



**HAL**  
open science

# Enhancing ferroelectric characterization at nanoscale: A comprehensive approach for data processing in spectroscopic piezoresponse force microscopy

H. Valloire, P. Quéméré, N. Vaxelaire, H. Kuentz, G. Le Rhun, Ł. Borowik

## ► To cite this version:

H. Valloire, P. Quéméré, N. Vaxelaire, H. Kuentz, G. Le Rhun, et al.. Enhancing ferroelectric characterization at nanoscale: A comprehensive approach for data processing in spectroscopic piezoresponse force microscopy. *Journal of Applied Physics*, 2024, 135 (19), 10.1063/5.0197226 . cea-04770018

HAL Id: cea-04770018

<https://cea.hal.science/cea-04770018v1>

Submitted on 6 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE | MAY 17 2024

# Enhancing ferroelectric characterization at nanoscale: A comprehensive approach for data processing in spectroscopic piezoresponse force microscopy

H. Valloire ; P. Quéméré ; N. Vaxelaire ; H. Kuentz ; G. Le Rhun ; Ł. Borowik 



*J. Appl. Phys.* 135, 194101 (2024)

<https://doi.org/10.1063/5.0197226>



## Articles You May Be Interested In

Switching spectroscopy piezoresponse force microscopy of polycrystalline capacitor structures

*Appl. Phys. Lett.* (January 2009)

Ferroelectric and magnetic properties of multiferroic BiFeO<sub>3</sub>-La<sub>0.7</sub>Sr<sub>0.3</sub>MnO<sub>3</sub> heterostructures integrated with Si (100)

*J. Appl. Phys.* (March 2015)

Interface and thickness dependent domain switching and stability in Mg doped lithium niobate

*J. Appl. Phys.* (December 2015)



Journal of Applied Physics

Special Topics Open  
for Submissions

[Learn More](#)

# Enhancing ferroelectric characterization at nanoscale: A comprehensive approach for data processing in spectroscopic piezoresponse force microscopy

Cite as: J. Appl. Phys. **135**, 194101 (2024); doi: [10.1063/5.0197226](https://doi.org/10.1063/5.0197226)

Submitted: 12 January 2024 · Accepted: 3 May 2024 ·

Published Online: 17 May 2024



View Online



Export Citation



CrossMark

H. Valloire,<sup>a)</sup> P. Quéméré,<sup>a)</sup> N. Vaxelaire,<sup>a)</sup> H. Kuentz,<sup>a)</sup> G. Le Rhun,<sup>a)</sup> and Ł. Borowik<sup>a)</sup>

## AFFILIATIONS

University Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France

<sup>a)</sup>Authors to whom correspondence should be addressed: [hugo.valloire@cea.fr](mailto:hugo.valloire@cea.fr) and [lukasz.borowik@cea.fr](mailto:lukasz.borowik@cea.fr)

## ABSTRACT

Switching Spectroscopy Piezoresponse Force Microscopy (SSPFM) stands out as a powerful method for probing ferroelectric properties within materials subjected to incremental polarization induced by an external electric field. However, the dense data processing linked to this technique is a critical factor influencing the quality of obtained results. Furthermore, meticulous exploration of various artifacts, such as electrostatics, which may considerably influence the signal, is a key factor in obtaining quantitative results. In this paper, we present a global methodology for SSPFM data processing, accessible in open-source with a user-friendly Python application called PySSPFM. A ferroelectric thin film sample of potassium sodium niobate has been probed to illustrate the different aspects of our methodology. Our approach enables the reconstruction of hysteresis nano-loops by determining the PR as a function of applied electric field. These hysteresis loops are then fitted to extract characteristic parameters that serve as measures of the ferroelectric properties of the sample. Various artifact decorrelation methods are employed to enhance measurement accuracy, and additional material properties can be assessed. Performing this procedure on a grid of points across the surface of the sample enables the creation of spatial maps. Furthermore, different techniques have been proposed to facilitate post-treatment analysis, incorporating algorithms for machine learning (K-means), phase separation, and mapping cross correlation, among others. Additionally, PySSPFM enables a more in-depth investigation of the material by studying the nanomechanical properties during poling, through the measurement of the resonance properties of the cantilever–tip–sample surface system.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0197226>

## I. INTRODUCTION

The Switching Spectroscopy Piezoresponse Force Microscopy (SSPFM) technique holds great promise for ferroelectric materials characterization, but it is still relatively new.<sup>1</sup> Only a small number of atomic force microscopy (AFM) manufacturers offer this measurement mode commercially. As a result, many SSPFM measurement setups are homemade<sup>2</sup> and require significant development efforts for data processing. The processing of SSPFM measurements is inherently dense and complex. Based on our knowledge, no open-source tool for SSPFM data processing currently exists, creating a genuine need for the analysis of this type of measurement. Various requirements within the scientific community must be addressed and resolved:

- With the existence of numerous measurement setups, a requisite solution is an adaptable tool that can accommodate all variations and, if needed, can be easily modified. This tool should establish a common foundation to ensure the accurate execution of diverse measurements.
- Given the diverse profiles of SSPFM experimenters with varying programming skills, a necessary solution is a tool that is user-friendly and comprehensible for all, including those without any programming skills.
- Quantifying measurements in Piezoresponse Force Microscopy (PFM) related techniques, presents a significant challenge.<sup>3–5</sup> Therefore, a data processing tool is essential, facilitating in-depth analysis and effectively addressing the challenge of quantification.

06 November 2024 14:58:49

- Various measurement techniques based on the SSPFM mode currently exist, and new ones are likely to be developed. Therefore, a crucial requirement is a tool that supports various measurement methods and can be modified to facilitate the accelerated development of new techniques based on the SSPFM mode.

In response to the diverse requirements within the scientific community, we have developed a comprehensive methodology that addresses the entire spectrum of SSPFM data processing. This methodology is made accessible through an open-source software called PySSPFM.

## II. DATA PROCESSING

### A. Overview

Figure 1 shows an overview of SSPFM measurement and data processing. Typically, an SSPFM measurement is conducted on a grid of points across the surface of the sample. For each measurement point, the AFM tip approaches the sample, maintains contact, while the SSPFM measurement is performed, and then retracts before moving to the next measurement point. An SSPFM measurement file is associated with each measurement point. PySSPFM performs data processing for the measurement file, resulting in an output list of values that correspond to specific ferroelectric material properties. Various artifact decorrelation protocols improve measurement accuracy and allow the measurement of

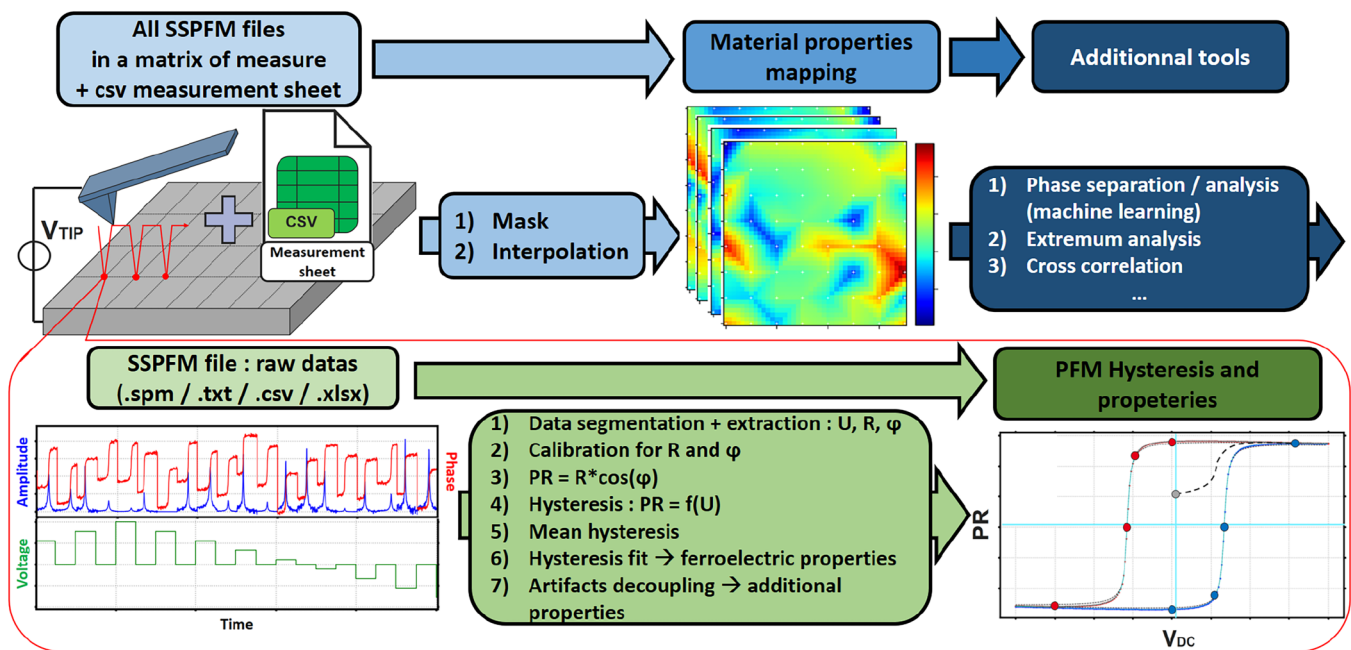
additional material properties. After analyzing all the measurement files, spatial maps are generated for each of the material properties. These cartographies provide a visual representation of the spatial distribution and variations in the material properties across the sample's surface, allowing for insights into the material's local behavior and heterogeneity.<sup>6–9</sup> The resolution of the image can be enhanced by applying 2D interpolation. A mask can also be applied to subtract the problematic pixel influence and not affect the contrast of the maps.

For each measurement file, to extract all material properties associated with a single pixel, the following set of steps is performed.

### B. Segmentation and extraction

The entire set of measurements is segmented to associate each polarization voltage value with a corresponding PFM amplitude and phase measurement. It is possible to separate the measurements into On-Field (amplitude and phase are measured simultaneously with the application of the electric field) and Off-Field (amplitude and phase are measured after the application of the electric field). Two separate procedures have been developed.

For a measurement conducted at a constant frequency, whether in resonance or not, the amplitude and phase PFM signals are usually relatively stable during the segment. Subsequently, the mean values, along with the standard deviations associated with these two parameters, are extracted over the entire segment.



06 November 2024 14:58:49

**FIG. 1.** Schematic representation of the PySSPFM code workflow. A series of voltage pulses are applied to the sample through the tip, and the amplitude–phase PFM signal is measured. The measurements are segmented, leading to the acquisition of On-Field and Off-Field hysteresis curves. Subsequently, the obtained curves undergo fitting, facilitating the extraction of ferroelectric material properties. Then, artifact effects are decorrelated from the overall signal. This measurement process can be repeated over a grid of points, enabling the reconstruction of material property maps.

