

PAPER • OPEN ACCESS

Propagation of priors for more accurate and efficient spectroscopic functional fits and their application to ferroelectric hysteresis

To cite this article: N Creange *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 045002

View the [article online](#) for updates and enhancements.

You may also like

- [The \(2+1\)-dimensional nonisospectral relativistic Toda hierarchy related to the generalized discrete Painlevé hierarchy](#)
Zuo-Nong Zhu
- [Hierarchy Optimization: A Means to Enhance Efficiency in E-Beam and Optical Lithography](#)
Anja Rosenbusch K. Kalus, Hiroyuki Endo
Hiroyuki Endo, Yasuki Kimura Yasuki Kimura et al.
- [BKP and CKP revisited: the odd KP system](#)
Aristophanes Dimakis and Folkert Müller-Hoissen



PAPER

OPEN ACCESS

RECEIVED
21 January 2021REVISED
5 April 2021ACCEPTED FOR PUBLICATION
26 April 2021PUBLISHED
15 July 2021

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Propagation of priors for more accurate and efficient spectroscopic functional fits and their application to ferroelectric hysteresis

N Creange¹ , K P Kelley¹ , C Smith¹, D Sando³, O Paull³, N Valanoor³, S Somnath², S Jesse¹, S V Kalinin¹ and R K Vasudevan^{1,*}

¹ Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States of America

² National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States of America

³ School of Material Science and Engineering, University of New South Wales, Sydney, NSW 2052, Australia

* Author to whom any correspondence should be addressed.

E-mail: vasudevanrk@ornl.gov

Keywords: piezoresponse force microscopy, ferroelectrics, spatial priors

Abstract

Multi-dimensional spectral-imaging is a mainstay of the scanning probe and electron microscopies, micro-Raman, and various forms of chemical imaging. In many cases, individual spectra can be fit to a specific functional form, with the model parameter maps, providing direct insight into material properties. Since spectra are often acquired across a spatial grid of points, spatially adjacent spectra are likely to be similar to one another; yet, this fact is almost never used when considering parameter estimation for functional fits. On datasets tried here, we show that by utilizing proximal information, whether it be in the spatial or spectral domains, it is possible to improve the reliability and increase the speed of such functional fits by $\sim 2\text{--}3\times$, as compared to random priors. We explore and compare three distinct new methods: (a) spatially averaging neighborhood spectra, and propagating priors based on functional fits to the averaged case, (b) hierarchical clustering-based methods where spectra are grouped hierarchically based on response, with the priors propagated progressively down the hierarchy, and (c) regular clustering without hierarchical methods with priors propagated from fits to cluster means. Our results highlight that utilizing spatial and spectral neighborhood information is often critical for accurate parameter estimation in noisy environments, which we show for ferroelectric hysteresis loops acquired on a prototypical PbTiO_3 thin film with piezoresponse spectroscopy. This method is general and applicable to any spatially measured spectra where functional forms are available. Examples include exploring the superconducting gap with tunneling spectroscopy, using the Dynes formula, or current–voltage curve fits in conductive atomic force microscopy mapping. Here we explore the problem for ferroelectric hysteresis, which, given its large parameter space, constitutes a more difficult task than, for example, fitting current–voltage curves with a Schottky emission formula (Chiu 2014 *Adv. Mater. Sci. Eng.* **2014** 578168).

1. Introduction

Spectroscopy enables essential insights into a wide range of fundamental processes ranging from astronomical to sub-atomic scales, and is utilized in a spectrum of methods such as hyperspectral imaging from satellites of Earth for identifying vegetation health [1–3], electron energy loss spectroscopy for elemental analysis in electron microscopy [4–6], and micro-Raman imaging for understanding symmetry breaking and structure in nanomaterials [7–10]. When implemented in an imaging framework, such spectral imaging enables multiple channels of data to be collected across a spatial grid of points, allowing correlations between functional properties and specific microstructural features to be studied.

In some cases, tractable models exist to describe individual spectra and the spectra can be fit to such models. Additionally, it is possible to perform inference on the given model parameters. Typical examples include the Dyne's functional equation for measurements of the superconducting gap in scanning tunneling spectroscopy [11], various functional relationships describing current–voltage relationships in oxides [12], and equations that describe features of hysteresis loops captured on ferroelectric materials with piezoresponse force spectroscopy (PFS) techniques [13]. Analysis in most cases of hyperspectral data of this sort is split into two general methods: (a) dimensionality reduction, often for visualization purposes [14–16], typically with some form of matrix [17] or tensor factorization [18] approach, and perhaps subsequent fitting of individual components of the spectra [19], or (b) individual fitting of spectra, independently, to tractable models where available. Option (a) is highly useful when no such models exist, or when the basic visualization of a feature's spatial variability in a given dataset is sought after. On the other hand, quantitative estimation of materials parameters usually requires inference of model parameters. The difficulty of applying (b) is, however, that the data, especially in many microscopy settings, is often noisy. Moreover, it is readily apparent that many spectra can be remarkably similar to one another, either due to spatial location (nearest neighborhood), or due to the generic properties of materials, such as segregation of responses by different phase mixtures and so on. As such, using independent functional fits from pixel to pixel that do not leverage this information is likely to lead to sub-optimal results. Previously, we have shown that even for simple four-parameter models such as that of a simple harmonic oscillator, standard least-squares based estimation can lead to poor results for noisy data, that can be alleviated via the use of trained artificial neural networks [20]. In that work, it was shown that utilizing the neural network to produce the prior, that was then ingested by the least squares optimizer, led to far superior results, increasing the signal-to-noise extraction by nearly an order of magnitude. The same type of use of neural networks for estimating priors to fit into more traditional optimization routines has also been used for e.g. in x-ray scattering for phase reconstructions [21]. Achieving better results from standard optimization routines is a straightforward way to improve the resolution of the derived fitting parameters maps from imaging studies. In principle, the resolution can be further enhanced via methods such as pan-sharpening for even greater impact [22].

