complex sample cannot be done using traditional centroid calculation methods. Two phases in the patterned sample correspond to slightly different lattice parameter of 4.155 Å (phase 1) and 4.131 Å (phase 2). For phase 1, a clear demarcation of intensity and lattice parameter distribution is observed between the up and the down domain consistent with polarized ferroelectric domains observed in PFM (Figure 1(a)). On the other hand, for phase 2, the intensity variation is relatively homogenous across the top and the bottom of the image, which could imply that this second phase is not ferroelectric in nature or was not switched between the two domains. Surprisingly, lattice parameter variation across the top and the bottom of the image is still observed. It is also interesting to note that the lattice parameter contrast is flipped between the two phases, i.e. for phase 1 the up domain has higher c-lattice parameter while phase 2 shows a lower c-lattice parameter.

The variation in the lattice parameter across the film is more likely to occur near PZT surface or PZT/LSMO interface in comparison to bulk, due to different electro-mechanical boundary conditions that can modify ferroelectric properties. It is then possible that the strain state closer to the film surface or near PZT/LSMO interface is not of purely ferroelectric origin and/or was not switched by PFM resulting in the observed behavior of intensity and lattice parameter distribution. As x-rays penetrate over the entire film thickness, it is challenging to discern whether this phase 2 is formed at the surface or the bottom interface. Another possibility for the second phase could be slightly different composition of PZT resulting in different tetragonality and ferroelectric behavior.

In order to identify the exact origin of this second phase, additional characterization techniques such as cross-sectional high-resolution TEM or atomic resolution spectroscopy are required.

The identification and isolation of two different phases with slightly different lattice parameter demonstrate the importance for developing and using k-means clustering approach to investigate complex samples showing presence of mosaicity, phase separation and nanoscale heterogeneities. While the traditional centroid calculation approach works well for ideal and epitaxial system, our results show that clustering algorithm-based approach can be utilized to study complex and heterogeneous samples. Furthermore, our x-ray nanodiffraction measurements also highlight that both strain and intensity distribution show a highly heterogeneous behavior, which was not observed in PFM images. Micron scale non-uniformity of strain states within ferroelectric domains has been previously reported in a multiferroic composite that includes a PMN-PbTiO₃ layer [39]. In PZT, we observe nanoscale non-uniformity of strain states within up and down domains with a lattice parameter variation of 0.02 Å within one domain. These measurements demonstrate higher sensitivity of x-ray nanodiffraction technique, which is able to measure existence of local strain or mosaicity due to resulting large variations in diffracted intensity.

6. Conclusions

In this article, we have utilized k-means cluster analysis for nanodiffraction measurements to investigate nanoscale morphology including phase separation and mosaicity in a patterned PZT sample. The use of clustering-algorithm overcomes the limitations of previous analysis methods and allows us to identify multiple distributions resulting in presence of two phases with different lattice parameters. While one phase shows a clear difference in intensity and lattice parameter distribution consistent with PFM images indicating ferroelectric domains, the second phase shows a relatively homogenous intensity distribution and could be present near the surface or the

PZT/LSMO interface due to electro-mechanical boundary conditions at surface and interfaces. This method provides a unique nondestructive approach for quantifying subtle structural distortions in ferroelectric domains. Local nanoscale structural effects probed by this method can have profound impact on manipulation of domains in ferroelectric based devices. Additionally, this approach can be also extended to other complex oxides, and can be utilized to understand the effects of local strain and phase separation on strongly inter-coupled electronic, optical and structural properties at nanometer lengthscales.

Funding Information

The work at UC Davis was supported by National Science Foundation (DMR-1902652). This research used resources of the Advanced Photon Source and the Center for Nanoscale Materials, both U.S. Department of Energy (DOE) Office of Science User Facilities operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Research at the Center of Nanoscience and Nanotechnology was supported by a grant from the DIM OXYMORE (Ile-de-France) and by the RENATECH network.