For the frequency sweep procedure in resonance, a peak of PFM amplitude is obtained.  $R(\omega)$  and  $\varphi(\omega)$  describe, respectively, the frequency  $\omega$  dependence of the PFM amplitude  $R$  and phase  $\varphi$  signal. A Simple Harmonic Oscillator (SHO) model in the forced oscillation regime is commonly used to describe the cantilever motion in contact with the electromechanically stimulated sample surface.<sup>10</sup>

$$R(\omega) = A_R \frac{\omega_0^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + \left(\frac{\omega_0 \omega}{Q}\right)^2}} + R_0, \quad (1)$$

$$\varphi(\omega) = A_\varphi \arctan\left(\frac{\omega_0 \omega}{Q(\omega_0^2 - \omega^2)}\right) + \varphi_0, \quad (2)$$

where  $\omega_0$  is the contact resonant frequency in Hz, approximately equal to 300 kHz, and  $Q$  is the quality factor, approximately equal to 100 (for a tip stiffness of 3 N/m). The amplitude of each model is, respectively, denoted as  $A_R$  and  $A_\varphi$  for the amplitude and phase. The amplitude of the phase model,  $A_\varphi$ , is approximately equal to 0.5, corresponding to a variation of 180° of the signal during the frequency sweep. We add to the model the amplitude background  $R_0$  and phase offset  $\varphi_0$ , respectively. PFM amplitude and phase are fitted with the SHO model, using an algorithm based on the lmfit module.<sup>11</sup> The measurements related to the peak's amplitude, its quality factor, the resonance frequency, background signal, and the associated phase value are extracted. In case of fitting errors, the same parameters are measured based on the determination of the

maximum PFM amplitude on the segment, utilizing peak tool processing.

For both procedures, the analysis of the segment can be performed on a user-defined specific portion, allowing, for instance, the mitigation of potential dynamic effects at the segment's onset. To enhance the robustness of the processing and mitigate measurement noise, the PFM signal can be filtered using a Butterworth or mean filter. Subsequently, the measurements are presented in the form of nano-loops, where each parameter is measured in relation to the polarization voltage.

### C. Postphase calibration

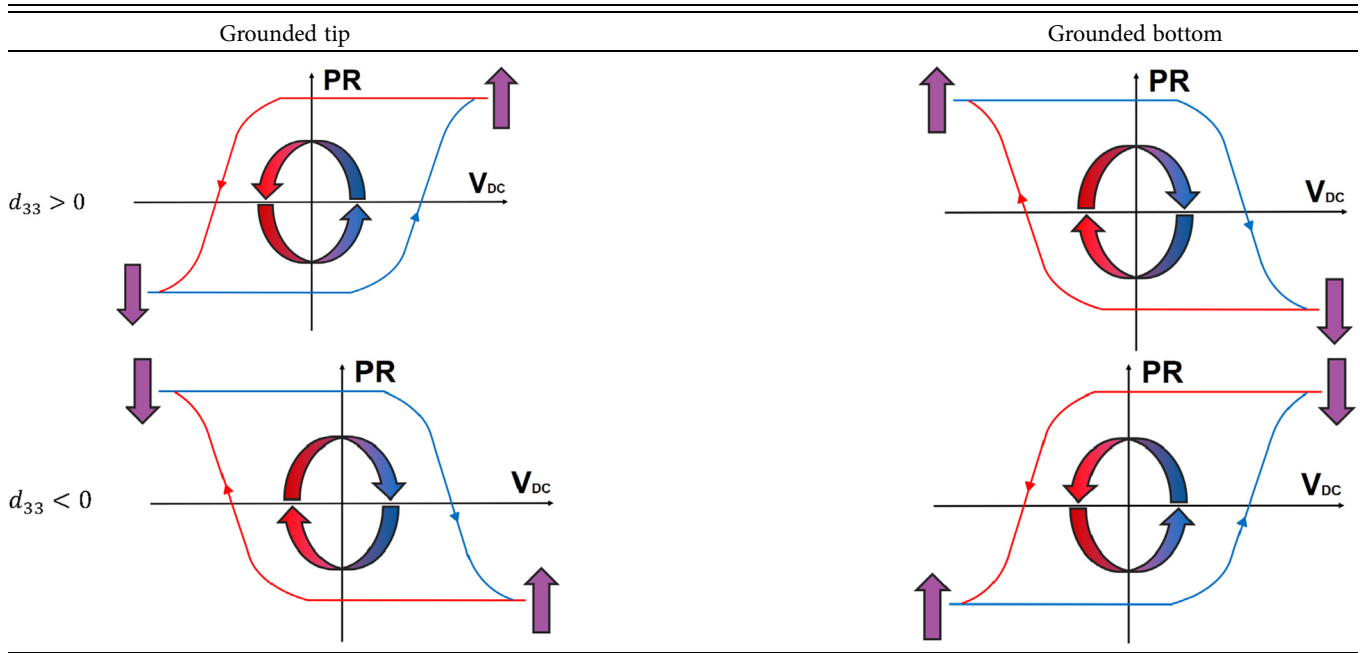
Phase is a key parameter of interest in PFM-based techniques. The pioneering study by Neumayer *et al.*,<sup>3</sup> initially proposed for PFM measurements, served in this study as a valuable foundation for developing the post-measurement phase calibration protocol, specifically in the case of SSPFM. Initially, the distinct hysteresis configurations in both Off-Field and On-Field scenarios based on experimental conditions can be identified.

The diverse configurations of hysteresis in the Off-Field mode are presented in Table I, contingent upon the specified experimental parameters (for more information, see Appendix A).

The different hysteresis configurations in the On-Field mode, characterized by a predominant electrostatic effect, are illustrated in Table II, depending on the specified experimental parameters (for more information, see Appendix B).

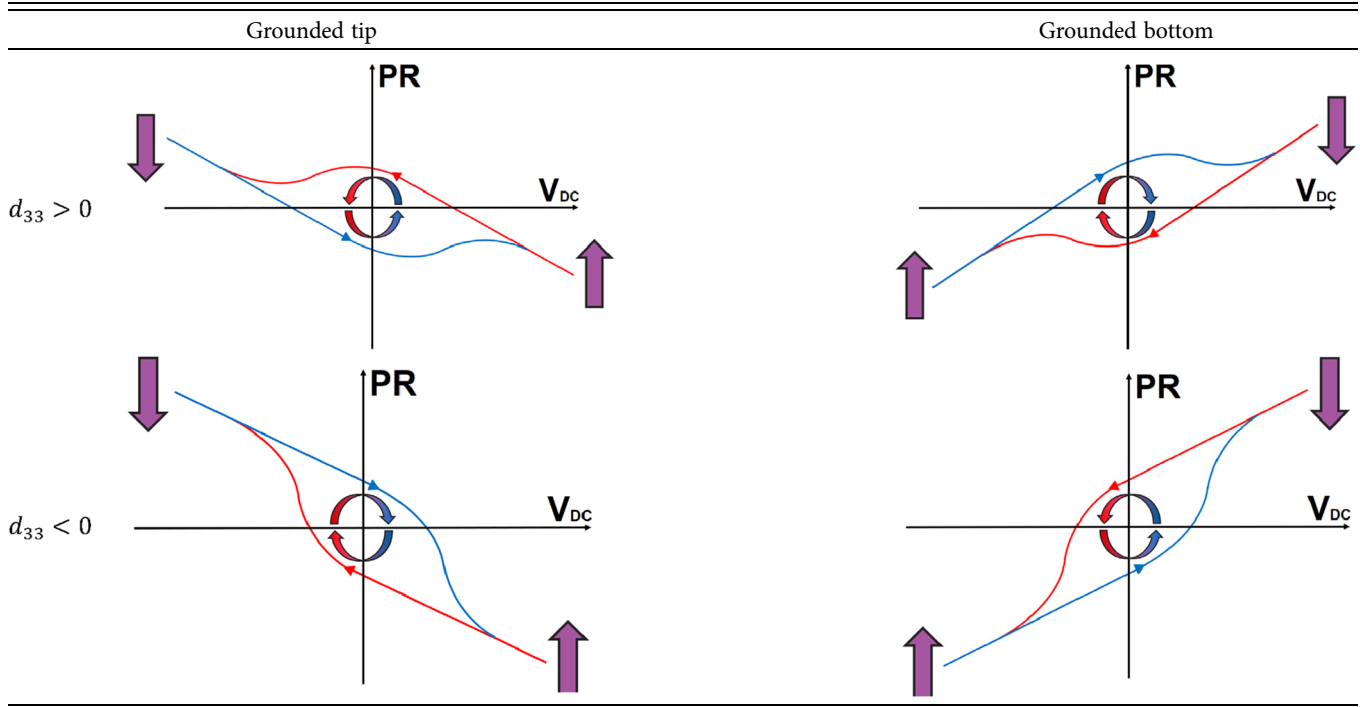
It is crucial for the user to determine the dominant component in the On-Field mode prior to phase analysis. If the

TABLE I. Hysteresis configuration indicating the polarization voltage direction and the sign of the piezoelectric coefficient for the Off-Field mode.



06 November 2024 14:58:49

TABLE II. Hysteresis configuration indicating the polarization voltage direction and the sign of the piezoelectric coefficient for On-Field mode.



ferroelectric component is dominant, the phase variation will be the same as in the Off-Field mode.

The inversion of phase values corresponding to the up and down direction of the vertical component of the polarization vector is experimentally possible in resonance-enhanced PFM. This phenomenon, resulting from the dynamic behavior of the cantilever, can be caused by a combination of experimental parameters such as the cantilever stiffness and the operating frequency.<sup>3</sup> However, in PFM-based techniques, only the amplitude and phase signals are obtained. If a phase inversion phenomenon occurs, the direction of hysteresis rotation is reversed. Therefore, the experimental phase variation with polarization voltage is analyzed and compared to the theoretical phase behavior. This step helps detect any potential phase inversions, and if present, the experimental phase values are corrected accordingly. It should also be noted that depending on the direction of the AC voltage application, the phase can be reverted. Subsequently, a phase histogram is constructed, where two distinct peaks corresponding to the up and down polarization states are expected to emerge.<sup>3,12</sup> Initially, the two peaks are centered within the phase measurement range to prevent phase switching, which can often occur for one of the polarities if this processing step is not performed. This approach helps avoid degradation in the appearance of nano-loops and facilitates subsequent measurement processing. In materials exhibiting a predominant piezo-ferroelectric response, a phase difference of approximately 180° is expected.<sup>3,12</sup> This phase difference plays a critical role in PFM as it provides insights into the relative influence of measurement artifacts on the observed piezoresponse (PR).<sup>13,14</sup> The phase

value associated with each of the peaks is extracted through Gaussian fitting or simply by using the maximum value of the peak.