Here, we explore three strategies to better inform the prior used in functional fits for hysteresis loop acquisition in PFS. These strategies utilize the averaging principle, where averaging like-spectra (spatially, or based on clustering) and fitting to the averages, can provide better priors for individual spectra, and leading to both better model parameter inference, as well as improved efficiency. We show our results on both synthetic and real datasets, focusing on a nine-parameter phenomenological model based on locally acquired hysteresis loops, to highlight the utility of this approach. Our results indicate that utilizing a nearest-neighbor or K-means clustering strategy not only averages ~50% less time for computation, but substantially increases the quality of the individual spectral fits compared to not using priors from similar pixels.

2. Main strategies

In PFS, acquired loops consist of two branches, a positive branch (PR^+) and a negative branch (PR^-), that can be fit via a phenomenological expression based on error functions, specifically:

$$PR^+(V) = a_0 + a_1 \times \left(g_1 \times \frac{\operatorname{erf}\left(\frac{v_1 - a_2}{g_1} + b_0\right)}{(b_0 + b_1)} \right) + a_4 \times v_1 \quad (1)$$

$$PR^-(V) = a_0 + a_1 \times \left(g_2 \times \frac{\operatorname{erf}\left(\frac{v_2 - a_3}{g_2} + b_2\right)}{(b_2 + b_3)} \right) + a_4 \times v_2 \quad (2)$$

where $g_1 = \frac{b_1 - b_0}{2 \times \operatorname{erf}\left(\frac{v_1 - a_2}{(v_1 - a_2) \times 1000} + 1\right)} + b_0$ and $g_2 = \frac{b_3 - b_2}{2 \times \operatorname{erf}\left(\frac{v_2 - a_3}{(v_2 - a_3) \times 1000} + 1\right)} + b_2$, $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \cdot \int_z^{t=0} \exp(-t^2)$ refers to the error function and v_1 and v_2 are the voltage values at the midpoint of the up and down voltage ramps, respectively. In the above case, the coefficients $a_0 \dots a_4$ and $b_0 \dots b_3$ are then used to calculate physical loop parameters such as nucleation threshold, imprint, work of switching, and switchable polarization [13, 23, 24]. The loss function is computed in terms of both branches of the loop, i.e. computing equations (1) and (2) simultaneously with the individual hysteresis loop. Unfortunately, the complicated nature of this expression tends towards difficult initialization procedures and loss landscapes with multiple local minima that are routinely far from the global optimal conditions.

We compare four distinct strategies for improving the fitting of hysteresis loops acquired using PFS: 'nearest neighbor', 'clustering', 'hierarchical', and 'random'. To evaluate the effectiveness of each strategy, a