References

- [1] M. Holt, R. Harder, R. Winarski, V. Rose, Nanoscale Hard X-ray Microscopy Methods for Materials Studies, *Annu. Rev. Mater. Res.* 43, 183-211 (2013)
- [2] J. Logan, R. Harder, L. Li, D. Haskel, P. Chen, R. Winarski, P. Fuesz, D. Schlager, D. Vine, C. Benson, I. McNulty, Hard X-ray polarizer to enable simultaneous three-dimensional nanoscale imaging of magnetic structure and lattice strain, *J. Synchrotron Radiation*. 23, 1210-1215 (2006)
- [3] S.O. Hruszkewycz, M.J. Highland, M. V. Holt, D. Kim, C.M. Folkman, C. Thompson, A. Tripathi, G.B. Stephenson, S. Hong, P.H. Fuoss, Imaging local polarization in ferroelectric thin films by coherent X-ray bragg projection ptychography, *Phys. Rev. Lett.* 110, 177601 (2013)

- [4] S. O. Hruszkewycz, M. V. Holt, M. Allain, V. Chamard, S. M. Polvino, C. E. Murray, P. H. Fuoss, Coherent Bragg nanodiffraction at the hard X-ray Nanoprobe beamline, *Philos. Trans. R. Soc.* 372, 20130118 (2014)
- [5] R.P. Winarski, M.V. Holt, V. Rose, P. Fuesz, D. Carbaugh, C. Benson, D. Shu, D. Kline, G.B. Stephenson, I. McNulty, J. Maser, A hard X-ray nanoprobe beamline for nanoscale microscopy, *J. Synchrotron Radiat.* 19, 1056-1060 (2012)
- [6] J.Y. Jo, P. Chen, R.J. Sichel, S. Baek, R.T. Smith, N. Balke, S. V Kalinin, M. V Holt, K. Evans-lutterodt, C. Eom, P.G. Evans, Structural Consequences of Ferroelectric Nanolithography, *Nano Lett.* 11, 3080-3084 (2011)
- [7] G. Rippey, L. Trinh, A. M. Kane, A. L. Ionin, M. S. Lee, R. V. Chopdekar, J. M. Christiansen-Salameh, D. A. Gilbert, A. J. Grutter, P. D. Murray, M. V. Holt, Z. Cai, K. Liu, Y. Takamura, R. Kukreja, X-ray nanodiffraction studies of ionically controlled nanoscale phase separation in cobaltites *Phys. Rev. Mater* 3, 082001(R) (2019).
- [8] A. Pagliero, L. Mino, E. Borfecchia, M. Truccato, A. Agostino, L. Pascale, E. Enrico, N. De Leo, C. Lamberti, G. Martínez-Criado, Doping change in the Bi-2212 superconductor directly induced by a hard X-ray nanobeam, *Nano Lett.* 14, 1583–1589 (2014)
- [9] Y. Zhu, Z. Cai, P. Chen, Q. Zhang, M.J. Highland, I.W. Jung, D.A. Walko, E.M. Dufresne, J. Jeong, M.G. Samant, S.S.P. Parkin, J.W. Freeland, P.G. Evans, H. Wen, Mesoscopic structural phase progression in photo-excited VO₂ revealed by time-resolved x-ray diffraction microscopy., *Sci. Rep.* 6, 21999 (2016)
- [10] Y. Zhao, G. Karypis, Data clustering in life sciences, *Mol. Biotechnol.* 31, 55–80 (2005)
- [11] Nugent R., Meila M. (2010) An Overview of Clustering Applied to Molecular Biology. In: Bang H., Zhou X., van Epps H., Mazumdar M. (eds) Statistical Methods in Molecular Biology. Methods in Molecular Biology (Methods and Protocols), Springer, Vol 620. Humana Press, Totowa, NJ
- [12] C.W. Hu, S.M. Kornblau, J.H. Slater, A.A. Qutub, Progeny clustering: A method to identify biological phenotypes, *Sci. Rep.* 5, 12894 (2015)
- [13] D. Fraix-Burnet, T. Chattopadhyay, A.K. Chattopadhyay, E. Davoust, M. Thuillard, A six-parameter space to describe galaxy diversification, *Astron. Astrophys.* 545, A80 (2012).
- [14] J. Sánchez Almeida, R. Terlevich, E. Terlevich, R. Cid Fernandes, A.B. Morales-Luis, Qualitative interpretation of galaxy spectra, *Astrophys J.* 756, 163 (2012)