It is known that a phase offset is always present in PFM-based technique measurements. This offset arises from various factors, including the AFM instrument itself, the response time of the electronic circuit, and the signal frequency, among others.<sup>3</sup> The position of the peaks is defined as

$$\varphi_1 = \varphi_{\text{offset}}, \tag{3}$$

$$\varphi_2 \approx \varphi_1 \pm 180^\circ = \varphi_{\text{offset}} \pm 180^\circ. \tag{4}$$

Identifying the positions of the two peaks facilitates the calculation of the phase difference. The user has the flexibility to select phase values (forward and reverse) based on the function employed for PR calculation (sine: forward and reverse phases are 90° and -90°, respectively, and cosine: 0° and 180°). To calibrate the phase, three distinct protocols can be applied:

- $\varphi_{\text{offset}}$  is subtracted from all phase values implemented to align the phase position, ensuring that the phase difference between the two peaks remains unchanged. This treatment method is designed to preserve the initial measurement as faithfully as possible.
- An affine relationship is applied to all phase values, adjusting the mean phase difference to 180°.

06 November 2024 14:58:49

- A threshold is established between the two peaks, and each phase value is assigned to a target value (phase forward or reverse values), based on its relative position to the threshold.

In the case of a phase inversion, all phase values are multiplied by  $-1$ ,

$$\varphi_{\text{calib}} = \pm(\varphi_{\text{meas}} - \varphi_{\text{offset}}). \quad (5)$$

Finally, in some cases, the hysteresis may exhibit a significant vertical offset [due to electrostatics<sup>15</sup> or frozen polarization effects],<sup>16</sup> and the phase signal may be unipolar. In this scenario, only one peak is detected on the phase histogram. The entire post-calibration phase procedure detailed here has also been adapted for this specific case. The user must, therefore, specify in advance whether the measurement is in phase inversion or not, based, for example, on automatic detection that can be performed on a bipolar measurement (see the [supplementary material](#) for phase calibration for bipolar and unipolar SSPFM measurements).

#### D. Hysteresis fitting and extraction of material properties

The PR is calculated based on the PFM amplitude ( $R$ ) and phase ( $\varphi$ ) values:  $PR = R * \text{func}_{\text{pha}}(\varphi)$  (with  $\text{func}_{\text{pha}}$  being either sine or cosine, depending on the selected target phase values). For both On-Field and Off-Field measurements, one or multiple hysteresis loops are determined:  $PR = f(U)$ . If multiple hysteresis loops are acquired at a measurement point, the average hysteresis loop is calculated separately for On-Field and Off-Field analyses (sometimes, the first hysteresis loop must not be considered for a pristine state sample), allowing for noise reduction in the measurements. The evolution of the hysteresis cycles throughout the measurement can be visualized to study the variation of properties during the

cycles, such as those arising from phenomena like the wake-up effect.<sup>17</sup>

The fitting process used in the study involves a model to determine the variation of  $PR$  with polarization voltage  $V$ , composed of sigmoidal (6) or arctangent (7) plus affine functions, where

$$PR^i(V) = G \times \left( \frac{1}{1 + \exp(-c^i(V - V_0^i))} - 0, 5 \right) + a \times V + b, \quad (6)$$

$$PR^i(V) = G \times \arctan(c^i \times (V - V_0^i)) + a \times V + b, \quad (7)$$

with  $i$  being the index of the branch ( $L$ : left,  $R$ : right). The complete set of fitting parameters is illustrated in Fig. 2(a). The sigmoidal (or arctangent) component represents the ferroelectric effect, while the affine component represents the contribution of artifacts (mainly of electrostatic origin). It should be noted that the fitting process is performed on the entirety of the hysteresis curve and not on each individual branch separately. Some parameters are global to the hysteresis, such as affine component parameters  $a$  and  $b$ , and hysteresis amplitude  $G$ , while some can vary for each branch, such as the position  $V_0^i$  and the dilatation coefficient  $c^i$ . Note that the user has the option to select a fit for either symmetric hysteresis (a single dilatation coefficient for the model) or asymmetric hysteresis (a distinct dilatation coefficient for each branch).

To enhance the robustness of the fitting process, the initialization of the fit parameters was implemented. This initialization step involves assigning an initial value and a range of variations for all the parameters in the fitting process. The detailed implementation of this step for each parameter will not be covered in this paper but can be found in the PySSPFM documentation available on a GitHub repository.<sup>18</sup> However, it is important to note that the

06 November 2024 14:58:49

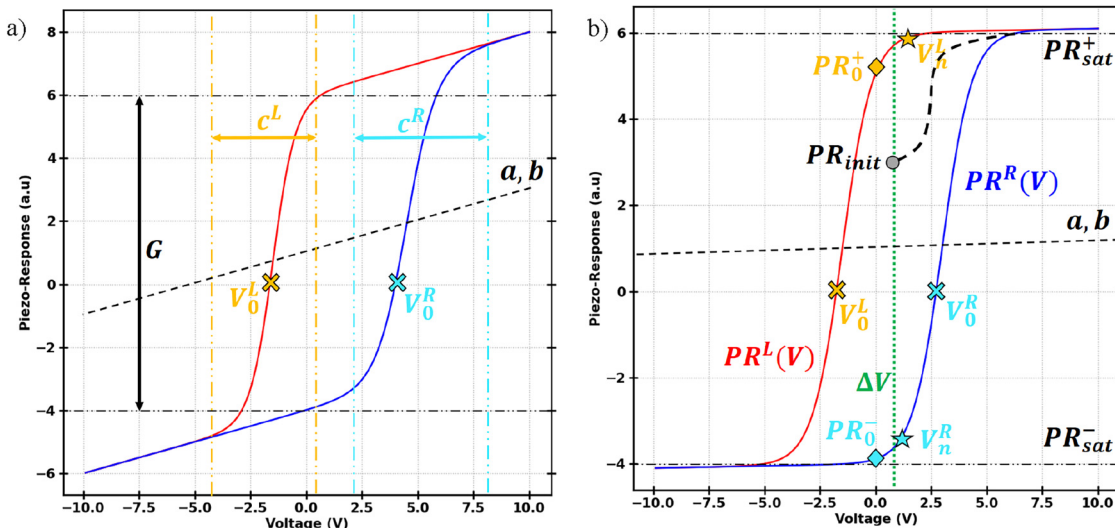


FIG. 2. Piezoresponse hysteresis nano-loop with fit parameters (a) and featuring ferroelectric characteristics (b).

direction of rotation of the hysteresis curve, determined by the sign of the hysteresis amplitude, is taken into consideration. Moreover, in the Off-Field mode, the slope of the affine relationship is fixed at 0, accounting for the absence of a linear component in this mode. Given the multitude of parameters in the hysteresis fit model and the potential deviations of measurements from the standard hysteresis form, it is recommended to use a fitting method for the fitting function of  $\text{lmfit}^{11}$  that guarantees robust convergence, even if it comes at the cost of an extended fitting duration. For instance, the Nelder–Mead method is preferred over the least squares method for this purpose.

Once the fitting step is completed, the analysis focuses on extracting a set of characteristic points from the hysteresis curves obtained during the On and Off-Field measurements, as shown in Fig. 2(b). These points form the measurement of the ferroelectric properties of the sample. In Table III, we provide a list of some of the main characteristic points.

The quality of the fit is expressed as the  $R^2$  value, which represents the mean square deviation between the fit and measurements. When analyzing ferroelectric properties, it is crucial to always consider the  $R^2$  value. This is because variations in these properties can arise from both physical nanoscale effects and poor quality fitting. For hysteresis curves with abnormally low  $R^2$  values, it is recommended to inspect each measurement individually to identify

potential issues with the data. Furthermore, additional properties are extracted from the signals of the height sensor and the tip deflection, such as their respective averages throughout the measurement, as well as the adhesion of the tip to the measured surface during retraction. Additionally, if the resonance frequency and quality factor are measured, the properties of the resonance peak are determined, encompassing the mean values of these parameters across all nano-loops. These extracted values facilitate measurements of the nano-mechanical elastic and dissipative properties of the material.<sup>26</sup>

### E. Artifact decoupling

A comprehensive set of five artifact decoupling procedures, primarily addressing electrostatic artifacts and informed by various publications,<sup>15,25,27,28</sup> has been integrated into the PySSPFM framework. Following these procedures, it becomes possible to extract the affine component (i.e., localized and non-localized electrostatic effects, electrostrictive effects, and Joule's effects: for more information, see Appendix C), as well as the nonlinear effects (Joule heating,<sup>28</sup> charge injection,<sup>28</sup> leakage current,<sup>29</sup> surface or tip degradation,<sup>30</sup> etc.)<sup>28,30,31</sup> that influence the measurement. The first procedure is designed to extract the contact potential difference (CPD) from the two minimum points of an On-Field PFM amplitude

TABLE III. Material properties extracted from hysteresis fitting.

Material property		Aspects of material physics
$PR_{\text{sat}}$	Saturation PR	Maximum PR of the polarized sample under electric field.
$PR_0$	Remnant PR	Maximum PR of the polarized sample under zero electric field.
$PR_{\text{init}}$	Pristine PR	PR of the sample in its pristine state.
$V_0$	Coercive voltage	The voltage corresponding to $PR = 0$ , representing the voltage required to perform a polarization switching.
$V_n$	Nucleation voltage	The voltage corresponding to 10% of the total switch, indicating the voltage needed to initiate domain nucleation. By comparing the values of the two branches, the relative influence of charged defects on domain wall mobility can be deduced. <sup>9</sup>
$\Delta V$	Imprint	Voltage offset corresponding to the internal field in the film, which can arise from <ul style="list-style-type: none"> <li>- Material stoichiometry<sup>19</sup></li> <li>- Internal stresses due to clamping<sup>20,21</sup></li> <li>- Charged defects<sup>20,21</sup></li> </ul>
$c$	Coef	The expansion coefficient of the branch corresponding to the switching speed and the velocity of domain walls. By comparing the values of the two branches, the relative influence of structural and/or charged defects on domain walls' mobility can be deduced. <sup>7-9</sup>
$a$	Slope	It originates from linear effects (electrostatics) and is influenced by <ul style="list-style-type: none"> <li>- Coupling between the stiffness of the cantilever and the tip-sample junction is influenced by the contact force between the tip and the sample.<sup>22</sup></li> <li>- Capacitive gradient of the tip-sample junction</li> <li>- Topography of the sample.<sup>23</sup></li> <li>- Nature of the sample.</li> </ul>
$b$	Offset PR	It can arise from <ul style="list-style-type: none"> <li>- Electrostatic effects: contact potential difference (CPD) between the tip and the sample.<sup>15</sup></li> <li>- Ferroelectric effect: frozen polarization,<sup>16</sup> size effects,<sup>16</sup> etc.<sup>24</sup></li> </ul>
$A_{\text{sw}}$	Work of switching	Hysteresis area corresponding to the total energy dissipated during the ferroelectric switching cycle. <sup>1,25</sup>



loop. The second procedure enables the extraction of coefficients for the affine component from the saturation domains of the hysteresis. The third procedure, previously outlined, involves a respective fit of On and Off-Field hysteresis loops, extracting the electrostatic component and all material properties with and without artifact influence. For detailed information on these procedures and their respective advantages, please consult the PySSPFM documentation available on the GitHub repository.<sup>18</sup>