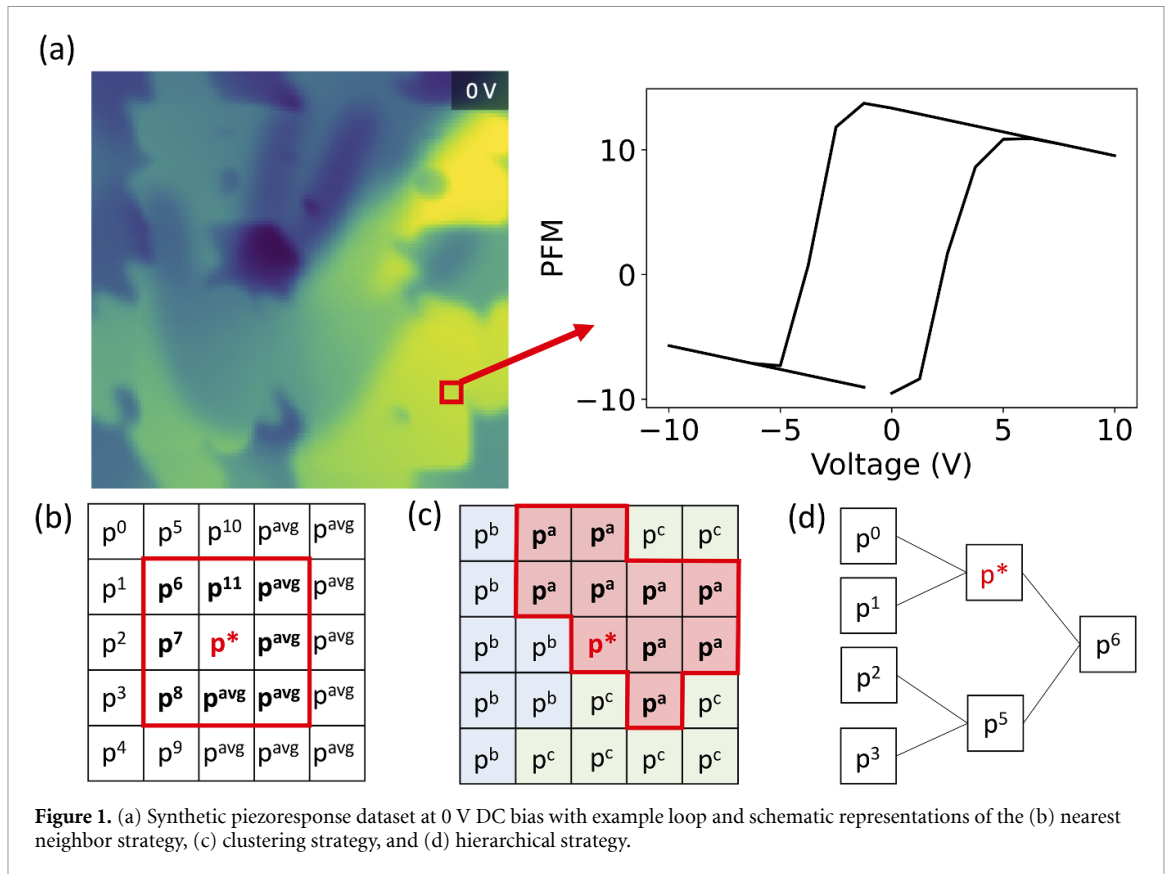
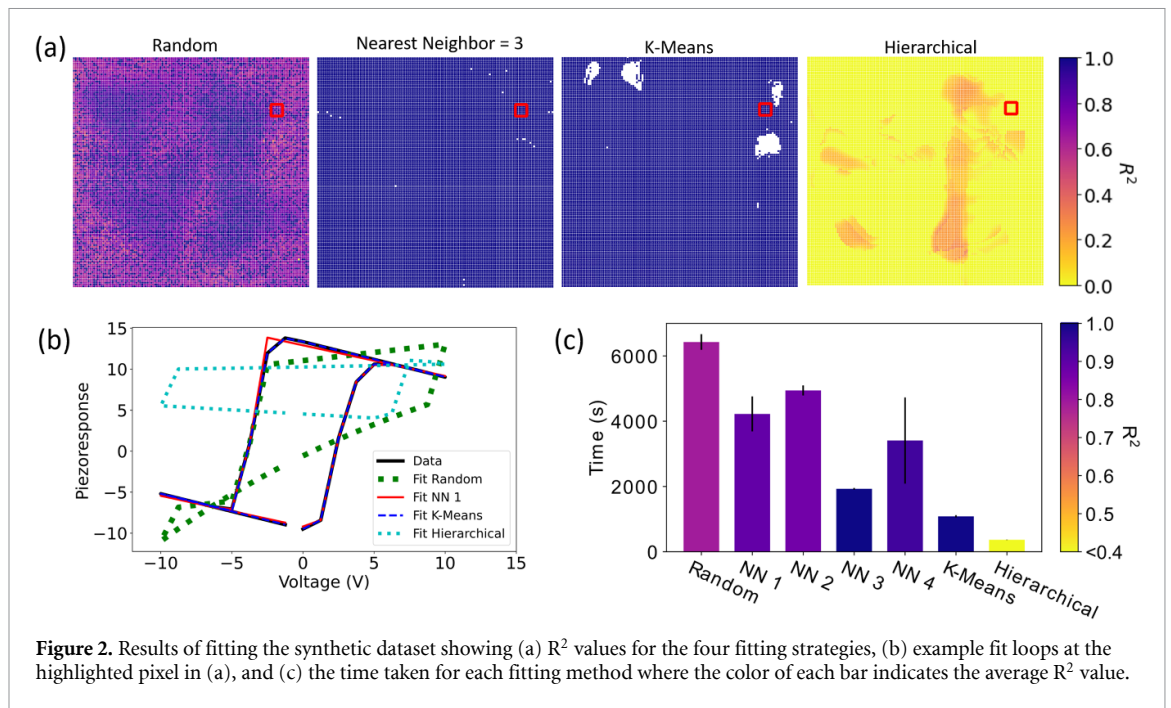


Figure 1. (a) Synthetic piezoresponse dataset at 0 V DC bias with example loop and schematic representations of the (b) nearest neighbor strategy, (c) clustering strategy, and (d) hierarchical strategy.

synthetic dataset of size 128×128 pixels was generated using equations (1) and (2) and nine spatial maps, derived from different images, such that the loop parameters vary spatially. An example of the piezoresponse, defined as the ferroelectric response at 0 V DC bias, from this simulated dataset, at 0 V DC bias, is shown in figure 1(a). The loop at this condition is rather typical of ferroelectric loops acquired during the so-called on-field setup, i.e. when the DC bias is active during the piezoresponse measurement, and the slope of the loop usually stems from the electrostatic contribution to the signal.

