Data-driven Scanning Probe Microscopy for Advanced Functional Materials

Boyuan Huang

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Jiangyu Li, Chair

I. Y. (Steve) Shen

Junlan Wang

Program Authorized to Offer Degree:

Mechanical Engineering

University of Washington

Abstract

Data-driven Scanning Probe Microscopy for Advanced Functional Materials

Boyuan Huang

Chair of the Supervisory Committee:
Professor Jiangyu Li
Department of Mechanical Engineering

Advanced functional materials have revolutionized our daily life and work in depth. Their applications are widely used in many fields, including but not limited to information technology, energy conversion and life science. However, the pace of improvement varies among different materials. For example, the trifecta of manufacturing, characterization, and theoretical understanding lays the foundation of Moore's law in the semiconductor industry, while the complex mechanisms reflected on coupled chemical, physical, and mechanical effects at the nanoscale evidently retard the progress of energy materials. Thus, a pressing as well as universal challenge facing development of advanced materials is how we can better understand their physics and various coupling at local length scales.

Scanning Probe Microscopy (SPM) is a powerful tool to study a wide variety of physical properties at the nanoscale which can be directly traced to their microstructures and further

linked to the performance on the device level. In this dissertation, we first introduce that SPM techniques have great potential to realize such promise by using halide perovskite solar cells as an example, which have emerged as one of the most promising next-generation photovoltaic materials. Yet their microscopic phenomena involving photo-carriers, ionic defects, spontaneous polarization are still inadequately understood. In this part, we highlight some recent progress and challenges of investigation toward local probing of its photocurrent, surface potential, spontaneous polarization, ionic motion, and chemical degradation via SPM. These findings resolve ambiguity regarding the crystalline nature of $CH_3NH_3PbI_3$ and its implication on photovoltaic conversion, reconcile the diverse and apparent contradictory data reported in literature, and point a direction toward engineering ferroic domains for enhanced efficiency.

We also summarize technical limitations and challenges encountered in this systematic study of $CH_3NH_3PbI_3$ and emphasize the need of innovative experimental methodologies based on SPM to acquire high quality, efficient, and physically relevant scientific data for deep analysis. To enable such vision and handle those challenges, the recent advances in big data inspire us to head for a data-driven SPM. In this part, a rough piezoelectric material is first examined using SPM combined with our newly developed sequential excitation (SE) method, which acquires multi-dimensional data over a range of frequencies excited in a sequential manner and enables us to map its intrinsic electromechanical properties at the nanoscale with high fidelity. To pursue a faster scanning speed, we then upgrade SE to the high-throughput sequential excitation which can capture full-time contact dynamics of probe-sample interaction of all pixels in just one scan. Using electrochemically active granular ceria as an example, we map both linear and quadratic electrochemical strain accurately across grain boundaries with high spatial resolution where the conventional approach fails. Both damped harmonic oscillator

(DHO) model and principal component analysis (PCA) are carried out to derive intrinsic electromechanical coupling of interests. It turns out that PCA can not only speeds up the analysis by four orders of magnitude, but also allows a physical interpretation of its modes in combination with the DHO model. This SE methodology can be easily adapted for other SPM modes to probe a wide range of microscopic phenomena.

Finally, the collected big data can not only pave the way for materials research, but also repay the development of SPM techniques. Here we demonstrate an artificial intelligence scanning probe microscopy (AI-SPM) for pattern recognition and feature identification in ferroelectric materials and electrochemical systems. This data-driven AI-SPM can respond to classification via adaptive experimentation with additional probing at critical domain walls and grain boundaries, all in real-time on the fly without human interference. Key to our success is an efficient machine learning strategy based on a support vector machine (SVM) algorithm capable of pixel-by-pixel recognition instead of relying on data from full mapping, making real-time classification and control possible during scan, with which complex electromechanical couplings at the nanoscale in different material systems can be resolved by the AI. For SPM experiments that are often tedious, elusive, and heavily rely on human insight for execution and analysis, this is a major disruption in methodology.

In conclusion, we believe such a data-driven SPM will not only facilitate the study of advanced functional materials, but also probably impact development for a wide range of scientific instruments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, my deepest gratitude goes to my advisor, Prof. Jiangyu Li, who has supported me with his guidance and expertise for the past four years. I could not accomplish my present achievements without his endless encouragement as well as trust. I would also like to express my sincere regards to my committee members, Prof. I. Y. (Steve) Shen, Prof. Junlan Wang, and Prof. Guozhong Cao for their valuable advice and time for evaluating my dissertation.

A special thank you to lab mates for all discussions and friendship during the past four years. I would also like to acknowledge my home department, Mechanical Engineering, and UW Clean Energy Institute for the financial and intellectual support.

Last but not least, I would like to thank my family who has always been the greatest support in my life and thank my friends for their companion and help.

Chapter 1. INTRODUCTION

1.1 MOTIVATION

Since Apple released its landmark product iPhone 4 in 2010, the last decade has witnessed a great boom of smartphone. No one could deny that smartphones have deeply reshaped our daily lives. Nowadays, new smartphones are annually announced with continuous progress in their speed, power, battery life, and general performance, behind which is also the rapid development of various functional materials. Nearly everywhere you look within a smartphone—whether an end component itself, the component's individual parts, or the manufacturing or processing steps that enabled that component— you can find advanced materials at some step along the way [1].

For example, a thin layer of indium tin oxide with high conductivity is particularly important to the touchscreen functionality [2]. Dielectric ceramics such as piezoceramics are not only utilized to pump out high-quality sound, but also provide antennas, inductors, and other radiofrequency (RF) components for sending and receiving signals from cell towers and connecting to wireless networks [1]. Temperature-compensated crystal oscillators with thermal compensation circuitry can improve GPS location accuracy [1,3]. Corning's Gorilla Glass is toughened by ion exchange to increase its durability [4], enabling the screen to better survive drops and abuse. Organic light emitting diode (OLED) with higher image and color quality is gradually taking the place of liquid crystal display (LCD) [5]. And there is also a predicted trend for foldable displays, which offers opportunities for flexible materials such as plastic in the cover screen market [1].

However, the pace of improvement varies among components and materials. There is a simple yet informative comparison between the evolvement of mobile processor and the battery capacity. The Apple A13 Bionic processor in the latest iPhone model is 12.5x faster than the A4 processor

in iPhone 4, while the difference between their battery capacity is only about twice [6]. Generally speaking, it is the trifecta of manufacturing, characterization, and theoretical understanding in semiconductor industry that contributes to such rapid progress of mobile processors. In contrast, the relatively slow progress in energy materials and devices is mainly due to their complex mechanisms. For example, the operation of battery often includes multiple electrochemical reactions and ionic and electron transport steps, liquid flows on multiple length scales, and coupled chemical, physical, and mechanical effects [7]. This undoubtedly brings huge difficulties to characterization, not to mention the construction of predictive theories and further guidance to manufacturing. It is not uncommon to confront such a situation in materials research, as many of advanced functional materials are expected to have outstanding performance in more than one aspect, which also means possible coupled chemical, physical, and mechanical effects waiting for investigation.

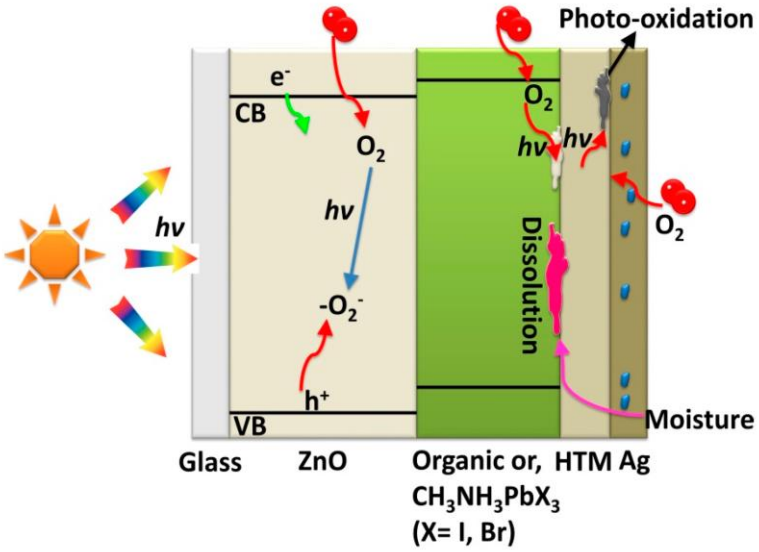

Figure 1.1. Schematic of a halide perovskite solar cell. This image is reprinted from [8,9].