The fourth method involves measuring multiple Off-Field hysteresis curves at different read voltages<sup>32</sup> [equivalent to the contact Kelvin Probe Force Microscopy (cKPFM) method introduced by Balke *et al.*<sup>27</sup>]. For each curve, a vertical offset is extracted by fitting the Off-Field hysteresis. Then, the vertical offset is determined as a function of read voltage, constituting the affine component. This approach is robust and precise as it is based solely on the Off-Field fits, with a broad range of validity. Moreover, it allows obtaining the entire affine component. By investigating the evolution of the PR not with respect to the writing voltage but with the reading voltage, we can obtain a more profound exploration of measurement artifacts, distinguishing between ferroelectric phenomena, electrostatic effects, charge injection, and more. However, the implementation of this method necessitates measuring several Off-Field hysteresis curves at various voltages, which is time-consuming and can disrupt the analysis of hysteresis reproducibility.

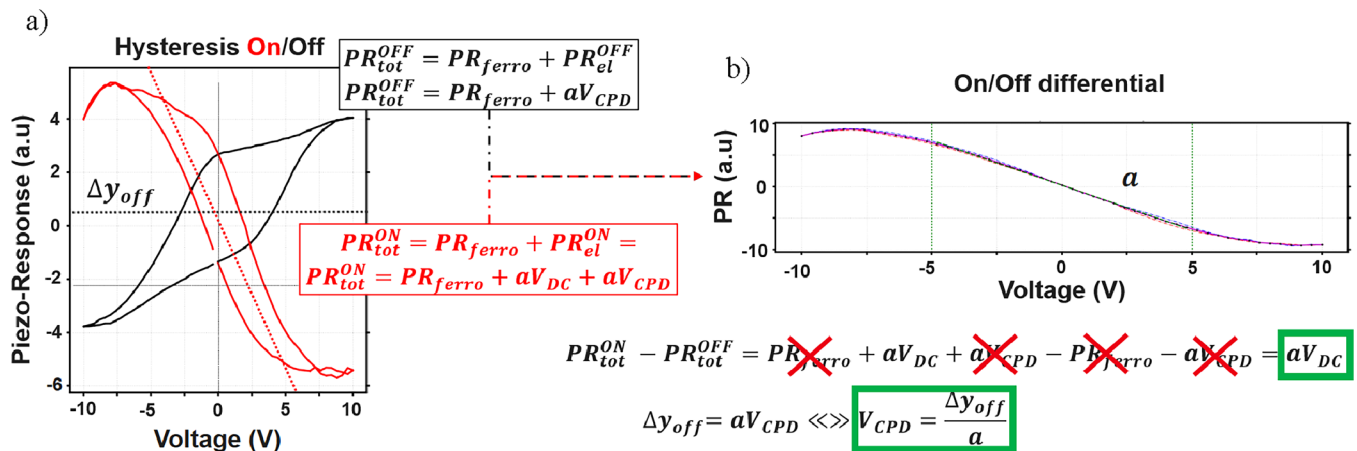
The last protocol involves subtracting On- and Off-Field hysteresis curves.<sup>15,28</sup> Figure 3 illustrates the procedure performed on a polycrystalline ferroelectric thin film sample of (K,Na)NbO<sub>3</sub> (KNN), deposited by sputtering, with a thickness of 500 nm and measured in a glovebox under a controlled argon atmosphere. Subsequently, a line intersecting zero is derived. A linear regression is performed to determine the slope, which is then used to divide the vertical offset of the Off-Field hysteresis curve to obtain the contact potential difference (CPD). Figure 4 shows a mapping of the CPD determined with this protocol, implemented on a spatial matrix of measurement points in SSPFM imaging. The mean CPD

value obtained is approximately -600 mV, in slight contrast to -900 mV that we measured using Kelvin Probe Force Microscopy (KPFM). It must be noted that CPD values derived from contact methods (e.g., SSPFM) and non-contact methods (e.g., KPFM) do not carry the same physical significance due to the gap between the tip and the sample. Consequently, they may not necessarily have the same value, especially for dielectrics as demonstrated by Balke *et al.*<sup>27</sup> Finally, this last procedure enables the determination of the entire affine component. It is both robust and precise since it primarily relies on Off-Field fits, and it is easy to implement. Additionally, the differential analysis allows for choosing the voltage domain to perform the linear regression. Indeed, in some cases, it is better to do it in the low voltage domain since non-linear effects could appear at high voltage, such as charge injection, leakage current, and Joule effect, as previously mentioned.

The various approaches for extracting artifacts can be applied independently, and their results can be compared to draw conclusions about the influence of artifacts on the measurement. It is crucial to note that each technique has its own conditions of validity, and it is necessary to verify whether these conditions are fulfilled.<sup>15</sup> Finally, it must be considered when conducting an SSPFM measurement, hysteresis may be observed in certain cases even for non-ferroelectric materials.<sup>25,33,34</sup> The signal may arise from artifacts of electrocapillary and/or electrostatic effects, which are closely associated with surface charges and are significantly influenced by ambient conditions and the presence of water layers on the surface, thereby strongly affecting surface charge dynamics.<sup>25,33,34</sup>

### III. PySSPFM TOOLBOX

After completing all the data processing and obtaining material properties for each file, additional tools integrated into the PySSPFM package can be used to gain further insights into the data analysis.



**FIG. 3.** Differential analysis between On- and Off-Field hysteresis of a KNN thin film: Hysteresis comparison between On-Field and Off-Field (PR vs  $V_{DC}$ ) (a). Differential PR vs  $V_{DC}$  (b). A linear regression analysis is performed to determine the slope of the line crossing through zero. The vertical offset of the Off-Field hysteresis is then divided by the slope to obtain the value of the Contact Potential Difference (CPD).

06 November 2024 14:58:49

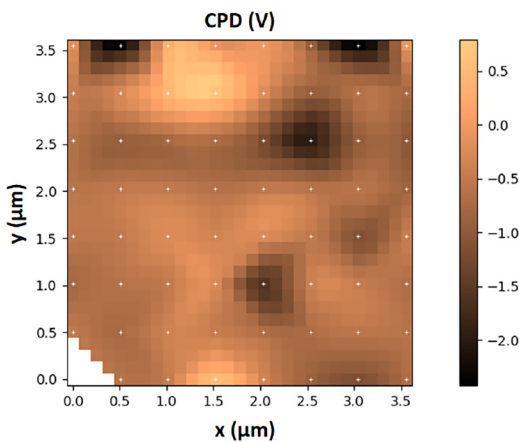


FIG. 4. Contact Potential Difference (CPD) map of a KNN thin film.

### A. Data readers

To facilitate data opening and visualization, dedicated reader modules have been developed to read datacube files (for both raw data and nano-loop visualizations), along with material properties' mapping generated at the end of the SSPFM data processing, as shown in Fig. 5. The mapping tools enable 2D interpolation to enhance map resolution. Additionally, these tools include functionalities for creating masks, allowing the mitigation of the influence of problematic pixels on map contrast or the isolation of specific regions within the map. Masks can also be generated based on conditions related to the values of a reference property to isolate specific phases of the material.

### B. Curve clustering K-means

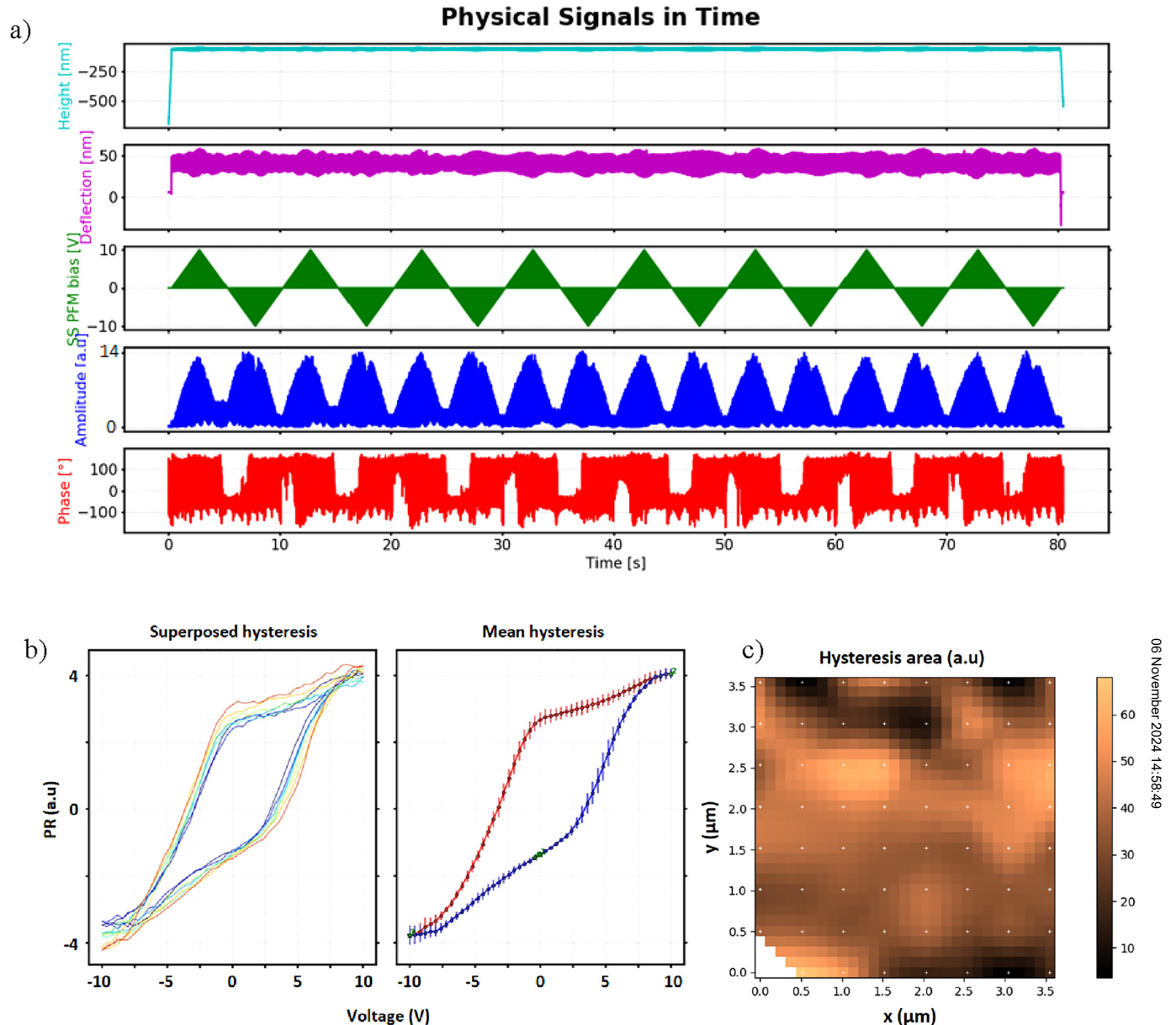
Recently, machine learning algorithms have experienced a surge in popularity, occupying an increasingly prominent role in our daily lives, exemplified by innovations such as ChatGPT. In data processing, clustering algorithms prove pertinent for segregating large volumes of multidimensional data, such as datacube files.<sup>10,19,35,36</sup> This explains the growing trend of publications utilizing these algorithms for the analysis of SSPFM mappings, even extending to the separation of ferroelectric phases<sup>36</sup> or the distinct physical contributions of the PFM signal.<sup>35</sup> The K-means algorithm is a method that separates multidimensional data points into clusters by iteratively minimizing the within-cluster sum of squares. The user specifies the number of clusters.