- [15] J.S. Almeida, J.A.L. Aguerri, C. Muñoz-Tuón, A. De Vicente, Automatic unsupervised classification of all sloan digital sky survey data release 7 galaxy spectra, *Astrophys. J.* 714, 487–504 (2010)
- [16] MacQueen, J. Some methods for classification and analysis of multivariate observations, *Proceedings of the 5th Symposium Math. Statist. Prob.* 281–297 (1967)
- [17] Jain, A. K. and Dubes, R. C. *Algorithms for Clustering Data*, Prentice Hall, 1998.
- [18] J.D. Simpson, P.L. Cottrell, C.C. Worley, Spectral matching for abundances and clustering analysis of stars on the giant branches of ω Centauri, *Mon. Not. R. Astron. Soc.* 427, 1153–1167 (2012)
- [19] L.G. Balazs, A.T. Garibjanyan, L.V. Mirzoyan, V.V. Hambaryan, M. Kun, A. Fronto, J. Keleman, Medium resolution spectroscopy of Ha objects in IC 1396, *Astron. Astrophys.* 311, 145–151 (1996)
- [20] I. Ordovás-Pascual, J. Sánchez Almeida, A fast version of the k-means classification algorithm for astronomical applications, *Astron. Astrophys.* 565, A53 (2014)
- [21] M. Haralick, L. Robert, Image Segmentation Techniques, *Computer Vision, Graphics, and Image Processing* 29, 100-132 (1985).
- [22] L. Juang, M. Wu, MRI brain lesion image detection based on color-converted K-means clustering segmentation, *Measurement* 43, 941-949 (2010)
- [23] X. Zheng, Q. Lei, R. Yao, Y. Gong, Q. Yin, Image segmentation based on adaptive K-means algorithm, *Image Video Proc.* 68, 7578 (2018)
- [24] J.F. Scott, “Applications of modern ferroelectrics”, *Science* 315, 949-959 (2007)
- [25] J.F. Scott, “Nanoferroelectrics: static and dynamics”, *J. Phys. Condens. Matter* 18, R361-386 (2006)
- [26] M. Dawber, K.M. Rabe and J.F. Scott, ‘Physics of thin-film ferroelectric oxides’, *Rev. Mod. Phys.* 77, 1083-1130 (2005)
- [27] N. Setter, D. Damjanovic, L. Eng, G. Fox, S. Gevorgian, S. Hong, A. Kingon, H. Kohlstedt, N. Y. Park, G. B. Stephenson, I. Stolitchnov, A. K. Taganstev, D. V. Taylor, T. Yamada, S. Streiffer, Ferroelectric thin films: Review of materials, properties, and applications, *Journal of Applied Physics*, 100, 051606 (2006).
- [28] S. Trolier-McKinstry, P. Muralt, Thin Film Piezoelectrics for MEMS, *Journal of Electroceramics* 12, 7-17 (2004)