Another example is photovoltaic materials like halide perovskite solar cells. Inside this multicomponent system are interacting photons and electrons, ionic defects, spontaneous

polarization, and molecular vibration and rotation on top of numerous grains, grain boundaries, and interfaces. Lack of clear fundamental understanding is hindering the optimization of their photovoltaic properties, device efficiency, and lifetime. A critical solution to remove this major roadblock will largely rely on the advancement in characterization approaches, especially characterization that can interrogate structure-property relations on the mesoscopic levels. In this sense, scanning probe microscopy to be introduced in the next section would be a perfect candidate for this purpose.

1.2 SCANNING PROBE MICROSCOPY

Since its invention in the 1980s, Scanning Probe Microscopy (SPM) has revolutionized local probing of a wide range of materials, devices, and systems with nanoscale resolution [10]. It utilizes a cantilever with a sharp tip radius unmatched by human finger and well below the optic diffraction limit, making it possible to achieve a spatial resolution as small as a few nanometers when touching and scanning the sample of interests. After decades of extensive development, SPM is no longer limited to probing morphology, its cantilever now can be modified to interrogate a range of valuable information, such as hardness, adhesion, temperature, and charging state under either local stimuli imposed by the scanning probe or global excitation applied externally [11,12], which are usually out of reach for either optic or electron microscopies. It is evident that SPM offers an unprecedented platform and has been used among many science fields, such as materials science, chemistry, physics, and biology.

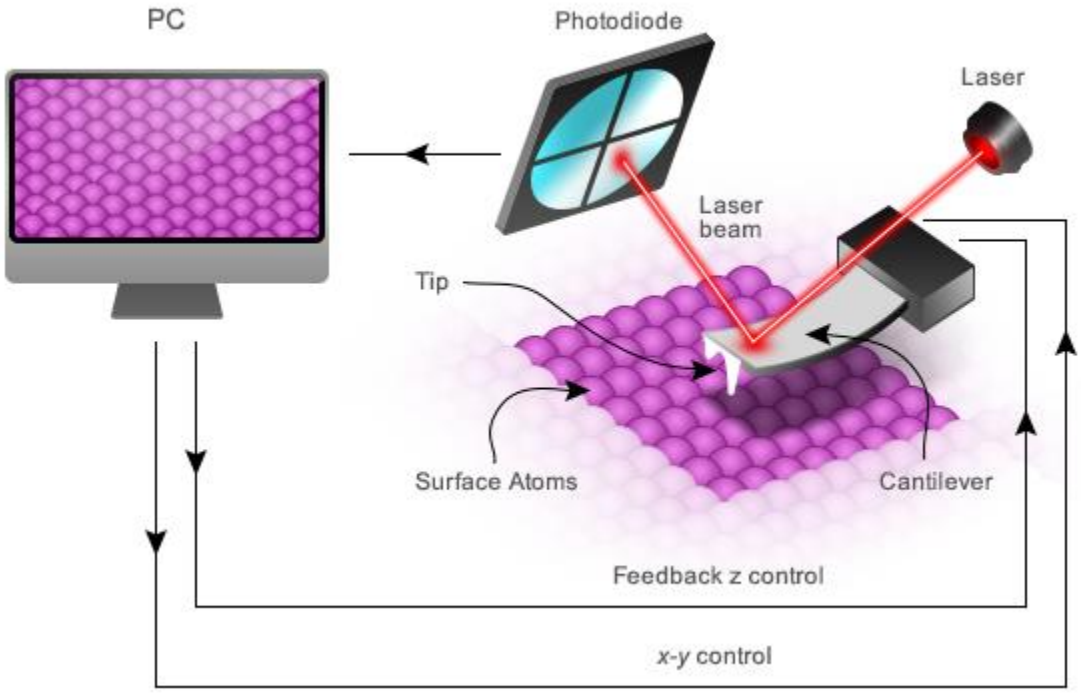

Figure 1.2. SPM uses beam deflection detection. As the cantilever is deflected via its interaction with the sample, the reflection of the laser beam will also be displaced on the photodiode. The difference between the amount of signal in the up and down halves of the photodiode is what we called deflection signal. This image is reprinted from [13].

According to tip-material interaction, there are two kinds of modes for topography. They are also basis of many other variant modes. The first one is contact mode, also known as constant force mode, in which the controller monitors the cantilever deflection when the tip scans the surface of sample. Before touching the sample, the cantilever is relaxed in the air. As it approaches the surface, the lever begins to bend and the reflected laser spot will shift in the photodetector, generating the digital signal of deflection [14]. Then a feedback loop that controls the height actuator will move the cantilever upward/downward correspondingly until the deflection signal reaches a setpoint. The distance of the movement is recorded as height information at each pixel.

The other mode is called tapping mode as the cantilever is typically oscillated mechanically by a small piezoelectric actuator. Prior to scanning, the drive frequency is set near the first resonance of the cantilever. The photodetector captures the oscillation and a built-in lock-in amplifier can yield its amplitude and phase for the height control and record. The detailed working principle is omitted here and can be found in ref [14], as it is irrelevant to our main topics. Since the probe tip just lightly taps the surface instead of scratching it, the tapping mode is more suitable for the measurement of soft materials. In addition, its setup integrating a lock-in amplifier can be easily reconfigured for other functional imaging.

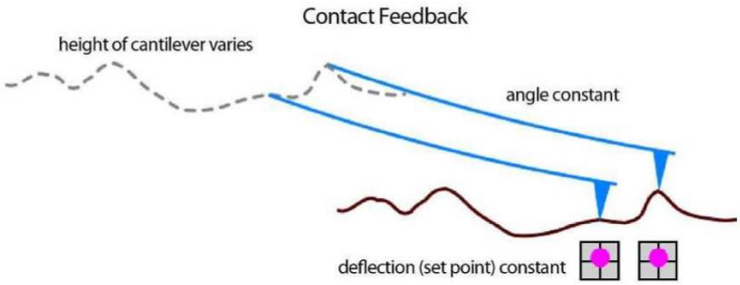

Figure 1.3. Contact mode tracking of the surface. Feedback keeps the cantilever deflection constant. This image is reprinted from [14].

Among various SPM functional imaging modes, there are three most representative methods widely used in studying functional materials. They are Piezoresponse Force Microscopy (PFM), Scanning Kelvin Probe Microscopy (SKPM), and Conductive Atomic Force Microscopy (CAFM).

1.2.1 *PFM*

Piezoelectric effect is a kind of coupling between electric field and mechanical strain that underpins a broader range of phenomena in materials without inversion symmetry [15]. As a variant of SPM, piezoresponse force microscopy based on contact mode is capable of imaging and

manipulation of piezoelectric materials and their microscopic domains [16,17]. This is achieved by bringing a sharp conductive probe into contact with a piezoelectric material and detecting the bias-induced piezoelectric surface deformation. Since the piezoresponse is very small, about several picometer (pm), it is usually boosted via resonance enhancement. In detail, a periodic bias $V_{tip} = V_{dc} + V_{ac}\cos(\omega t)$ is applied to sample via the probe, the converse piezoelectric effect results in deflection of the probe cantilever and can be detected through photodiode detector and then demodulated by a lock-in amplifier, as schematically shown in Figure 1.4. The dynamics of the interaction between the probe and sample surface can be described well by the damped harmonic oscillator (DHO) model [18],

$$A(\omega) = \frac{A_0\omega_0^2}{\sqrt{(\omega_0^2-\omega^2)^2+(\frac{\omega_0\omega}{Q})^2}} \qquad (1.1)$$

$$\phi(\omega) = tan^{-1}[\frac{\omega_0\omega}{Q(\omega_0^2-\omega^2)}] + \phi_0 \qquad (1.2)$$

where A_0, ω_0, and Q are intrinsic electromechanical response (piezoelectricity), resonant frequency (elasticity), and quality factor (energy dissipation) of the resonant system that we are interested in learning, while ω is the excitation frequency with $A(\omega)$ and $\phi(\omega)$ as the corresponding amplitude and phase that are directly measured in experiment.