For each of the modes (On-Field, Off-Field, and differential), the entirety of nano-loops associated with each file is subjected to classification using the K-means function from `sklearn.cluster`.<sup>37</sup> Nano-loops can manifest in various measurements, including PR, amplitude, phase, and potential resonance frequency or quality factor for nano-mechanical analysis vs the polarization voltage. The script facilitates the selection of multiple measurements for constructing a curve: these measurements are then normalized between 0 and 1 and concatenated. This procedure was inspired by Ref. 35.

For instance, to discern the influence of various physical phenomena on the PFM signal, it proves more effective to construct a curve composed of both the amplitude and phase rather than solely relying on the PR.<sup>35</sup> The code facilitates the assignment of a reference cluster, arbitrarily defined as the cluster with the maximum number of elements. The indices of other clusters are determined based on the distance from their centroid to that of the reference cluster. In other words, the clustering indexing offers the user insights into the separation of each cluster relative to the reference cluster.

For each mode, three figures are generated, as depicted for the Off-Field hysteresis [Figs. 6(a)–6(c)] and the differential components [Figs. 6(d)–6(f)] in a point measurement network on the surface of the KNN sample. The first figure comprises the entire set of curves [Fig. 6(a) and 6(d)], the second presents the average curves for each cluster [Figs. 6(b) and 6(e)], and the third depicts a spatial map of the assigned clusters [Figs. 6(c) and 6(f)]. In each case, a color is assigned based on the index of the corresponding cluster. For Off-Field hysteresis, different regions are highlighted and spatially represented in Fig. 6(c), each with specific ferroelectric and electrostatic properties. Specifically, we observe that clusters A, B, C, and D exhibit more distinct differences in their hysteresis shapes, indicative of varying ferroelectric properties. Clusters A and B correspond to local regions with a less pronounced maximum electromechanical response compared to clusters C or D. As for clusters E and F, there is a greater variation in the hysteresis offset suggesting an intensified electrostatic effect likely caused by a significant alteration in the CPD at these four measurement points. The correlation between the CPD mapping in Fig. 4 and cluster C in Fig. 6(c) is noteworthy. The CPD is also determined using the slope of the differential component, as outlined in the latest artifact decorrelation protocol detailed in Sec. II E. This parameter is sensitive to both the stiffness of the lever and the contact junction between the tip and the sample, as well as the capacitive gradient between the tip-sample system. Both can be significantly influenced by the surface of the sample adjacent to the tip,<sup>23</sup> as it affects both the contact surface between the tip and the sample, and consequently, the tip-sample junction stiffness<sup>38</sup> and the distribution of local electrostatic field lines between the tip and the sample.<sup>23</sup> The set of differential curves in Fig. 6(d) demonstrates less variability compared to the Off-Field hysteresis in Fig. 6(a), indicating a more uniform distribution of electrostatic properties on the sample surface compared to ferroelectric properties. In one case, the electrostatic force acts at varying distances between the tip and the lever, while in the other, tip displacement is induced by the sample in contact via the inverse piezoelectric effect. As such, it appears that two clusters effectively separate one cluster composed of almost all curves, while the remaining cluster consists of three curves with lower slopes. By considering both clustering methods, Off-Field and electrostatic, a strong correlation with the CPD is observed. While the explanation of local CPD variations is beyond the scope of this paper, one could hypothesize that this change is attributed to fluctuations in localized charge densities within these regions or, alternatively, to surface impurities present on the sample. These results also illustrate one of the advantages of CPD measurement in contact SSPFM over the non-contact KPFM mode, allowing for

06 November 2024, 14:58:49



06 November 2024 14:58:49

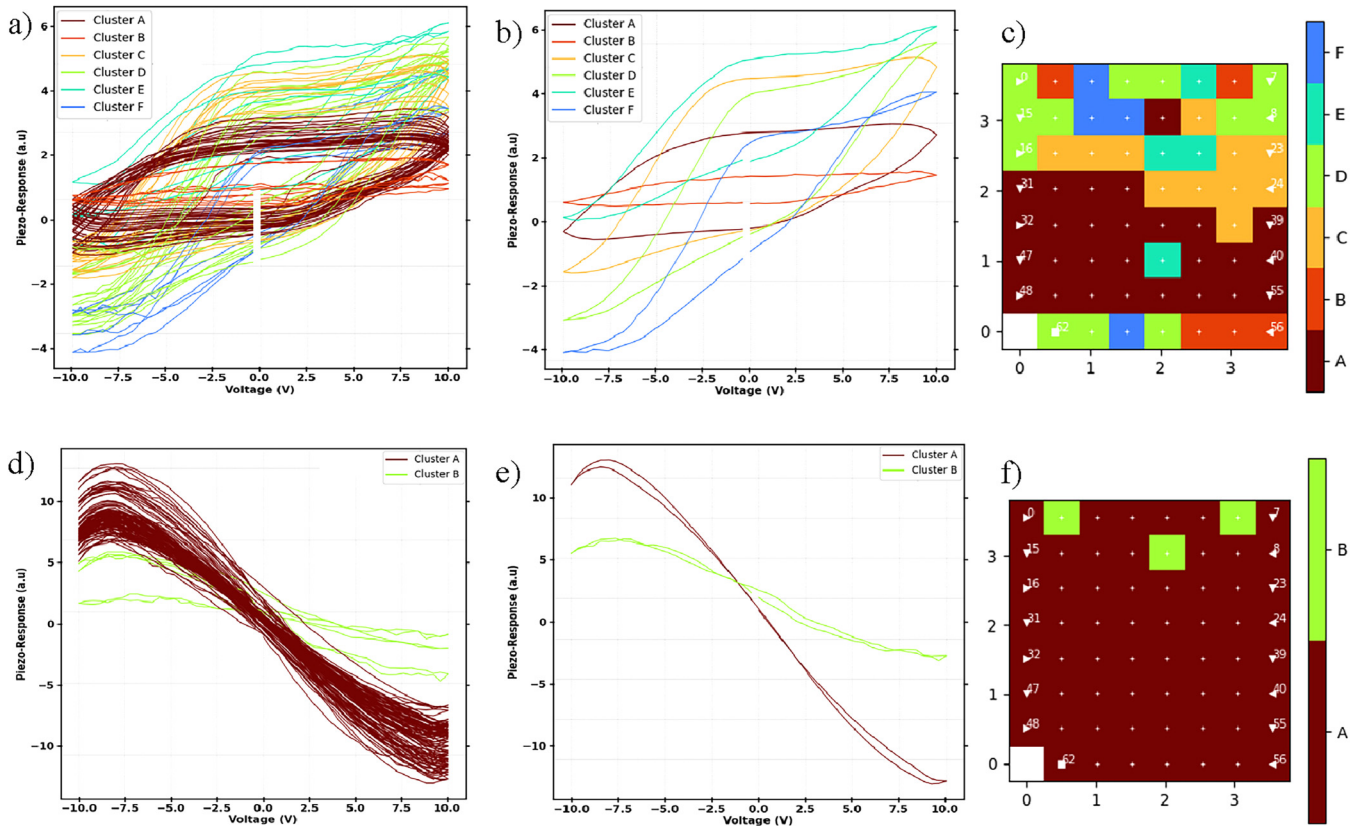
FIG. 5. Representation of Reader Outputs for measurements on a KNN thin film: SSPFM raw data (a), superimposed and mean hysteresis nano-loops (b), hysteresis area map (c).

better lateral resolution by confining the electrostatic field lines more effectively between the sample and the tip lever.<sup>27</sup>

### C. Cross correlation between maps

Correlating different maps of the sample's properties allows us to gain a better understanding of the origin of the signals and the interactions between various physical phenomena. Analyzing the

maps of different properties enables the identification of correlations between these maps and the underlying microstructure's influence.<sup>19</sup> In this work, we utilize the Pearson product-moment correlation coefficients, which express the correlation between two maps as a scalar value. The coefficient can vary between 1 (red), indicating perfect correlation, and  $-1$  (blue), indicating perfect anti-correlation. For coefficients close to 0 (white), the correlation between the maps is low. The relationship between the correlation



06 November 2024 14:58:49

**FIG. 6.** Hysteresis separation through machine learning (clustering with K-means methodology) performed on a KNN thin film for Off-Field hysteresis loop (a)–(c) and differential curves (d)–(f): set of analyzed curves (a) and (d), average curve per cluster (b) and (e), and spatial distribution of curve according to their assigned cluster (c) and (f).

coefficient matrix,  $R_{ij}$ , and the covariance matrix,  $C$ , for two variables  $i$  and  $j$ , is

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \quad (8)$$

The Pearson product–moment correlation coefficient is calculated using the “corrcoef” function from NumPy.<sup>39</sup> Subsequently, a visual table of cross correlation between different maps corresponding to specific properties of the KNN sample can be obtained, as presented in Fig. 7.

For instance, it is observable that the contrasts in the maps of positive and negative coercive voltages are strongly correlated ( $R_{23} = -0.89$ ). In fact, both properties are governed by the activation energy of domain switching. Furthermore, considering the maximum piezoresponse at saturation, a key ferroelectric property, the contrast of the latter shows no correlation with the vertical offset, influenced by electrostatic properties ( $R_{15} = 0.08$ ) or the imprint, dependent on the internal field within the thin film ( $R_{16} = -0.05$ ).