Multiple strategies can be envisioned to obtain better priors based on the idea of fitting to a mean of ‘similar’ pixels to derive the prior for the present fit. One of the simplest is perhaps the ‘nearest neighbor’ strategy, which averages the surrounding N nearest neighbors around the pixel of interest, p^* . The number of neighboring pixels can be chosen to adjust for different feature length scales, where $N = 1$ would consider only the 1st shell of surrounding pixels, schematically shown in figure 1(b), and $N = 2$ would consider the 1st and 2nd shell of surrounding pixels. Prior to any fitting, the average loop found and fit 20 consecutive times using non-linear least squares, with each fit starting at a random initial guess. Of the 20 fits, the fit with the lowest sum of squares error is chosen and each pixel is initially assigned the average loop fit coefficients, denoted as p^{avg} in figure 1(b). At the pixel of interest, p^* , the coefficients from the surrounding pixels are averaged and used as the initial guess for fitting the loop at p^* . After a successful fit, the coefficients at pixel p^* are updated to the coefficients found after fitting. Schematically in figure 1(b), the indicated p^* would be replaced with p^{12} . The next pixel is then considered, moving linearly through each row and column until each pixel has been fit. This strategy is likely to work well for samples that display high degrees of local spatial homogeneity. On the other hand, it is also possible that similar responses are not necessarily correlated in this manner. For instance, meandering domain walls which are only one or two pixels wide in the dataset would have similar response across the domain wall, but this usually traces out a non-trivial path through the images.

In contrast, clustering methods can help to identify these pixels, and can subsequently be used as the initial seed, as shown schematically in figure 1(c). In the ‘clustering’ strategy, the loops are K-Mean clustered into $M/100$ clusters, where M is the total number of pixels. The average loop of each cluster is then fit using non-linear least squares 20 times, each time with a random initial guess of coefficients. Fits at the pixel of interest, p^* , then uses the coefficients of the average loop for its corresponding cluster. This strategy helps resolve fitting loops which have a low spatial correlation to the surrounding pixels. One challenge with this method is that the number of pixels in each cluster varies, often substantially. This can lead to poorer priors for clusters with comparatively small numbers of members, due to insufficient averaging.

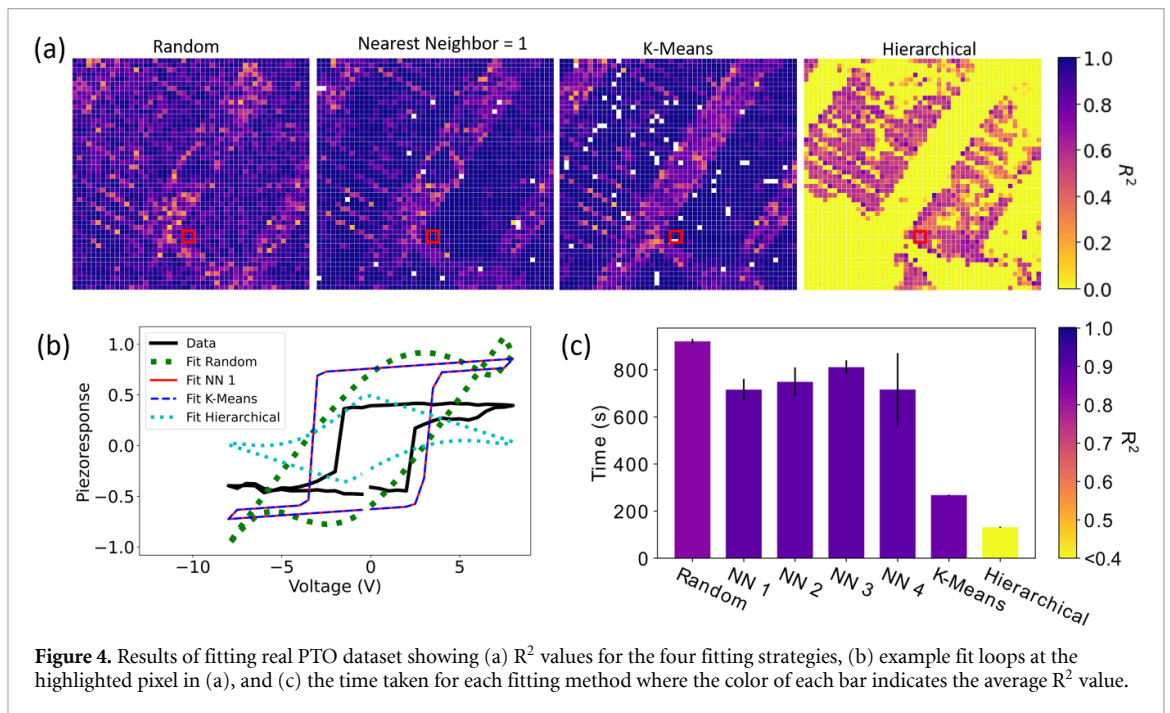
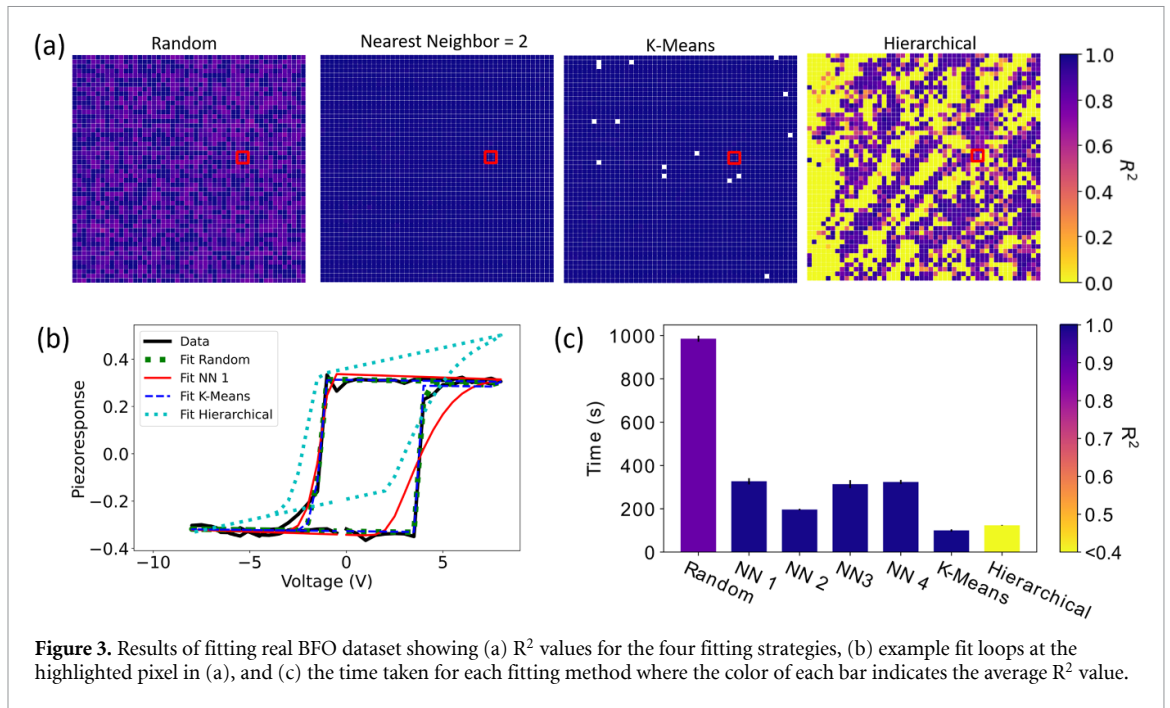