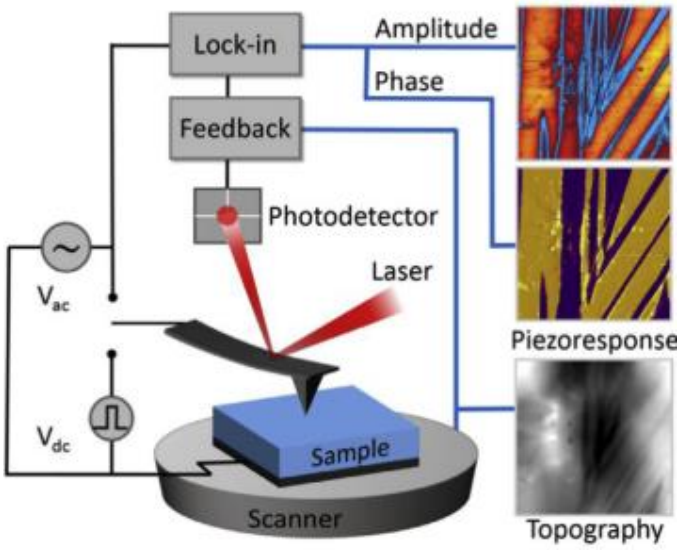

Figure 1.4. Configurations of a typical PFM system and typical PFM maps. This image is reprinted from [12].

It is worthy to note that the phase offset ϕ_0 can yield information on the polarization direction below the tip, which is valuable in the characterization of ferroelectric domains. Ferroelectric material is a special type of piezoelectric material with a spontaneous electric polarization that can be reversed by the application of an external electric field. Microscopically, adjacent dipoles pointing in the same direction will form domain structures. For so-called c- domains with polarization vector oriented normal to the surface and pointing downward, the application of a positive tip bias results in the expansion of the sample and surface oscillations are in phase with the tip voltage, $\phi = 0°$, whereas for opposite c^+ domains, $\phi = 180°$ [17]. In this way topography and ferroelectric domains can be imaged simultaneously with high resolution.

Depending on measured strain and corresponding deflection, PFM can be categorized into two types as shown in Figure 1.5. One is vertical PFM for normal strain, and the other is lateral PFM for shear strain. The combination of them can provide comprehensive information on the

underlying microstructure and microscopic processes. A detailed example will be seen in Section 2.3.

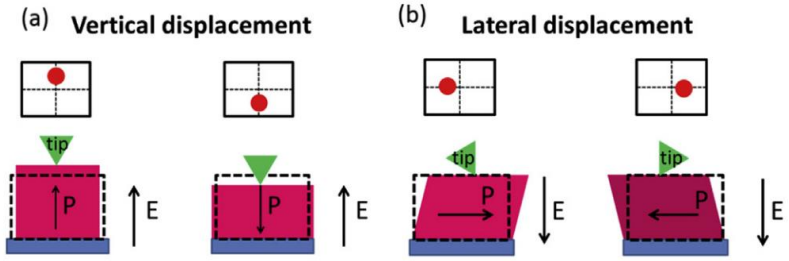

Figure 1.5. Schematics of (a) vertical and (b) lateral PFM mode. This image is reprinted from [12].

1.2.2 *SKPM*

Scanning Kelvin probe microscopy (SKPM) is a technique that measures the potential difference between the probe tip and the sample surface [19,20]. It has extensive applications in characterizing the electric and electronic properties of semiconductors and insulator materials. The technique relies on the electric force produced from the capacitor between the probe tip and the sample. By keeping tip and sample at a constant separation, the electric force is proportional to the square of voltage according to Equation (1.3), where $V = (V_{DC} - V_{sp}) + V_{AC} \sin(\omega t)$.

$$F = \frac{1}{2}\frac{\partial(C)}{\partial(z)}V^2 \quad (1.3)$$

The potential difference V_{sp} is what we are trying to measure. V_{DC} and V_{AC} are DC and AC voltages applied to help the measurement. C is the capacitance and z is the distance. After squaring V into Equation (1.3) and rearranging the result, we have

$$F = \frac{1}{2}\frac{\partial(C)}{\partial(z)}\left(\left[\left(V_{DC} - V_{sp}\right)^2 + \frac{1}{2}V_{AC}^2\right] + 2\left[\left(V_{DC} - V_{sp}\right)V_{AC}\sin(\omega t)\right] - \frac{1}{2}V_{AC}^2\cos(2\omega t)\right) \quad (1.4)$$

The principle of this method is to set the ω to the resonant frequency of the cantilever so that DC and 2ω signals can be filtered out [14]. We only need to focus on the middle term of the above

equation and tune V_{DC} to minimize the force, which means $V_{DC} = V_{sp}$. In the most commonly used dual-pass SKPM, the probe works in tapping mode to obtain surface topography information in a first pass before lifting to a certain height to measure the surface potential [14], as schematically shown in Figure 1.6. During the second pass the vibration of the probe comes from the applied AC bias instead of a piezo-driver.

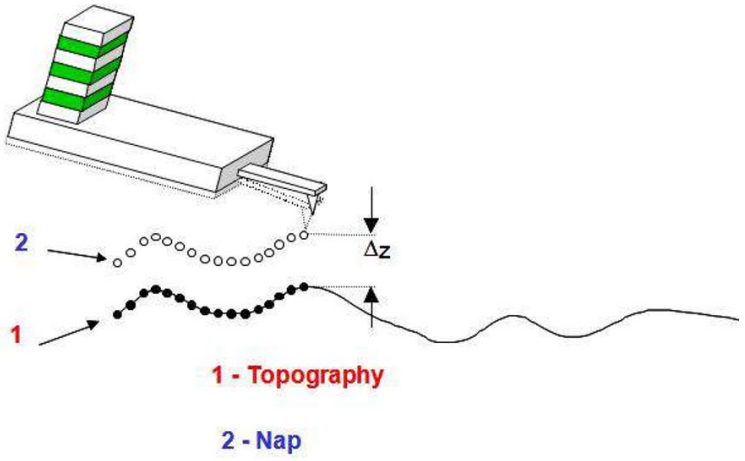

Figure 1.6. Schematic of two-pass mode used in SKPM. For each scan line, the system captures the topography. The system then retraces that topography on the same line in order to keep a constant distance from the sample surface. This image is reprinted from [14].

1.2.3 *CAFM*

Conductive atomic force microscopy (CAFM) is a technique implemented for the simultaneous measurement of the topography of a material and its electric current information [21,22]. Unlike scanning tunneling microscopy (STM) which is difficult to discern if the current fluctuations originate from surface roughness or intrinsic inhomogeneities of local conductivity, CAFM avoids this confusion by using two different detection systems, scanning for the topography and preamplifier for the current [7,22]. In CAFM, the probe is brought close to the surface until its deflection reaches the set value, which is the principle of contact mode, while a voltage is applied between the probe covered with conductive coating and the substrate. The current is

recorded at each pixel using a current-to-voltage preamplifier, yielding a current map with spatial resolution close to the tip radius [7], as schematically shown in Figure 1.7. Besides, a local I-V curve can be acquired by holding the tip position and sweeping the bias.

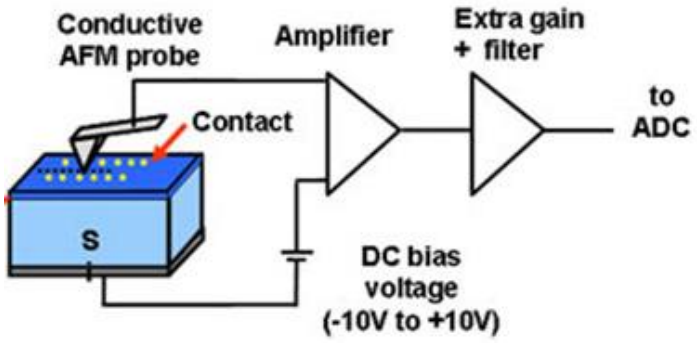

Figure 1.7. Schematic of CAFM. This image is reprinted from [23].

1.3 CURRENT CHALLENGES OF CHARACTERIZATION WITH SPM

Although it is an exciting time in the field of SPM and a toolbox of SPM techniques has been invented and played important roles in materials research, the demand of advanced SPM methods for informative as well as efficient characterization never stopped. In fact, conventional SPM techniques have two long-standing limitations: (1) The data collected for each pixel is limited and vulnerable to measurement errors, making it hard for deep analysis. For instance, PFM characterizes the dynamics of the cantilever motion through lock-in amplifiers, where the responses probed from the sample are confined to a particular excitation frequency and harmonic, risking missing critical information. The resonance tracking based on multiple PID controllers also results in a system with complex configuration, slow scanning speed, and susceptible to crosstalk; (2) their operations are very empirical and rely heavily on users' experiences in choosing appropriate imaging modes and tuning the experimental parameters. During scanning, parameters

are preset and remain static, incapable of responding to the probed data in an adaptive manner unless user intervenes. Such operations are tedious and very often users would miss key information and critical spots during experiments, only to realize later, often too late to get back to the exact same spots.