The correlation can be determined among parameters within the same mode (On-Field, Off-Field, differential) or between identical parameters between On-Field and Off-Field modes to assess the influence of linear artifacts on the measurements.

#### D. Mean hysteresis

The mean of all hysteresis can be determined over a list of pixels selected under certain conditions, which can allow the separation of different phases and provide a clearer understanding of the sample’s behavior.<sup>40</sup> Moreover, material properties and artifact decoupling can be performed. Suppose an SSPFM data set for a sample with two distinct phases of ferroelectric properties. A threshold value for the hysteresis amplitude can be set to separate the two phases based on this criterion. Then, the mean hysteresis loop can be calculated for each phase to compare their respective responses. Additionally, the distribution of each phase on the sample surface can be visualized. The pixel list can also be manually selected, for instance, to isolate properties specific to a particular area of the mapping.

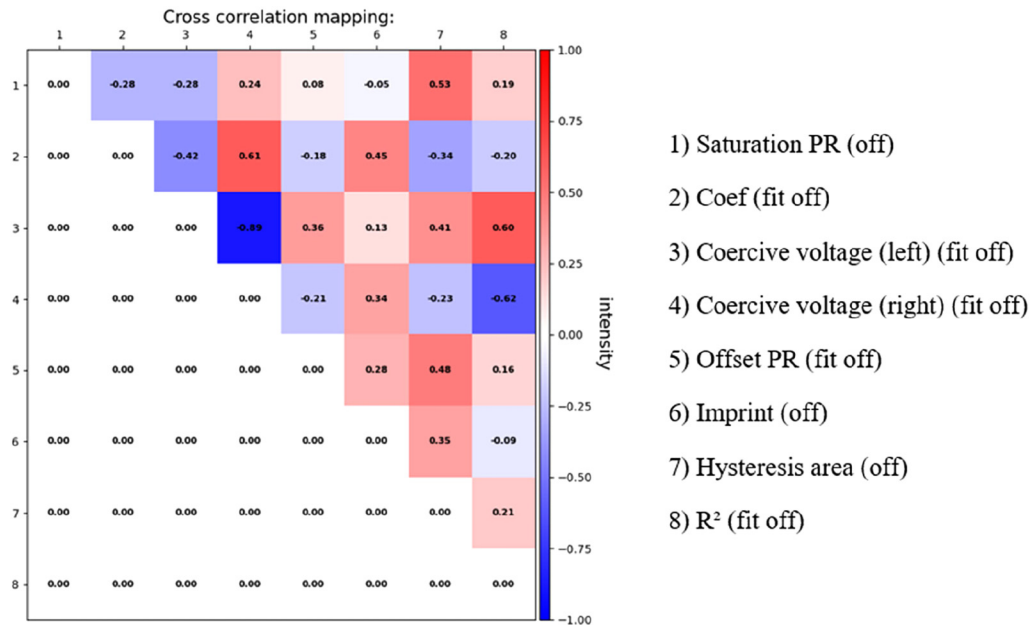


FIG. 7. Cross correlation analysis among several SSPFM maps of properties of the KNN sample.

### E. Pixel extremum

A distinct tool facilitates the identification of extreme pixel values for a given material property. The corresponding hysteresis loops are graphically represented and contrasted with the average hysteresis loop derived from all pixels. The pixel's location is displayed on the property map, and a histogram of property values, including the value of the specific pixel, is determined.

## IV. APPLICATION USAGE

We introduce PySSPFM as a user-friendly approach, featuring several key attributes. First, we have developed an inclusive and intuitive graphical user interface (GUI) to guide users through processing parameters and executable scripts. Second, adhering to the coding standard of PEP-8, PySSPFM ensures high code quality and readability. Each module has been evaluated using PyLint, a Python code analysis tool, with the majority receiving scores above 9.5 out of 10, demonstrating adherence to PEP-8 guidelines. The architecture of PySSPFM employs a modular structure with dedicated folders for specific tasks, promoting a systematic and streamlined workflow and facilitating code navigation and reusability. Detailed function headers provide comprehensive descriptions of inputs, outputs, type hints, and the function's purpose. Moreover, this paper and the code documentation available on the GitHub repository<sup>18</sup> aim to provide a comprehensive description of PySSPFM, covering its functionality, workflow, and practical utilization. The user is guided in their measurement process through the provided measurement sheet templates in the GitHub repository.<sup>18</sup> Moreover, PySSPFM incorporates quantitative analysis techniques based on established publications, offering diverse solutions

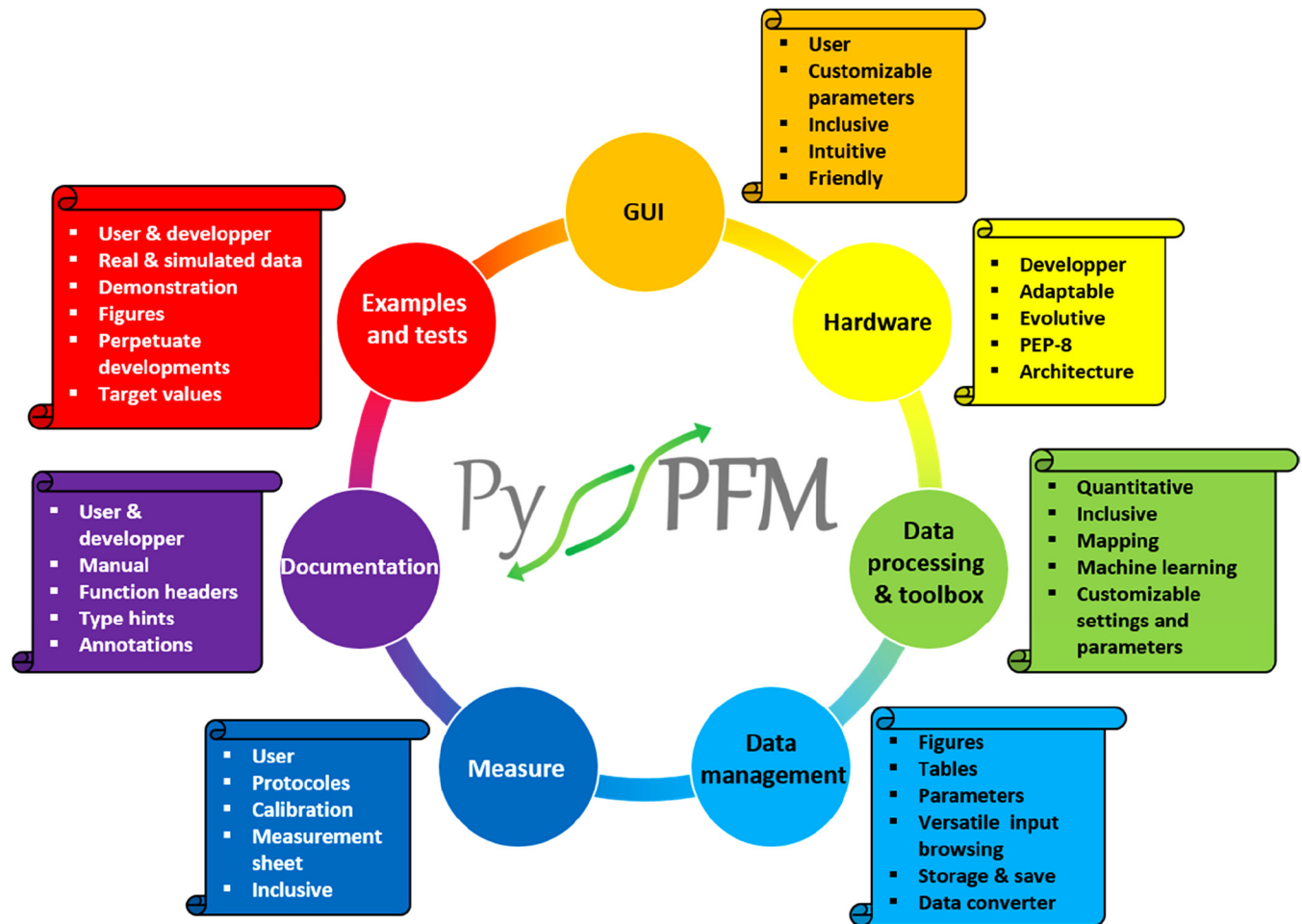
for calibration and artifact decoupling. PySSPFM includes modules for generating simulated SSPFM data, relying on physical equations derived from physical models for both ferroelectric and artifact components, allowing for realistic simulations. Users can add noise to the measurements with adjustable levels to increase fidelity. Simulated data serve to understand Python modules and their functionality, while also providing examples and insights into the measurement process and its influencing factors. A mind map showing the different aspects of PySSPFM is shown in Fig. 8.

### A. Executable modules

In general, two types of modules can be distinguished: executable modules and libraries. All executable modules rely on libraries. In the case of executable modules, processing parameters must be defined by the user. Various procedures enable users to adjust parameters, through the corresponding graphical user interface (GUI) script, or users can adjust parameters directly either in a pre-filled JSON file, in a TOML file, or in the executable Python code, within the "parameters" function located at the end of the module. Additionally, a description of the parameters is provided either in a tooltip of the GUI when hovering the mouse over the button or in the header of the parameters function. The "data\_processing," "toolbox," and "gui" directories each contain the complete set of corresponding executable codes.