The third strategy involves hierarchical clustering, where fits are propagated down the hierarchical tree, schematically shown in figure 1(d). The principle is similar to the previous clustering method but differs in that the clustering itself is hierarchical in nature, i.e. the data is repeatedly split to form a hierarchical representation of similarity. One advantage of this approach can be that at the top of the tree, the numbers of loops in every node is very high, leading to high signal-to-noise ratios and presumably better fits; however, it drops substantially down the tree, which could again cause issues with insufficient averaging, especially if the data is noisy and similarity between nodes on the same level is weak. Finally, we also tested against a ‘random’ strategy, which uses random coefficients as the initial guess for fitting. Coefficients are each chosen at random within an $N(0.1,5)$ distribution. The fit is performed twice at each pixel of interest, each with different initial guesses. Below, we compare the performance of these three methods, as compared with random initialization.

3. Synthetic dataset results

We compare R^2 results for the random guesses, nearest neighbor ($N = 3$), K-means, and hierarchical clustering strategies, as shown in figure 2(a). For the nearest neighbor method, the number of shells taken into account during averaging was varied and only the best performing nearest neighbor method is shown in figures 2(a) and (b). It can be seen that in this instance, the hierarchical method completely fails with a very low average R^2 value. Random guessing fairs better, but still has a mean R^2 map of 0.825. Meanwhile, the best performing methods are the nearest-neighbor and K-means strategies, which indicate almost perfect results, with mean R^2 values of 0.999 and 0.998 respectively. However, it should be noted that certain pixels are present for which no optimization converged with these initializations; these are the white pixels in figure 2(a). This dropout occurs when the initial guesses of the coefficients deviate too far from a minima and thus the maximum iterations of the non-linear least squares algorithm is reached, or alternatively, the proposed coefficient value is outside of the specified bounds.