These limitations can be better illustrated by the perovskite $CH_3NH_3PbI_3$ in Chapter 2, which is widely pursued for the next generation solar cells [24–30], yet suffering from degradation due to environmental sensitivity. The microscopic processes in $CH_3NH_3PbI_3$ are highly complex, involving generation and transport of photo-carriers, ionic migration, defects, and possible polarization domain and switching [31–35]. In order to understand the effects of such microscopic processes on photovoltaic conversion, detailed local distributions need to be mapped out, including microstructure morphologies, surface potential, photocurrent, piezoresponse, and ionic activities. Obtaining such comprehensive information with high fidelity using a conventional SPM requires continuous scanning as well as switching among various imaging and spectroscopic modes, which is not only boring but also virtually impossible to do for sensitive materials like $CH_3NH_3PbI_3$ without significant degradation. The issues discussed here are by no means limited to halide perovskite and are in fact quite common for advanced materials and other complex systems.

Nonetheless, the recent advances in big data inspire us to overcome such limitations by heading for a data-driven SPM. Nowadays, the fusion of scientific research and big data has provided unprecedented opportunity for accelerated discovery, understanding, and innovation [36–42], yet it also imposes new challenges for scientists to adjust, adapt, and thrive in the face of daunting data volume [42,43]. As scanning probe microscopy can acquire high-resolution images and spectroscopy [44], it also provides us an ideal playground for big data methodology. To enable such vision and handle those challenges, there will be three key questions to be answered:

(1) how do we develop innovative experimental methodologies based on SPM to acquire high quality, efficient, and physically relevant scientific data to enable deep analysis. As is discussed, the information yielded from conventional techniques such as PFM, SKPM, and CAFM is insufficient due to its configuration, while multi-dimensional physical datasets are highly needed in current research for advanced material with complex coexisting effects. To this end, we design a new scheme for excitation and data acquisition of SPM from both hardware and software, which will be demonstrated in Chapters 3 and 4.

(2) how can we accelerate and enhance physical understanding from SPM data by informed and targeted big data analytics. Ever increasing hardware capabilities and computation powers have made acquisition and analysis of big scientific data routine. A good example in this field is the recently developed G-mode SPM that relies on brute force data analytics without pre-committing to a particular physical process [42,45]. However, much of the data acquired often turns out to be redundant, noisy, and/or irrelevant to the problems of interests, and it remains nontrivial to draw clear mechanistic insights from brute force data analytics [37,42,46]. As such, there is a strong desire to fuse scientific knowledge of physics, chemistry, and biology into big data via data mining, correlation analysis, and unsupervised classification [37,38,40,41,46,47], and thus from brute forces to informed and/or targeted strategies. In Chapter 3 and 4, we adopt a more targeted approach, taking advantage of both physical understanding and data power.

(3) how can we make better use of big data to repay the development of SPM techniques. It is quite common in the research that only a small portion of data is finally published, while most of them just fade into memory. Thanks to the recent advances in the field of artificial intelligence (AI), we are inspired to design and train an AI-SPM by reusing data accumulated in our lab in the past ten years to realize autonomous material classification and feature recognition. As mentioned

above, operating SPM is tedious at times for experienced users, while very often important yet subtle information is overlooked by novices when conducting experiments, and insights are only realized during the post data processing afterward, which is often too late - it is virtually impossible to get back to the critical locations again for further probing, such as defects, heterogeneities, and interfaces, where the most interesting physics occur. Our proposed data-driven AI-SPM in Chapter 5 will be a feasible solution for this challenge and can be applied to broader material systems, and thus harnessing the data revolution in materials research.

1.4 Overview of Dissertation

This dissertation covers the research materials over the past four years on the development of advanced SPM methodology and its application on functional materials. Chapters 2 first demonstrates great potential of SPM techniques in studying functional materials at the nanoscale by using perovskite solar cells as an example, and highlight some of the recent progress along the line of investigation toward local probing of photocurrent, work function, ionic activities, polarization switching, and chemical degradation. Detailed evidence is then presented that single crystal perovskite $CH_3NH_3PbI_3$ consists of ferroic domains with alternating polar and nonpolar orders, in contrast to previous experimental and theoretical expectations, and polar domains possess reduced photocurrent. The major limitation and emerging challenges of current SPM techniques with respect to data acquisition, quality and analysis are discussed, which becomes the primary motivation of the following chapters.

In Chapter 3, a developed sequential excitation technique for collecting high-quality, efficient and physically relevant data with SPM is introduced in detail. A rough piezoelectric material has been probed using such a strategy, yielding the map of its intrinsic electromechanical properties at the nanoscale with high fidelity, where conventional methods fail. Finally, principal component

analysis (PCA) is carried out, which not only speeds up the data analysis by four orders of magnitude, but also allows a clear physical interpretation of its modes in combination with DHO analysis.

Chapter 4 demonstrates an upgraded version of the above technique, that is, a high-throughput sequential excitation which can acquire the spectrum of data on each pixel over a range of frequencies excited sequentially. Different from the previous one requiring multiple scans, the new method is capable of accurately capturing the cantilever dynamics in just one scan, showing a scanning speed comparable to commercial dual-amplitude resonance tracking (DART) [48], yet with much enhanced spatial resolution and quantitative accuracy.

In Chapter 5, an artificial intelligence scanning probe microscopy (AI-SPM) is proposed, trained and tested with a large amount of mapping data accumulated in our lab in the past decade. The AI-SPM can not only do pattern recognition and feature identification for ferroelectric materials and electrochemical systems, but also respond to classification via adaptive experimentation with additional probing at critical domain walls and grain boundaries, all in real-time on the fly without human interference.

Finally, Chapter 6 summarizes the research covered in this dissertation and discuss some future directions for further studies.

Chapter 2. STUDYING PEROVSKITE SOLAR CELLS WITH CONVENTIONAL SPM

2.1 MOTIVATION

As one of the most promising candidates in the next generation photovoltaic technologies, halide perovskite solar cells (PSCs) based on methylammonium lead triiodide ($CH_3NH_3PbI_3$) and related materials have been continuously attracting attention. Since it was first reported in 2009 [24], there has been a spectacular rise in its photovoltaic conversion efficiency (PCE) in past 10 years, from 3.81% measured in laboratory to over 21% [49]. Such a rapid advance in photovoltaic performance with ultimate low-cost synthesis process is unparalleled in other types of solar cells. However, as the best PCE reaches commercial viability, intense worldwide competition for the new performance record has gradually cooled down and much of research begins to focus on achieving longer term stability or more fundamental understanding of photovoltaic characteristics of halide perovskite, which is more urgent for their commercialization.

Macroscopic tests have been extensively adopted to investigate PSCs, such as current-voltage (IV) curve, quantum efficiency, impedance photoluminescence, and transient photovoltage/current decay test, but local heterogeneity is usually averaged out at the device level, making it difficult to correlate microstructure and performance. Thus, there is an increasing attention on microscopical measurements. A number of studies have attempted to reveal the microstructural mechanisms responsible for the macroscopic performance by mapping the composition, phase, and structure of photovoltaic materials via various imaging and analytic techniques at nanoscale. Nevertheless, the connection between the local structures and global performance is still vague, and thus an effective way to probe the local photovoltaic characteristics of solar cells that can be directly traced to their microstructures is highly demanded. For a system as intricate as halide

perovskite solar cells with interacting photo-carriers, ionic defects, spontaneous polarization, and molecular vibration and rotation on top of numerous grains, grain boundaries, and interfaces, such correlation is necessary for elucidating their macroscopic photovoltaic properties, device efficiency, and degradation. In this Chapter, we will first show the great potential of SPM techniques to realize such promise at the nanoscale, then highlight progresses on the study of ferroic nature of PCSs, and finally summarize challenges along this line of investigation.