- [29] J. Hoffman, X. Pan, J. W. Reiner, F. J. Walker, J. P. Han, C. H. Ahn, T. P. Ma, Ferroelectric field effect transistors for memory applications, *Adv. Mater.* 22, 2957-61 (2010)
- [30] J.A. Sulpizio, S. Ilani, P. Irvin, J. Levy, Nanoscale Phenomena in Oxide Heterostructures, *Annu. Rev. Mater. Res.* 44, 117 (2014)
- [31] P. Gao, J. Britson, C.T. Nelson, J.R. Jokisaari, C. Duan, M. Trassin, S. Baek, H. Guo, L. Li, Y. Wang, Y. Chu, A.M. Minor, C. Eom, R. Ramesh, L. Chen, X. Pan, Ferroelastic domain switching dynamics under electrical and mechanical excitations, *Nat. Commun.* 5, 3801 (2014)
- [32] J.A. Mundy, J. Schaab, Y. Kumagai, A. Cano, M. Stengel, I.P. Krug, D.M. Gottlob, H. Doğanay, M.E. Holtz, R. Held, Z. Yan, E. Bourret, C.M. Schneider, D.G. Schlom, D.A. Muller, R. Ramesh, N.A. Spaldin, D. Meier, Functional electronic inversion layers at ferroelectric domain walls, *Nat. Mater.* 16, 622–627 (2017)
- [33] J. C. Agar, A. R. Damodaran, M. B. Okatan, J. Kacher, C. Gammer, R. K. Vasudevan, S. Pandya, R. V. K. Mangalam, G. A. Velarde, S. Jesse, N. Balke, A. M. Minor, S. V. Kalinin, L. W. Martin, Highly-mobile ferroelastic domain walls in compositionally-graded ferroelectric thin films, *Nat. Mater.* 15, 549–556 (2016)
- [34] S. O. Hruszkewycz, M. V. Holt, M. Allain, V. Chamard, S. M. Polvino, C. E. Murray, P. H. Fuoss, Efficient modeling of Bragg coherent x-ray nanobeam diffraction, *Opt. Lett.* 40, 3241-4 (2015)
- [35] Dal-Hyun Do, Paul G. Evans, Eric D. Isaacs, Dong Min Kim, Chang Beom Eom & Eric M. Dufresne, Structural visualization of polarization fatigue in epitaxial ferroelectric oxide devices, *Nature Materials* 3, 365-369 (2004)
- [36] Alexei Grigoriev, Dal-Hyun Do, Dong Min Kim, Chang-Beom Eom, Bernhard Adams, Eric M. Dufresne, Paul G. Evans, Nanosecond domain wall dynamics in ferroelectric Pb(Zr, Ti)O₃ thin films, *Phys. Rev. Lett.* 96, 187601 (2006)
- [37] MathWorks, k-means (2019). Available: <https://www.mathworks.com/help/stats/k-means.html>
- [38] Lloyd, Stuart P. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory* 28, 129–137 (1982)
- [39] Roberto Lo Conte, Zhuyun Xiao, Cai Chen, Camelia V. Stan, Jon Gorchon, Amal El-Ghazaly, Mark E. Nowakowski, Hyunmin Sohn, Akshay Patabi, Andreas Scholl, Nobumichi Tamura, Abdon Sepulveda, Gregory P. Carman, Robert N. Candler, Jeffrey Bokor, Influence of