### B. Input files

To employ PySSPFM, a straightforward prerequisite involves having a directory encompassing all measurement files. These files can exist as either spreadsheets, denoted by columns of value lists



06 November 2024 14:58:49

FIG. 8. PySSPFM mind map, visually depicting various aspects of the application along with their key features.

for each measurement, with extensions TXT, CSV, or XLSX or directly in the Bruker datacube SSPFM file format with the SPM extension. It is crucial to emphasize that, for Bruker SPM file data extraction, the installation of Dynamic Link Library (DLL) files alongside the Nanoscope Analysis software (Bruker) is mandatory. It is also possible to process measurement files from other manufacturers by converting them into spreadsheets with one of the valid extensions prior to processing. To our knowledge, there are no open-source libraries available that facilitate the extraction of SSPFM measurements from datacube files across various manufacturers offering this measurement mode. In our experience, PySPM<sup>41</sup> primarily supports the extraction of files originating from AFM scans and imaging modes like PFM, but does not encompass SSPFM measurements. Subsequently, there is pycroscopy,<sup>42</sup> a library developed by Oak Ridge Laboratory, which is capable of extracting data, including SSPFM datacube files. However, to our knowledge, it is exclusively compatible with devices manufactured by Oxford Instruments,

such as the Cypher AFM. The customization settings include header dimensions, the column separator in spreadsheet files for extraction, and the parameters to be extracted. The current version handles the direct measurement of height sensor, deflection, polarization voltage, PFM amplitude, and phase, signal frequency, but the code is adaptable to accommodate additional measured parameters. A converter from SSPFM datacube files from Bruker (SPM extension) to spreadsheets (TXT, CSV, or XLSX extensions) is available. In conjunction with the SSPFM data files, users must choose and complete a measurement sheet, available on the GitHub repository,<sup>18</sup> utilizing the supplied CSV template. The specified measurement parameters will be used during SSPFM data processing. A dedicated tool designed for SSPFM measurement sheet generation is also available. PySSPFM is capable to handle SSPFM-DFRT<sup>43,44</sup> mode measurements in addition to constant or sweep frequency measurements, broadening the scope of data analysis. This aspect is developed further in our dedicated paper.<sup>44</sup>

### C. Output file management

The first step of the data processing procedure in PySSPFM involves the segmentation of raw data to create nano-loop curves. This is accomplished by executing the “data\_processing/atacube\_to\_nanoloop\_s1.py” script. Upon launching this module, the user selects an initial measurement file, which undergoes processing to generate a comprehensive set of figures. These figures encompass various aspects of the data, including visualizations of the raw data in the form of graphs and maps, extraction plots for the first and last segments of the file, histograms, phase graphs, and Off-On-Field nano-loop cycles. The generated figures can be automatically saved in a new processing directory, facilitating easy access and organization. Subsequently, the entire set of files is analyzed, and for each file, measurements such as amplitude, phase, resonance frequency, and quality factor are recorded as nano-loop curves with respect to the applied bias voltage. Separate text files are created for Off- and On-Field measurements. A text file containing the complete set of measurement and data processing parameters is generated within the processing directory.

The second and final steps of the data processing involve performing phase calibration, creating and fitting the hysteresis curves, artifact decoupling, and extracting material properties. This step is executed with the “data\_processing/nanoloop\_to\_hyst.py” script. After launching the module, the user selects the folder containing the nano-loop measurements stored in TXT files. The first file is open, and corresponding figures of the fits for both On- and Off-Field measurements, as well as the artifact decoupling, are generated and optionally saved in the measurement processing folder. Subsequently, all the files are analyzed, the best nano-loop curve is determined for each file, and the material properties are extracted, separately for On- and Off-Field measurements. Last, the processing parameters for this second step are also saved in the text file along with other parameters.

For the executable files in the toolbox, a designated directory is created to save all figures and user-input parameters upon execution (optional).

### D. Tests and examples

Examples and tests are readily accessible for nearly all functions within the PySSPFM application. They adhere to the same structural framework as the primary scripts. Their purpose is to provide users with a usage guide for the associated functions and offer an idea of the operations performed by the function, often manifested through graphical representations. As for the tests, they also conform to the same framework and directly rely on the corresponding examples. They rely on quantifiable results, subsequently compared against target values using the `pytest`<sup>45</sup> library to ensure the proper functionality of the entire script suite. These tests also help in identifying code sections affected during modifications or maintenance. Importantly, all examples and tests are grounded in actual SSPFM data or data simulated through purpose-built application scripts. In its current version, PySSPFM tests are executed on GitHub servers and are compatible with Python versions ranging from 3.7 to 3.10. The tests and examples can be downloaded from GitHub and run locally by the user, provided that the `pytest` library is downloaded.

### E. Settings

The measurement processing can be tailored to the user's specific application case by adjusting the entire set of settings and processing parameters. A JSON file named “settings” encompasses all the settings associated with PySSPFM comprehensively. Users can also choose the method for managing data processing parameters (JSON, TOML, Python) and adjust settings for extracting input SSPFM files (property names, header size, column separator, etc.) or other measurement processing settings, such as the fitting method associated with the `lmfit`<sup>11</sup> fitting function. Some of the settings are linked to path management for result storage and various display options for figures. Another JSON file, “default\_settings,” should not be modified by the user as it contains settings for running tests and examples specific to the input data used for their execution. The Python script named “settings” handles the entire extraction and processing related to these two JSON files.

### F. Installation and dependencies

PySSPFM can be installed directly from PyPI using the command “pip install PySSPFM” in the IDE terminal or through GitHub using Poetry with “poetry add git + <https://github.com/CEA-MetroCarac/PySSPFM.git>” or using pip with “pip install git + <https://github.com/CEA-MetroCarac/PySSPFM.git>.” The functionality of PySSPFM relies on the use of modules such as `lmfit`,<sup>11</sup> `scikit-learn`,<sup>37</sup> `NumPy`,<sup>39</sup> `pandas`,<sup>46</sup> `openpyxl`, `matplotlib`,<sup>47</sup> `Pillow`,<sup>48</sup> `scipy`,<sup>49</sup> and `nanoscope`. Additional dependencies are optional, including `pytest`<sup>45</sup> for running test files or `toml` for managing user parameters when using TOML files. It is important to note that, for extracting SPM files from Bruker for SSPFM, DLL files installed with the Nanoscope Analysis software (Bruker) are required.

### V. CONCLUSION

In this paper, we provide a comprehensive and reliable state-of-the-art solution for SSPFM data processing with the PySSPFM application, available in open-source on a GitHub repository.<sup>18</sup> We used a thin film sample of KNN to demonstrate the feasibility and significance of the various processing steps. The code has been developed in Python, the most widely used programming language known for its ease of understanding and relevance in scientific data analysis and processing. Moreover, PySSPFM adheres to PEP-8 programming standards. In response to the defined requirements outlined as a statement of need,

- This publication, along with available documentation<sup>18</sup> and files, enables users to familiarize themselves with the tool and customize it for specific measurement scenarios if needed. Various file extensions are accommodated for processing to adapt to different experimental setups. Measurement sheet, provided in the GitHub repository,<sup>18</sup> guides users as a shared basis during measurement and is also used for extracting measurement parameters for the data processing.
- PySSPFM is controllable through a GUI designed to be as simple and intuitive as possible, or through the execution of corresponding Python scripts, with or without the use of JSON or TOML files for managing measurement parameters.

- Particular efforts have been dedicated to addressing the challenge of measurement quantification by providing a state-of-the-art quantitative data processing approach for SSPFM measurements. This approach involves the study of system resonance, various calibration procedures, and artifact decorrelation protocols.
- The PySSPFM application currently manages various SSPFM measurement modes (frequency sweep in resonance, single frequency in or out of resonance, SSPFM-DFRT,<sup>43,44</sup> cKPFM).<sup>27</sup> Moreover, PySSPFM is developed to accommodate additional measurement modes that users may implement within their measurement systems. Recently, the implementation of the new SSPFM-DFRT mode in PySSPFM<sup>43,44</sup> has been demonstrated. The entirety of the available documentation (on GitHub or within the code) empowers users to make these additions.

Future developments are currently underway to enhance PySSPFM, including the implementation of hysteresis fine-grain structure<sup>8</sup> to delve deeper into the analysis of the influence of point defects (either charged or not) on the ferroelectric properties of the material. Progress in machine learning has been leveraged to increase the signal-to-noise ratio in PFM-based techniques, enabling measurements across a broader range of samples or yielding a more accurate signal near coercive voltages.<sup>50,51</sup> Such procedures could be adapted in PySSPFM modules.

## SUPPLEMENTARY MATERIAL

See the supplementary material for additional data, including an application case of phase calibration for bipolar and unipolar SSPFM measurements.

## ACKNOWLEDGMENTS

We express gratitude to all the developers of the Python open-source packages on which PySSPFM relies, more particularly `lmfit`,<sup>11</sup> `scikit-learn`,<sup>37</sup> `NumPy`,<sup>39</sup> `Matplotlib`,<sup>47</sup> and `SciPy`,<sup>43</sup> among others. This work, carried out on the Platform for Nanocharacterization (PFNC), was supported by the “Recherches Technologiques de Base” program of the French National Research Agency (ANR) and the “IMINEN” research project (No. ANR-22-CE-09-0032 IMINEN).

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**H. Valloire:** Conceptualization (lead); Data curation (lead); Formal analysis (lead); Investigation (lead); Methodology (lead); Software (lead); Writing – original draft (lead); Writing – review & editing (equal). **P. Quéméré:** Software (supporting); Supervision (supporting); Writing – review & editing (equal). **N. Vaxelaire:** Project administration (lead); Resources (equal); Supervision (lead); Validation (equal); Writing – review & editing (equal). **H. Kuentz:** Resources (equal). **G. Le Rhun:** Resources (equal). **L. Borowik:** Resources (equal); Supervision (lead); Validation (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

## APPENDIX A: OFF-FIELD HYSTERESIS CONFIGURATION

The direction of vertical polarization, assuming a predominantly ferroelectric effect, depends on the applied voltage between the tip and the material’s bottom electrode. Voltages exceeding the low and high coercive voltages of the hysteresis are categorized as low (–) and high (+) voltages, respectively. Two distinct scenarios are distinguished: one for the grounded tip configuration and the other for the grounded bottom configuration. [Table IV](#) succinctly illustrates the polarization direction relative to the applied voltage for both scenarios.

**TABLE IV.** Dependence of induced polarization on applied voltage for the material in the case of grounded tip and bottom configurations.

Grounded <sup>a</sup>	Voltage <sup>b</sup>	Polarization <sup>c</sup>
Tip	–	↓
	+	↑
Bottom	–	↑
	+	↓

<sup>a</sup>Grounded, direction of external polarization voltage applied.

<sup>b</sup>Voltage, sign of external polarization voltage applied.

<sup>c</sup>Polarization, direction of induced polarization.

With the sign of the piezoelectric coefficient of the material  $d_{33}$ , we can determine the dependence between the direction of the vertical component of the polarization vector of the material and the corresponding measured PFM phase value, as shown in [Table V](#).

Based on the information extracted from the initial two tables, one can deduce a resultant [Table VI](#) illustrating the orientation of hysteresis rotation, which is entirely dependent on the experimental conditions.

**TABLE V.** Dependence of the measured PFM phase value on the polarization of the material, corresponding to the positive and negative sign of  $d_{33}$ .

Sign of $d_{33}$	Polarization <sup>a</sup>	Phase <sup>b</sup>
+	↓	Reverse
	↑	Forward
–	↓	Forward
	↑	Reverse

<sup>a</sup>Polarization, direction of polarization.

<sup>b</sup>Phase, PFM phase value measured.