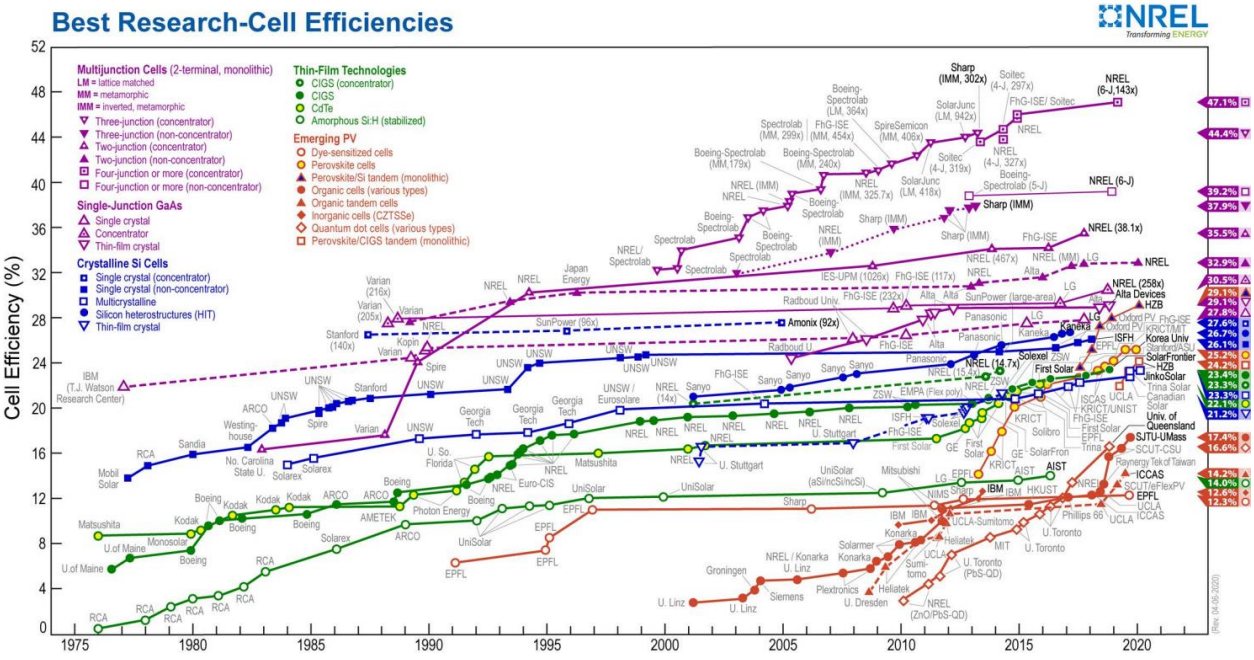

Figure 2.1. Perovskite solar cells have increased in power conversion efficiency at a phenomenal rate compared to other types of photovoltaics. This image is reprinted from [50].

2.2 MICROSCOPIC PHOTOVOLTAIC PROPERTIES

As introduced in the previous section, SPM is a powerful as well as flexible platform that allows exploration of material at nanoscale. For solar cells, monitoring its morphological, structural and electrical changes under real working conditions (illumination) is significant to understand photovoltaic characteristics and degradation process. Here I want to introduce some

simple results from the performance and stability study of $CH_3NH_3PbI_3$ thin films via photo-assisted SPM techniques, including photoconductive atomic force microscopy (pcAFM), SKPM, and local IV curve measurement. All of them could play important roles in investigating the local properties corresponding to macroscopic figure of merit like short-circuit photocurrent density (J_{SC}) and open-circuit voltage (V_{OC}). Probing such structure/function relationships is an important step toward the improvement of intact devices, and here we just want to showcase the capability of the techniques, without diving into detailed analysis and interpretation for mechanistic understanding.

2.2.1 *Results of pcAFM, SKPM and local IV curve*

$CH_3NH_3PbI_3$ samples used in these experiments were synthesized via one-step solution processing by our collaborator Prof. Jinjin Zhao, where $CH_3NH_3PbI_3$ is spin-coated either on a transparent poly(3,4-ethylenedioxythiophene):polystyrenesulfonate (PEDOTS:PSS) substrate or on a titanium dioxide (TiO_2) substrate. These substrates serve as a hole/electron transport layer to reduce the recombination of the free charge carriers (electrons and holes) with their counterparts on defects.

Although conductive atomic force microscopy (cAFM) can map local current distribution with a high sensitivity of picoampere (pA), it cannot provide information on how the solar cell operates. To this end, we can equip cAFM setup with additional illumination focusing on the sample, which is so-called photoconductive atomic force microscopy (pcAFM), as shown in the Figure 2.2. This technique has been utilized to evaluate variations in local photocurrent characteristics with inhomogeneities in the film structures such as grains and grain boundaries.

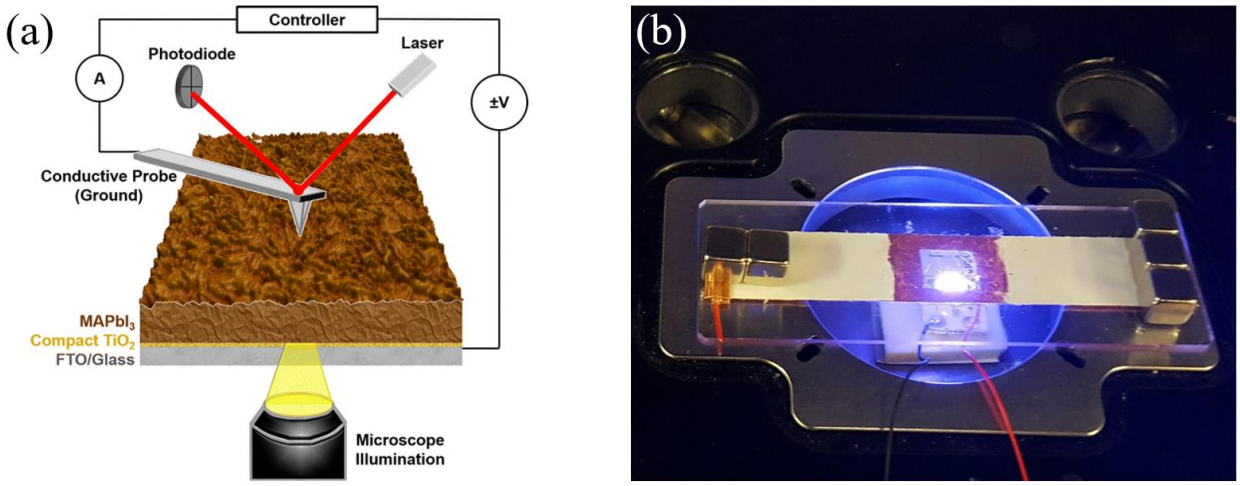

Figure 2.2. (a) Schematics of pcAFM measurement for perovskite solar cells. This image is reprinted from [51]. (b) Picture of bottom illumination for photo-assisted SPM.

In this experiment, pcAFM measurements were carried out on an Asylum Research MFP-3D AFM using ORCA conductive AFM holder with a gain of 5×10^8 V/A and a Nanosensors PPP-EFM conductive probe with a spring constant of 3.18 N/m. During the measurement, the sample was biased with specified DC voltage and illuminated in air through the bottom FTO cathode using an unfiltered CREE MK-R 12V LED.

Compared with grains, the probe may collect more current near the grain boundaries due to the larger contact area from adjacent grains. However, Figure 2.3(b) reveals that the transport of photocurrent is clearly hindered by some grain boundaries (marked by blue), which excludes possible influences of the topography on pcAFM. The similar contrast of photocurrent distribution on grains and grain boundaries have been also reported before [51], and there is significant variation from grain to grain as well.

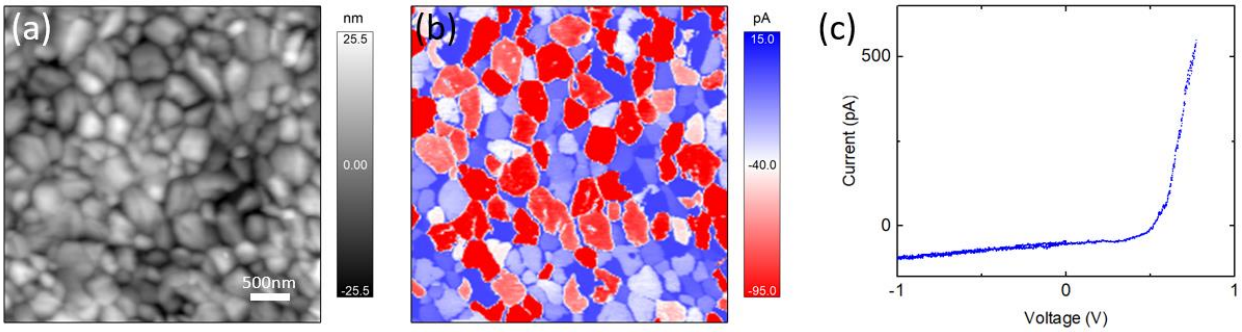

Figure 2.3. pcAFM results of $CH_3NH_3PbI_3$. (a) topography; (b) photocurrent; (c) local IV curve for single pixel.