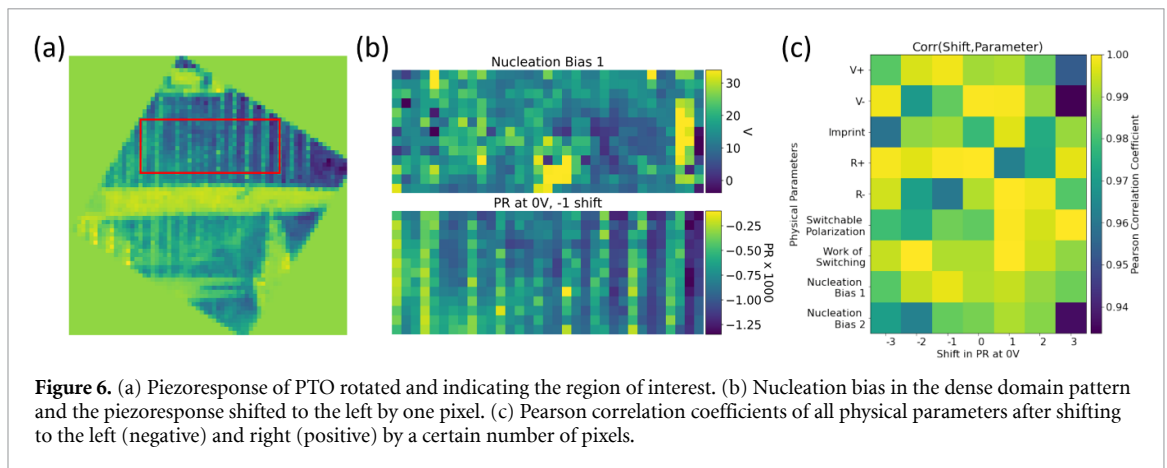
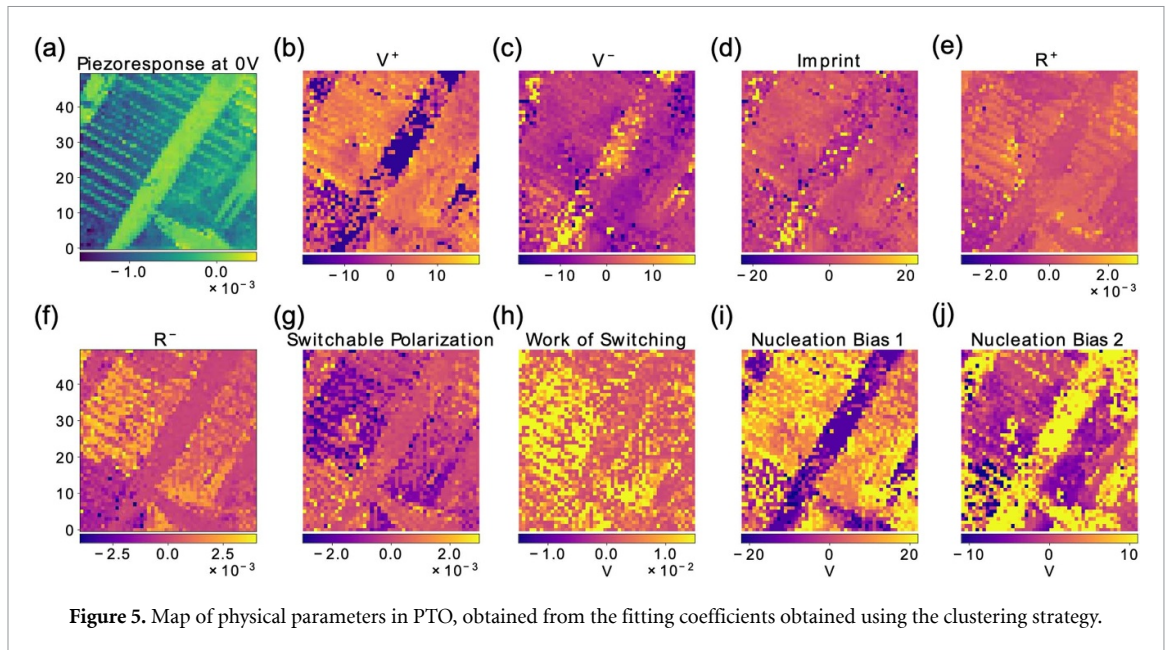
An example of the loop fits using the different methods are shown in figure 2(b) where the corresponding pixel is outlined in red (figure 2(a)). The hierarchical method and random initialization both perform poorly. The time taken for each method to fit the simulated dataset is shown in figure 2(c) where the error bars indicate three separate runs for statistics, and the color signifies the average R^2 value. As such, the random initialization takes the longest, whereas nearest neighbors takes less time, whilst K-means and hierarchical methods take the least amount of time. The time savings are substantial, between 50% and 80% compared with the random seeds. This is likely because the good initial estimates for the parameters reduce the number of iterations for the least squares optimizer to arrive at a value within the prescribed tolerance value for optimization termination. The reason for the very small time taken for hierarchical clustering might be that the initial guesses were extremely close to a sub-optimal loss minimum. In that case the time taken to run the optimization would be very small, but the result would be highly inaccurate, which appears



to have been the case here. Additionally, the inaccurate results may have been caused by a few bad fits high up on the tree, corrupting the fits which use that corrupted branch. As for the nearest neighbor method, it can be seen in figure 2(c) that $N = 3$ performs sufficiently better than $N = 1$ and $N = 2$. This is likely to be due to the length scale of features in the synthetic data set, which are quite large, thus a consideration of a larger number of similar pixels, producing an initial guess which is likely to be closer to the global minima.