**TABLE VI.** Dependence of the orientation of hysteresis rotation on experimental parameters.

	Grounded tip	Grounded bottom
$d_{33} > 0$	Counterclock	Clock
$d_{33} < 0$	Clock	Counterclock

**APPENDIX B: ON-FIELD HYSTERESIS CONFIGURATION**

In the specific context of On-Field measurements, marked by a predominant electrostatic component (where the phase value is determined by the electrostatic component), the polarity of the electrostatic component's slope is dictated by the direction of the applied voltage. Consequently, the voltage value can be directly correlated with the phase, as illustrated in Table VII, without any modification in the hysteresis rotation direction. The transition in the measured PFM phase value for the affine component occurs at the CPD.

**TABLE VII.** Relationship between the sign of the affine slope and the measured PFM phase value with applied polarization voltage, considering both grounded tip and bottom configurations.

Grounded <sup>a</sup>	Sign of affine slope	Voltage <sup>b</sup>	Phase <sup>c</sup>
Tip	–	–	Forward
		+	Reverse
Bottom	+	–	Reverse
		+	Forward

<sup>a</sup>Grounded, direction of external polarization voltage applied.<sup>b</sup>Voltage, sign of external polarization voltage applied.<sup>c</sup>Phase, PFM phase value measured.**APPENDIX C: DECOMPOSITION OF TOTAL PIEZORESPONSE TERMS IN SSPFM MEASUREMENTS**

The total piezoresponse of the first-order harmonic component  $PR_{\omega}^{\text{tot}}$  can be decomposed into two terms:  $PR_{\omega}^{\text{piezo,ferro}}$  is the first term, due to piezo- and ferroelectric effects, and is manifested as hysteresis, while the second term  $PR_{\omega}^{\text{artefacts}}$  includes the influences of the considered artifacts (i.e., localized and non-localized electrostatic effects, electrostrictive effects, and Joules effects),

$$PR_{\omega}^{\text{tot}} = PR_{\omega}^{\text{piezo,ferro}} + PR_{\omega}^{\text{artefacts}},$$

$$PR_{\omega}^{\text{tot}} = PR_{\omega}^{\text{piezo,ferro}} + PR_{\omega}^{\text{elec, cant}} + PR_{\omega}^{\text{elec, tip}} + PR_{\omega}^{\text{electrostriction}} + PR_{\omega}^{\text{joules}},$$

$$PR_{\omega}^{\text{tot}} = d_{33}^{\text{eff}}(V_{dc})V_{ac} + \frac{1}{k_{\omega}} \frac{\partial C^{\text{cant}}(z)}{\partial z} (V_{dc} - V_{\text{CPD}}^{\text{glob}})V_{ac} + \frac{1}{k_{\omega}^*} \frac{\partial C^{\text{tip}}(z)}{\partial z} (V_{dc} - V_{\text{CPD}}^{\text{loc}})V_{ac} + \frac{2M_{333}}{t} V_{ac}V_{dc} + \frac{2\beta}{R} V_{ac}V_{dc}.$$

**REFERENCES**

- S. Jesse, A. P. Baddorf, and S. V. Kalinin, *Appl. Phys. Lett.* **88**, 062908 (2006).
- F. Flores-Ruiz, J. Gervacio-Arciniega, E. Murillo-Bracamontes *et al.*, *Measurement* **108**, 143–151 (2017).
- S. M. Neumayer, S. Saremi, L. W. Martin *et al.*, *J. Appl. Phys.* **128**, 171105 (2020).
- N. Balke, S. Jesse, P. Yu *et al.*, *Nanotechnology* **27**, 425707 (2016).
- J. O'Donnell, E. U. Haq, C. Silien *et al.*, *J. Appl. Phys.* **129**, 185104 (2021).
- I. K. Bdkin, A. L. Kholkin, A. N. Morozovska *et al.*, *Appl. Phys. Lett.* **92**, 182909 (2008).
- S. V. Kalinin, B. J. Rodriguez, A. Y. Borisevich *et al.*, *Adv. Mater.* **22**, 314–322 (2010).
- S. V. Kalinin, S. Jesse, B. J. Rodriguez *et al.*, *Phys. Rev. Lett.* **100**, 155703 (2008).
- S. Jesse, B. J. Rodriguez, S. Choudhury *et al.*, *Nat. Mater.* **7**, 209–215 (2008).
- B. Huang, E. N. Esfahani, and J. Li, *Natl. Sci. Rev.* **6**, 55–63 (2019).
- M. Newville, T. Stensitzki, D. B. Allen, and A. Ingargiola, LMFIT, version 1.2.2, see <https://lmfit.github.io/lmfit-py/> for “Non-linear least-square minimization and curve-fitting for Python,” 2014, Zenodo. 10.5281/zenodo.11813.
- S. Bradler, A. Schirmeisen, and B. Roling, *J. Appl. Phys.* **122**, 065106 (2017).
- S. Jesse, A. P. Baddorf, and S. V. Kalinin, *Nanotechnology* **17**, 1615–1628 (2006).
- T. Jungk, Á Hoffmann, and E. Soergel, *J. Microsc.* **227**, 72–78 (2007).
- D. Alikin, A. Abramov, A. Turygin *et al.*, *Small Methods* **6**, 2101289 (2022).
- M. Alexe, C. Harnagea, D. Hesse *et al.*, *Appl. Phys. Lett.* **79**, 242–244 (2001).
- N. Balke, T. Schenk, I. Stolichnov *et al.*, in *Ferroelectricity in Doped Hafnium Oxide* (Woodhead Publishing, 2019), Chap. 7.1, pp. 291–316. doi: 10.1016/B978-0-08-102430-0.00014-0.
- H. Valloire and P. Quéméré, PySSPFM GitHub repository (2023) to be inserted once the repository is made public: [Link].
- Y. Hiranaga, T. Mimura, T. Shimizu *et al.*, *Jpn. J. Appl. Phys.* **61**, SN1014 (2022).
- D. Lee, A. Yoon, S. Y. Jang *et al.*, *Phys. Rev. Lett.* **107**, 057602 (2011).
- A. Gruverman, B. J. Rodriguez, A. I. Kingon *et al.*, *Appl. Phys. Lett.* **83**, 728–730 (2003).
- S. V. Kalinin and D. A. Bonnell, *Phys. Rev. B* **65**, 125408 (2012).
- N. Balke, S. Jesse, B. Carmichael *et al.*, *Nanotechnology* **28**, 065704 (2017).
- A. L. Kholkin, V. V. Shvartsman, A. Y. Emelyanov *et al.*, *Appl. Phys. Lett.* **82**, 2127–2129 (2003).
- S. Jesse, H. N. Lee, and S. V. Kalinin, *Rev. Sci. Instrum.* **77**, 073702 (2006).
- A. Gannepalli, D. G. Yablou, A. H. Tsou *et al.*, *Nanotechnology* **22**, 355705 (2011).
- N. Balke, P. Maksymovych, S. Jesse *et al.*, *ACS Nano* **8**, 10229–10236 (2014).
- N. Balke, S. Jesse, Q. Li *et al.*, *J. Appl. Phys.* **118**, 072013 (2015).
- B. Kim, D. Seol, S. Lee *et al.*, *Appl. Phys. Lett.* **109**, 102901 (2016).
- S. Gonzalez Casal, Ph.D. thesis (INSA de Lyon, 2022).
- B. Gautier and D. Albertini, *Technique de l'ingénieur: Mesures mécaniques et dimensionnelles R6719*, 19–23 (2022).
- Bruker application note, see <https://www.bruker.com/en/products-and-solutions/microscopes/materials-afm/afm-modes/ss-pfm.html> for “Characterizing ferroelectric materials with SS-PFM and DCUBE PFM.”
- H. Miao, C. Tan, X. Zhou *et al.*, *Europhys. Lett.* **108**, 27010 (2014).
- N. Balke, P. Maksymovych, S. Jesse *et al.*, *ACS Nano* **9**, 6484–6492 (2015).
- G. K. Ligonde, K. N. Williams, I. Gaponenko *et al.*, *Adv. Phys. Res.* **2**, 2200090 (2023).
- A. Kumar, O. Ovchinnikov, S. Guo *et al.*, *Phys. Rev. B* **84**, 024203 (2011).
- F. Pedregosa, G. Varoquaux, A. Gramfort *et al.*, *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
- S. Jesse, B. Mirman, and S. V. Kalinin, *Appl. Phys. Lett.* **89**, 022906 (2006).
- C. R. Harris, K. J. Millman, S. J. van der Walt *et al.*, *Nature* **585**, 357–362 (2020).
- K. P. Kelley, A. N. Morozovska, E. A. Eliseev *et al.*, *Nat. Mater.* **22**, 1144–1151 (2023).

- <sup>41</sup>O. Scholder, pySPM, version 0.2.16, 2019, Zenodo. 10.5281/zenodo.998575, see <https://pypi.org/project/pyspm/> for information on the Python library and its installation.
- <sup>42</sup>S. Somnath, C. Smith, N. Laanait, N. Laanait, R. K. Vasudevan, A. V. Ievlev, A. Belianinov, A. R. Lupini, M. (Arjun) Shankar, S. V. Kalinin, and S. Jesse, “USID and Pycroscopy – Open frameworks for storing and analyzing spectroscopic and imaging data,” arXiv: Data Analysis, Statistics and Probability (2019).
- <sup>43</sup>H. Valloire, D. Mariolle, and Ł. Borowik, F.R. patent 2,204,531 (12 May 2022).
- <sup>44</sup>H. Valloire, P. Quéméré, N. Vaxelaire *et al.*, *Rev. Sci. Instrum.* (in press) (2024).
- <sup>45</sup>H. Krekel, B. Oliveira, R. Pfannschmidt, F. Bruynooghe, B. Laughner *et al.*, pytest, version 7.4.3, 2004, see <https://github.com/pytest-dev/pytest> for information on the Python library and its installation.
- <sup>46</sup>W. McKinney *et al.*, in *Proceedings of the 9th Python in Science Conference* (Wes McKinney, 2010), Vol. 445, pp. 51–56.
- <sup>47</sup>J. D. Hunter, *Comput. Sci. Eng.* **9**, 90–95 (2007).
- <sup>48</sup>A. Clark, Pillow, version 10.1.0, 2015, Zenodo 10.5281/zenodo.596518, see <https://pypi.org/project/Pillow/> for information on the Python library and its installation.
- <sup>49</sup>P. Virtanen, R. Gommers, T. E. Oliphant *et al.*, *Nat. Methods* **17**, 261–272 (2020).
- <sup>50</sup>P. Sriboriboon, H. Qiao, O. Kwon *et al.*, *npj Comput. Mater.* **9**, 28 (2023).
- <sup>51</sup>H. S. Yuchi, K. N. Williams, G. K. Ligonde *et al.*, *AI for Accelerated Materials Design NeurIPS ‘22 Workshop* (NeurIPS, New Orleans, 2022).