Besides, point-wise spectroscopy study like local IV curves measurement is accessible by measuring current while simultaneously sweeping DC voltage biased via bottom electrode, as shown in Figure 2.3(c). For example, high hysteresis of IV curve at grain boundaries was reported using this method [52]. However, it is relatively inefficient for a high density grid of points to achieve desired spatial resolution. Given that a single local IV curve usually takes about 1s for each pixel, acquiring all curves for a 128×128 pixels map would take around 4.5hr. Such a long measurement is not practical due to drift of scanning and degradation of materials, which can be seen in the next subsection 2.2.2.

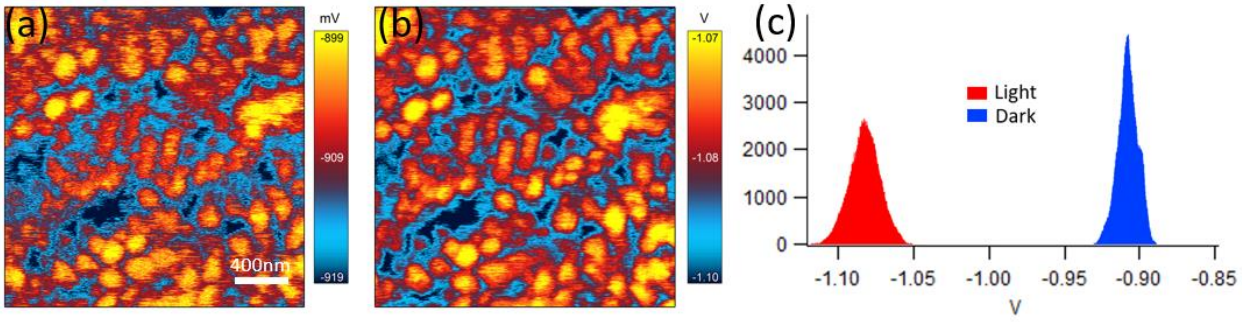

Figure 2.4. Surface potential mappings of $CH_3NH_3PbI_3$ films on the FTO/PEDOTS:PSS substrate. (a) the dark condition, (b) the light condition, (c) the histogram of two mappings.

As the current flow is inherently global, pcAFM is only pseudo-local in nature. Thus, alternative techniques like SKPM would be a good supplement. Since the absolute surface

potential of thin film varies due to the nonuniformity of topography, the potential difference between dark and light conditions is important. Figure 2.4 shows results of $CH_3NH_3PbI_3$ thin films with/out light and the comparison of their histograms. Under light illumination, holes migration downward to the hole collecting PEDOTS:PSS, while electrons migrate upward to the sample surface, resulting negative shift in surface potential as observed. In addition, SKPM can be used to probe the cross-section of a perovskite solar cell, yielding valuable information on the interface, such as band bending and charge accumulation [53].

2.2.2 *Stability study*

In spite of the fast power conversion efficiency evolution of PSCs, it is widely recognized that the perovskite films normally lack chemical and structural stability, which means they will readily degrade under moisture, heat, light and the presence of an electric field. Although there are some reports claiming the stability could be remarkably enhanced through composition engineering, the cost is often their efficiency. Except tuning PSC's compositions, appropriate surface modification can also greatly improve the film's stability. More importantly, this strategy can change only the surface properties of the films without significantly modifying the bulk's photoelectric properties [54].

In the following controlled experiment, pcAFM mappings show an obvious improvement of stability based on surface modification. Two samples are provided by our collaborator Dr. Hongmei Zhu and Prof. Wei Chen of Huazhong University of Science and Technology. The control group is pristine $CH_3NH_3PbI_3$ film (or call it $MAPbI_3$ film), while the experimental group is the film modified with a hydrophobic organic acid molecule, 4-dimethylaminobenzoic acid (4-DMABA) on its top surface. A white light LED lamp with a light intensity of 80 mW/cm^2 illuminated the samples through an optical fiber. The relative humidity of the testing room was

well controlled around 34% by two dehumidifiers. The p-type NiMgLiO coated FTO glass served as the anode, while the SPM probe served as the cathode. This configuration could be regarded as an electron transport layer-free PSC.

The pcAFM results in Figure 2.5 shows photocurrent mapping and corresponding topography information along with aging time for (a-b) the pristine $CH_3NH_3PbI_3$ film and (c-d) the 4-DMABA modified $CH_3NH_3PbI_3$ film. The scale bars in (a–d) are 1 μm. The dashed squares highlight the same region (position shifted) of the films during pcAFM tests due to inevitable drift of long-time scanning. Figure 2.6 shows that the average photocurrent decreases quickly from an initial value of 231.841 pA to 22.388 pA after 3 hours, which indicates the severe degradation of the pristine $CH_3NH_3PbI_3$ film. Consistently, the film morphology becomes much blurrier and the grain boundaries gradually melt together with increasing aging time.

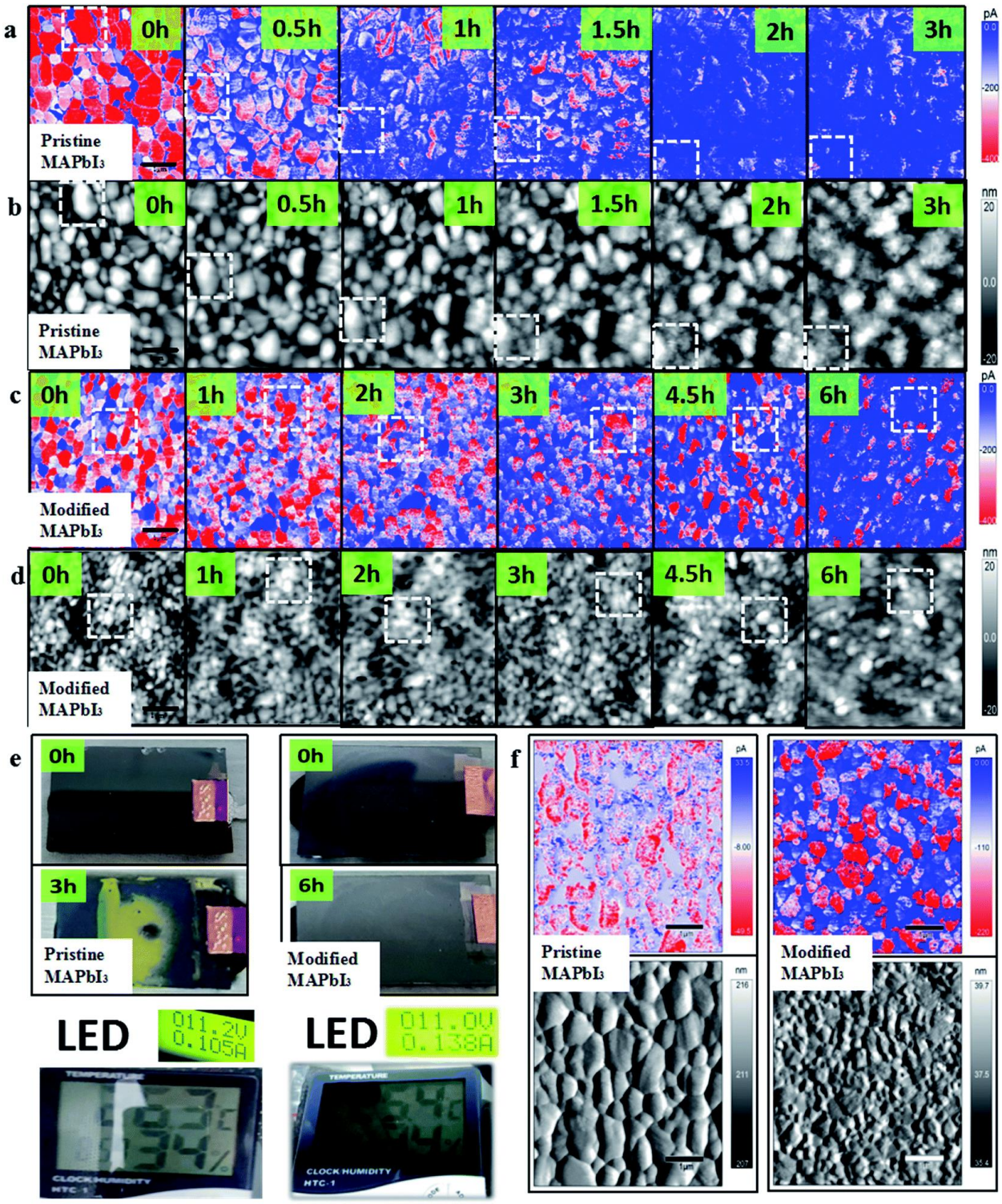

Figure 2.5. The pcAFM mapping results and corresponding topography information along with aging time for (a and b) the pristine $CH_3NH_3PbI_3$ (or call it $MAPbI_3$) film and (c and d) the 4-DMABA modified $CH_3NH_3PbI_3$ film. The scale bars in (a–d) are 1 μm. The dashed squares highlight the same region (position shifted) of the films during pcAFM tests. (e) Optical images

of the two kind perovskite films before and after several hours of aging. (f) The pcAFM results and corresponding topography information of the two kinds of perovskite film after keeping them under ambient conditions in the dark for 1 month.