4. Experimental results

Bismuth ferrite (BFO) and lead titanate (PTO) are two well-explored ferroelectrics [25] with two different domain structures owing to their distinct crystal structures (rhombohedral vs. tetragonal), and are here used as model systems to explore generalizability of the four fitting strategies. PFS was performed on both materials using an Asylum Cypher AFM, resulting in 50×50 pixel spectral maps for each material. All four fitting strategies were performed in an identical manner as the synthetic dataset with the results for BFO and PTO shown in figures 3 and 4, respectively. A similar trend in R^2 and total fitting time are seen for BFO and



PTO loops, the K-means based clustering strategy is the fastest and contains the highest R^2 values in both data sets. Due to this consistency, the K-means based clustering strategy reveals that it can be generalizable to a wide range of feature sizes and obtains accurate fits. It should be noted, the R^2 are worse for the PTO case for all methods; this is expected, however, given that the spatial map indicates lower values within the large in-plane domains which are not expected to display switching behavior in a vertical PFS measurement, neglecting shear effects [26].

For measurements taken on BFO, the loops show an almost ideal behavior; thus, the nearest neighbor and clustering strategies work seemingly well with R^2 values above 0.993. In stark contrast, the loops in PTO are less ideal, as compared to BFO, and thus the R^2 values trend lower than those in BFO. However, due to the high level of detail in the spatial maps for PTO, we can now inspect the switching near a dense ferroelastic structure where the domain width is only a few pixels. Using the obtained fits from the clustering strategy, the coefficients are used to calculate the physical parameters, namely the imprint, coercive biases (V^+ , V^-), remnant polarizations (R^+ , R^-), switchable polarization, work of switching, and nucleation bias, shown in figure 5 [13]. A clear distinction can be seen between the a- and c-domains in all parameters which follows the domain pattern seen in the PFS maps.

Regions of dense a-c domain structure, like the region observed in the upper left-hand corner of the image, are expected to contain interesting physics [27–29]. Thus, the region is isolated and rotated, as shown in figure 6(a), to allow for easier mapping and further correlation analysis. The correlation between physical parameters is explored using Pearson correlation coefficients, where a coefficient of 1 indicates perfect correlation [30]. The piezoresponse at 0 V is shifted to the left and right and compared to each physical parameter map, as shown in figure 6(b), to identify the shift which produces the maximum correlation, where a negative shift indicates a shift to the left. Some parameters such as work of switching seem to vary

less with shifting the piezoresponse map, as those parameters are likely less sensitive to domain orientation. Conversely, imprint and nucleation bias seem more sensitive to the shift as they are very sensitive to the specific geometry of the domains [13, 18]. Shifting the parameters reveal a maximum correlation with a piezoresponse shift of one pixel to the right, indicating that the maxima and minima are located to the right of the domain wall. This may be due to wall twist effect lowering nucleation bias at some finite distance [31].

5. Conclusions

We have shown three strategies for fitting PFS loops which takes into account the surrounding loop information. It was shown overall that taking priors from similar loops increases the accuracy of the fit and reduces the processing time. Using the clustering strategy provided the highest R^2 values, as compared to the other strategies, for all scenarios presented. In addition, clustering is less sensitive to the feature size and domain shape, which is not the case for the nearest neighbor strategy. By obtaining accurate PFS loop fits, quantitative physical parameters can be determined, and interesting physics can be extracted. The developed strategies for prior determination are broadly applicable to many fields, further increasing the speed and accuracy of any fitting process. In addition, these methods could be combined with neural networks for prior generation to further improve the spatial resolution of multi-dimensional hyperspectral datasets.

Data availability statement

The code used here is openly available via the Github repository BGLib, a part of the pycroscopy ecosystem [32]. The data that support the findings of this study are openly available at the following URL/DOI: [10.13139/ORNLNCCS/1761194](https://doi.org/10.13139/ORNLNCCS/1761194).

Acknowledgments

This work (fitting algorithm) was performed and supported (N C, R K V, S V K) at the Oak Ridge National Laboratory's Center for Nanophase Materials Sciences (CNMS), a US Department of Energy, Office of Science User Facility. A portion of the work (PFM data acquisition, K P K) is supported by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), Materials Sciences and Engineering Division. Partial support for sample synthesis was provided by Professor Hiroshi Funakubo at the Tokyo Institute of Technology.

Materials

The BiFeO₃ investigated here consisted of a 20 nm thick thin film deposited via pulsed laser deposition, grown on a conductive 2 nm thick La_{0.67}Sr_{0.33}MnO₃ back electrode and LaAlO₃ (310) substrate, using conditions previously reported [33, 34]. Additional details on sample preparation and properties can be found in work published by Paull *et al* [35]. The polydomain PbTiO₃ investigated here consisted of a 700 nm pulsed metal-organic chemical vapor deposited thin film grown on KTaO₃ (100) with a SrTiO₃ back electrode. Further details concerning the PbTiO₃ synthesis can be found in work published by Nakashima [36].