Figures

Figure 1

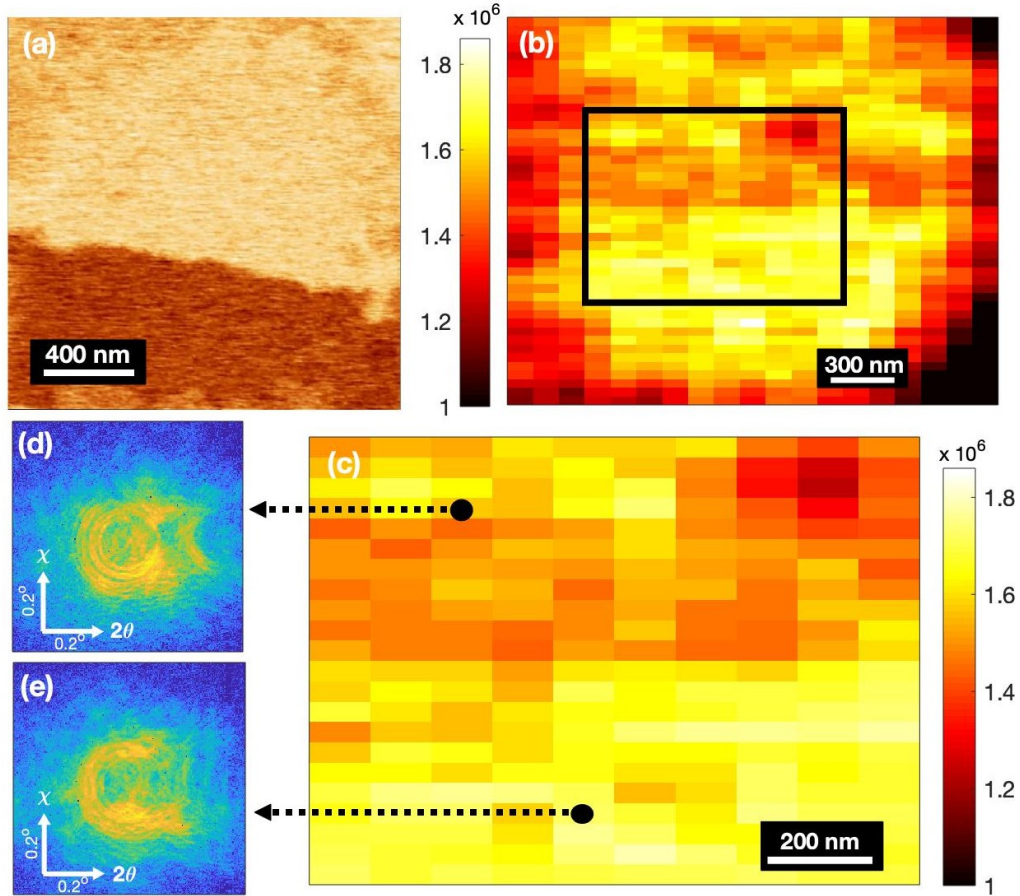


Figure 1: a) PFM phase image of one PZT 2 μm x 2 μm island showing up (light brown) and down (dark brown) domain, b) nanodiffraction map over the same region as a) using integrated intensity of the diffraction pattern obtained near the (002) Bragg peak. A clear difference in intensity is observed between top and bottom of the image due to up and down domain, c) nanodiffraction map of a selected region in the box which was used for further analysis, d) and e) detector image showing diffraction pattern for up and down domain respectively.

Figure 2

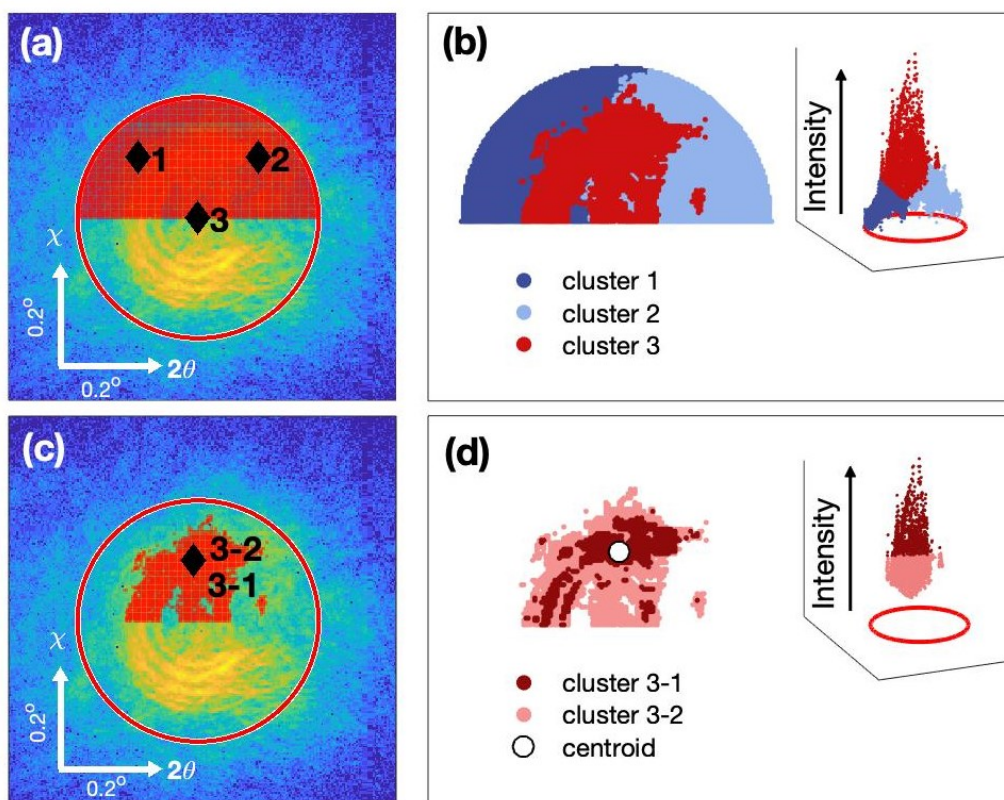


Figure 2: Cluster algorithm for identifying a local maximum within a ROI: a) initial guesses for clusters within ROI, b) resulting clusters and segmentation of diffraction data, c) further refinement of highest intensity cluster obtained in b) by using cluster algorithm, the initial guesses are also shown; d) resulting clusters and further segmentation of data. Centroid of cluster 3-1 is the local maxima of ROI.