For the modified $CH_3NH_3PbI_3$ film, Figure 2.5(c-d) show its photocurrent and topography maps under the same aging conditions for 6 hours. It is clear that the average photocurrent almost maintains its initial value after 3 hours of aging; and even after aging for 6 hours, the average photocurrent of the modified $CH_3NH_3PbI_3$ film (141.651 pA) is still much higher than that of the pristine one after aging for 3 hours (22.388 pA). Consistently, the morphology change for the modified $CH_3NH_3PbI_3$ film is also much smaller and slower, indicating its retarded decomposition in comparison to the pristine $CH_3NH_3PbI_3$ film.

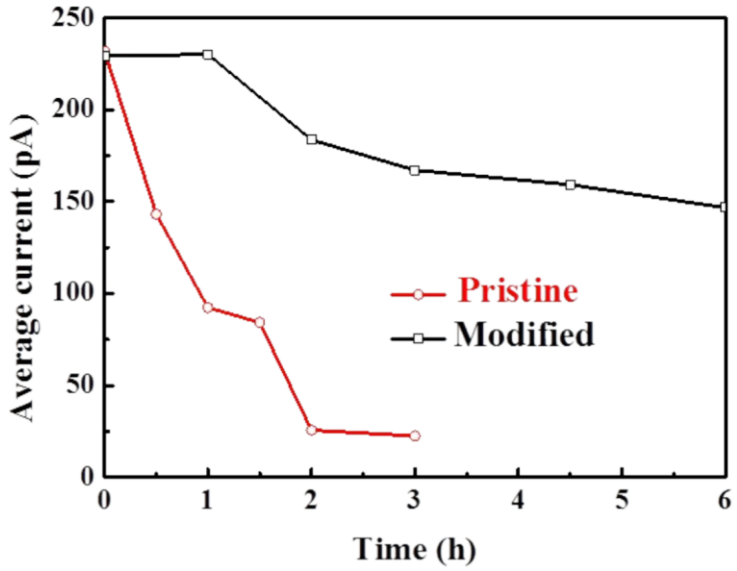

Figure 2.6. The degradation of average current of the pristine and 4-DMABA modified $CH_3NH_3PbI_3$ films along with the aging time during the pcAFM tests. Average current is obtained from mapping results.

Moreover, the distinct difference in decomposition speed between the pristine and modified $CH_3NH_3PbI_3$ films can also be identified by looking at their optical images which are shown

in Figure 2.5(e). The color of the pristine $CH_3NH_3PbI_3$ film turned to yellow after 3 hours of light-soaking, indicating its decomposition from perovskite to PbI_2, while the color of the modified $CH_3NH_3PbI_3$ film after 6 hours of aging was nearly unchanged. After we left the pristine and modified $CH_3NH_3PbI_3$ films in a dark drawer without humidity control for 1 month, and then tested the films using pcAFM, we found that the average photocurrent of the pristine film had degraded to 12.671 pA while the modified film still held a value of 105.734 pA (Figure 2.5(f)).

2.3 FERROELASTICITY AND FERROELECTRICITY

After getting a taste on the capability of SPM, we move on to more serious investigation, and focus on a problem that is hotly debated in the community. Indeed, even after extensive researches in the past a few years, the very crystalline structure of $CH_3NH_3PbI_3$ remains ambiguous and generates a heated debate on whether ferroic order of $CH_3NH_3PbI_3$ is polar or not. Ferroic refer to a generic class of materials with ferro-(magnetic, electric, elastic) orders [55]. As mentioned earlier, ferroelectricity means materials have a spontaneous electric polarization, which is hypothesized to play an important role in reducing charge recombination for solar cells [56,57], and therefore is of great interest for further investigation [58]. In contrast, ferroelasticity is a phenomenon in which a material exhibits a spontaneous strain [59]. In brief, ferroelastic material is not necessary to have polar order, while the ferroelectric must. To date, there has been little agreement on the polar nature of $CH_3NH_3PbI_3$. As summarized in Table 2.1, advocates for either opinion can find their evidences in X-ray and neutron diffractions [60–67], optic second harmonic generation (SHG) [68,69], macroscopic pyroelectric and ferroelectric measurement [68,70,71], microscopic piezoresponse force microscopy (PFM) [34,72–80], and density functional theory (DFT) and molecular dynamics (MD) simulations [57,80–85]. Although early experimental and computational work observed

switchable ferroelectric-like domains [72,86], implying possible polar order, this concept has been challenged soon by other studies demonstrating the existence of nonpolar ferroelasticity [76,78].

In this section, we will present compelling evidences that $CH_3NH_3PbI_3$ crystals self-grown on FTO/TiO_2 substrate consist of ferroic domains with alternating polar and nonpolar orders, in contrast to previous experimental and theoretical expectations, and polar domains possess reduced photocurrent. We found that polar and nonpolar orders of $CH_3NH_3PbI_3$ can be distinguished from their distinct lateral piezoresponse, energy dissipation, first and second harmonic electromechanical couplings, and temperature variation, even though their difference in crystalline lattice is very subtle, and they possess two-way memory effect through cubic-tetragonal phase transition. These findings resolve key questions regarding polar nature of $CH_3NH_3PbI_3$ and its implication on photovoltaics, reconcile contradictory data widely reported, and point a direction toward engineering ferroic domains for enhanced PCE.

Table 2.1. Survey of selected literature report on the structure of tetragonal $CH_3NH_3PbI_3$.

Technique	Non-polar I4/mcm	Polar I4cm	Noncommittal
X-ray and neutron diffractions	*J. Mater. Chem. A*. **1**, 5628 (2013); *Chem. Commun.* **51**, 4180 (2015); *Adv. Funct. Mater.* **25**, 2378 (2015); *Sci. Rep.* **6**, 35685, (2016); *Sci. Rep.* **6**, 30680, (2016);	*Inorg. Chem.* **52**, 9019 (2013); *CrystEngComm*. **17**, 665 (2015).	*J. Mater. Chem. A*. **3**, 9298 (2015)
Optic SHG	*J. Phys. Chem. Lett.* **7**, 2412 (2016).	*Proc. Natl. Acad. Sci.* **114**, E5504 (2017).	
Macroscopic measurements	*Appl. Phys. Lett.* **106**, 173502 (2015); *ACS Energy Lett.* **1**, 142 (2016).	*Proc. Natl. Acad. Sci.* **114**, E5504 (2017).	
Microscopic PFM	*J. Phys. Chem. Lett.* **6**, 1155 (2015); *Appl. Phys. Lett.* **113**, 072102 (2018); *Nat. Mater.* **17**, 1013 (2018); Adv. Optical Mater., **7**, 1901451, (2019); *Energy Environ. Sci.* **12**, 2537 (2019).	*J. Phys. Chem. Lett.* **5**, 3335 (2014); *J. Phys. Chem. Lett.* **6**, 1729 (2015); *J. Mater. Chem. A*. **3**, 7699 (2015); *J. Mater. Chem. A*. **4**, 756 (2016); *Nanoscale* **9**, 3806 (2017); *Energy Environ. Sci.* **10**, 950 (2017)	*J. Phys. Chem. C*. **120**, 5724 (2016); *Science Advances* **3**, e1602165 (2017)
TEM			*Nat. Commun.* **8**, 14547 (2017)
DFT and MD simulations	*J. Mater. Chem. A*. **3**, 8926 (2015)	*Nano Lett.* **14**, 2584 (2014); *Chem. Mater.* **26**, 6557 (2014); *J. Phys. Chem. Lett.* **6**, 693 (2015); *J. Phys. Chem. Lett.* **6**, 2223 (2015); *J. Phys. Chem. Lett.* **6**, 31 (2015); *J. Phys. Chem. Lett.* **6**, 1155 (2015)	

2.3.1 *Materials and Methods*

Crystal growth: In this work, we probe $CH_3NH_3PbI_3$ films self-grown on FTO/TiO_2 substrate provided by our collaborator Prof. Jinjin Zhao. These $CH_3NH_3PbI_3$ are single crystalline in nature with good ambient stability and respectable PCE [87]. They give us a clean crystalline system to study without the complexity of grains and grain boundaries that are common in polycrystalline $CH_3NH_3PbI_3$, and they provide a photovoltaic relevant system to correlate photocurrent with ferroic domains.