ORCID iDs

N Creange  <https://orcid.org/0000-0003-2764-7384>

K P Kelley  <https://orcid.org/0000-0002-7688-0484>

R K Vasudevan  <https://orcid.org/0000-0003-4692-8579>

References

- [1] Jay S, Baret F, Dutartre D, Malatesta G, Héno S, Comar A, Weiss M and Maupas F 2019 *Remote Sens. Environ.* **231** 110898
- [2] Yue J, Yang G, Tian Q, Feng H, Xu K and Zhou C 2019 *ISPRS J. Photogramm. Remote Sens.* **150** 226–44
- [3] Li Z, Li Z, Fairbairn D, Li N, Xu B, Feng H and Yang G 2019 *Comput. Electron. Agric.* **162** 174–82
- [4] Pennycook S J and Boatner L A 1988 *Nature* **336** 565–7
- [5] Batson P E 1993 *Nature* **366** 727–8
- [6] Campos A, Troc N, Cottancin E, Pellarin M, Weissker H C, Lermé J, Kociak M and Hillenkamp M 2019 *Nat. Phys.* **15** 275–80
- [7] Pezzotti G 2013 *J. Appl. Phys.* **113** 211301
- [8] Ryu S K, Zhao Q, Hecker M, Son H Y, Byun K Y, Im J, Ho P S and Huang R 2012 *J. Appl. Phys.* **111** 063513
- [9] De Wolf I 1999 *J. Raman Spectrosc.* **30** 877–83
- [10] Colomban P 2002 *Adv. Eng. Mater.* **4** 535–42

- [11] Dynes R C, Narayanamurti V and Garno J P 1978 *Phys. Rev. Lett.* **41** 1509–12
- [12] Chiu F C 2014 *Adv. Mater. Sci. Eng.* **2014** 578168
- [13] Jesse S, Lee H N and Kalinin S V 2006 *Rev. Sci. Instrum.* **77** 73702
- [14] Jesse S and Kalinin S V 2009 *Nanotechnology* **20** 085714
- [15] Bonnet N 1998 *J. Microsc.* **190** 2–18
- [16] Bonnet N 2000 *Advances in Imaging and Electron Physics* vol 114, ed P W Hawkes (San Diego: Academic) pp 1–77
- [17] Kannan R, Ievlev A V, Laanait N, Ziatdinov M A, Vasudevan R K, Jesse S and Kalinin S V 2018 *Adv. Struct. Chem. Imaging* **4** 6
- [18] Kelley K P et al 2020 *npj Comput. Mater.* **6** 113
- [19] Strelcov E, Belianinov A, Hsieh Y H, Jesse S, Baddorf A P, Chu Y H and Kalinin S V 2014 *ACS Nano* **8** 6449–57
- [20] Borodinov N, Neumayer S, Kalinin S V, Ovchinnikova O S, Vasudevan R K and Jesse S 2019 *npj Comput. Mater.* **5** 1–8
- [21] Chan H, Nashed Y S G, Kandel S, Hruszkewycz S, Sankaranarayanan S, Harder R J and Cherukara M J 2020 Real-time 3D nanoscale coherent imaging via physics-aware deep learning (arXiv:2006.09441v1)
- [22] Borodinov N, Bilkey N, Foston M, Ievlev A V, Belianinov A, Jesse S, Vasudevan R K, Kalinin S V and Ovchinnikova O S 2019 *npj Comput. Mater.* **5** 1–9
- [23] Jesse S, Baddorf A P and Kalinin S V 2006 *Appl. Phys. Lett.* **88** 62908
- [24] Jesse S, Maksymovych P and Kalinin S V 2008 *Appl. Phys. Lett.* **93** 112903
- [25] Tagantsev L E C A K and Fousek J 2010 *Domains in Ferroic Crystals and Thin Films* (Berlin: Springer) (<https://doi.org/10.1007/978-1-4419-1417-0>)
- [26] Guyonnet J, Béa H, Guy F, Gariglio S, Fusil S, Bouzehouane K, Triscone J M and Paruch P 2009 *Appl. Phys. Lett.* **95** 132902
- [27] Wicks S, Seal K, Jesse S, Anbusathaiah V, Leach S, Garcia R E, Kalinin S V and Nagarajan V 2010 *Acta Mater.* **58** 67–75
- [28] Ivry Y, Chu D P, Scott J F, Salje E K H and Durkan C 2011 *Nano Lett.* **11** 4619–25
- [29] Balke N, Gajek M, Tagantsev A K, Martin L W, Chu Y H, Ramesh R and Kalinin S V 2010 *Adv. Funct. Mater.* **20** 3466–75
- [30] Pedregosa F et al 2011 *J. Mach. Learn. Res.* (Available at: <https://jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>)
- [31] Aravind V R et al 2010 *Phys. Rev. B* **82** 1–11
- [32] Somnath S, Smith C R, Lannait N, Vasudevan R K, Ievlev A V, Belianinov A, Lupini A R, Shankar M, Kalinin S V and Jesse S 2019 (arXiv:1903.09515)
- [33] Sando D, Young T, Bulanadi R, Cheng X, Zhou Y, Weyland M, Munroe P and Nagarajan V 2018 *Japan. J. Appl. Phys.* **57** 0902B2
- [34] Sando D et al 2020 *Adv. Funct. Mater.* **30** 2000343
- [35] Paull O et al 2021 Super-R BiFeO₃: epitaxial stabilization of a low-symmetry phase with giant electromechanical response (arXiv:2102.00558)
- [36] Nakashima T, Ichinose D, Ehara Y, Shimizu T, Kobayashi T, Yamada T and Funakubo H 2017 *Appl. Phys. Lett.* **110** 122902