Figure 3

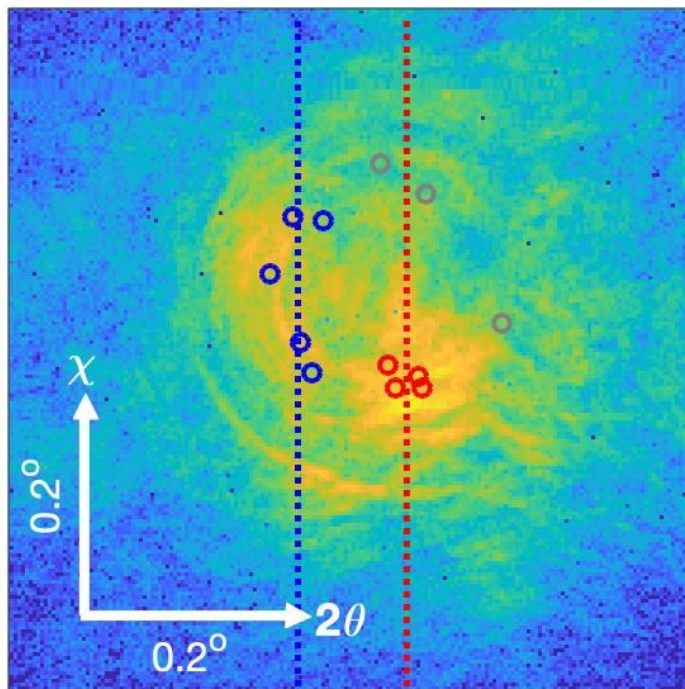


Figure 3: Local maxima (or centroids) obtained for twelve different ROIs. Centroids below 40% intensity threshold are discarded (gray circles). Centroids above this threshold can be grouped in two groups (blue and red) using k-means. X-location of these two groups is shown using red and blue dashed line.

Figure 4

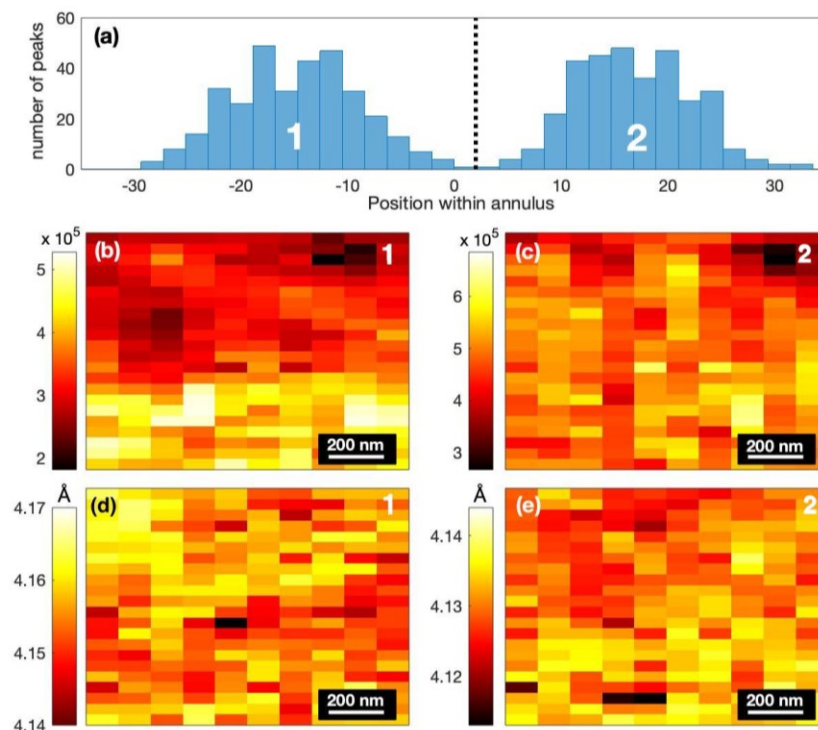


Figure 4: a) Histogram of local maxima location for all sample positions showing a bimodal distribution, b) and c) integrated intensity map for distribution 1 and 2 respectively, and d) and e) lattice parameter map for distribution 1 and 2 respectively.