SEM and EBSD: The morphologies of $CH_3NH_3PbI_3$ crystals were examined by scanning electron microscopy (SEM, ZEISS GeminiSEM 300). Electron backscatter diffraction (EBSD) Kikuchi patterns were obtained on Oxford NordlysMax3 detector.

Optical microscopy: Polarized optical microscopy was carried out in reflection-mode using a commercial Nikon Eclipse 80i microscope with a Lumenera Infinity 1-3C camera. The sample was fixed and illuminated by Nikon LV-UEPI2 universal EPI illuminator (a LV-HI50W 12V-LI Halogen Lamp) through top objective lens. The domain pattern was observed by rotating the inside polarizer at an angle θ with respect to the x-axis of images.

Scanning probe microscopy:

SPM measurements, including topography, PFM, SKPM, and pcAFM, were performed on Asylum Research Cypher AFM and MFP-3D Bio AFM. Nanosensors PPP-EFM conductive probes with a spring constant around 2.8 Nm^{-1} were used for most measurements. Olympus AC160TS-R3 probes with resonance frequency of 288 kHz were also used.

Topography: For topography mapping, the less invasive tapping mode was adopted to avoid possible damage to samples. The probe was excited at its resonance frequency and the scan rate is set to be about 1.0 Hz for visualizing the domain structure. The amplitude channel was used to

present topography due to its higher sensitivity to the topography variance. Topography mappings were also acquired from the simultaneous deflection channel of PFM and pcAFM.

PFM: For PFM measurements, we employed four different techniques. The first is single frequency PFM scanning, where the driving frequency is preset and constant during the scanning, thus a very fast scan rate up to 2 Hz can be reached without sacrificing image quality. The vertical and lateral PFM responses were detected by using vertical and lateral deflection signals of cantilever. An AC voltage of 2 V was applied to the probe near the sample–probe resonance frequency f_0 to enhance the sensitivity. The f_0 of the first and second mode of vertical PFM are 357 kHz and 1079 kHz, separately. The f_0 of lateral PFM is 658 kHz.

The second is the dual-amplitude resonance tracking (DART) PFM scanning [48]. Since microstructure determine the properties of materials, the resonance frequency of tip-sample system varies at the nanoscale scale. Therefore, driving at a fixed frequency can sometimes lead to enormous measurement errors, which is so-called topographic cross-coupling. As shown in Figure 2.7, the principle of DART is to apply two AC bias to the probe at two frequencies, f_1 and f_2. One on either side of resonance peak f_0. The corrected amplitude, phase, resonant frequency, and quality factor can be calculated with two data points (A_1, f_1) and (A_2, f_2) by using DHO model as long as the f_0 is between them [88], otherwise the model will fail. During scanning, the difference between the two amplitudes A_1 and A_2 can be used as the feedback signal to track the resonance peak. For example, when $A_2 > A_1$, the excited frequencies will shift to larger side until new amplitudes become nearly equivalent, thereby tracking the new f_0. Compared with single frequency mode, DART can provide relatively reliable data at a slower scanning speed due to the lag of tracking. In this experiment, we used an AC amplitude of 1 V for both vertical and lateral PFM and a scan rate below 0.6 Hz to ensure resonant frequency tracking.

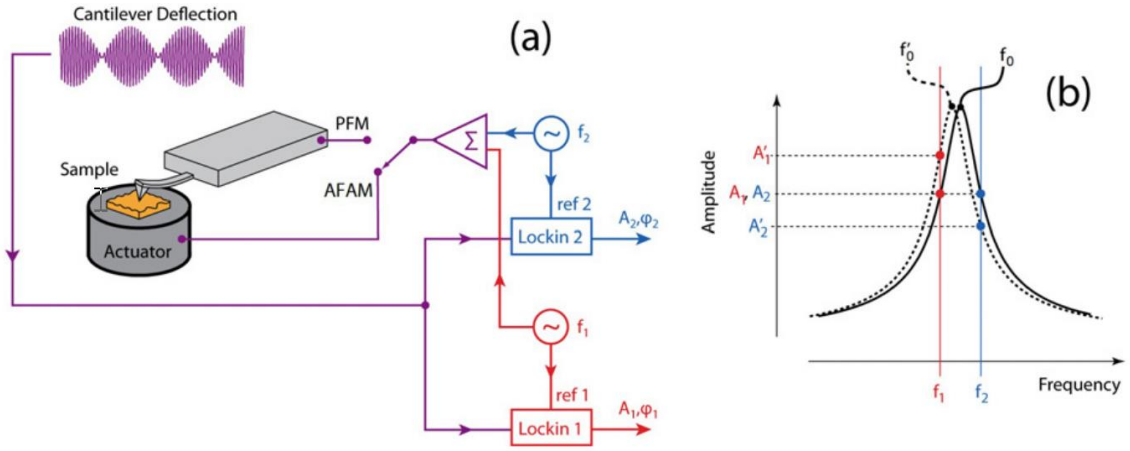

Figure 2.7. (a) Schematic of a DART experimental setup using two sinusoidal drives with frequencies such that one (f_1) is below and the other (f_2) is above the contact resonance frequency (f_0) of the cantilever. (b) An illustration of the principle of DART. This image is reprinted from [48].

The third is a home-developed first and second harmonic PFM scanning, which was implemented using a Zurich Instrument HF2LI lock-in amplifier in combination with Cypher AFM. As electromechanical strains may also originate from induced ionic motion and many groups have reported active ion activity inside $CH_3NH_3PbI_3$, examining first and second harmonic strain responses is necessary and helpful to distinguish them [89,90]. Electromechanical strains from both induced and spontaneous polarizations are electrostrictive, which is given by $\epsilon \propto P^2$. $P = \chi E + P_S$ is the polarization consisting of induced polarization χE and spontaneous polarization P_S, and χ is the dielectric susceptibility. Suppose that $E = E_0 \sin(\omega t)$ is applied AC electric field via the probe, ϵ will consist of first harmonic term $\epsilon_1 \sim \chi P_S E_0 \sin(\omega t)$ and second harmonic term $\epsilon_2 \sim \chi^2 E_0^2 \cos(2\omega t)$. For typical ferroelectric materials exhibiting large spontaneous polarization, ϵ_1 is overwhelmingly greater than ϵ_2, while for non-ferroelectric materials ϵ_2 is not negligible due to active ionic motion. In order to obtain the second harmonic signals, the sample response was

measured at f_0 while the excitation voltage was applied at $f_0/2$, as schematically shown in Figure 2.8.

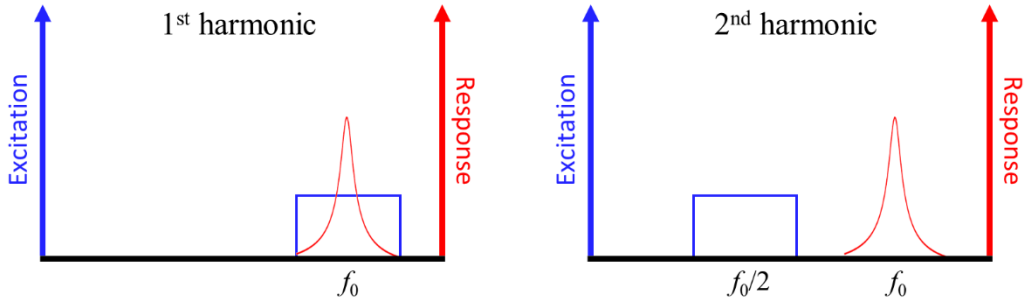

Figure 2.8. Schematics of first and second harmonic piezoresponse measurements.

The fourth is the pointwise first and second harmonic PFM measurement, which can eliminate crosstalk with topography. For each single point, a series of AC voltages from 0.5 to 2.9 V were applied to the probe with an increment of 0.3 V. At each voltage step, the probe-sample system is excited around f_0 first and then around $f_0/2$, thus generating two set of tuning data around f_0. The corresponding first and second harmonic responses, including amplitude, quality factor, and resonance frequency, were extracted by fitting the raw data with the DHO model.

pcAFM: pcAFM mapping is based on the cAFM mode using Asylum Research ORCA module, which includes a transimpedance of an amplifier with a gain value of 5×10^8volts per amp. The samples were illuminated by two different sources depending on SPM. For Cypher, the built-in LED sheds light on the sample from the top. For MFP-3D Bio, the illumination was from below through a glass fiber connected to a Nikon C-HGFIE illuminator with an ultrahigh pressure 130W mercury lamp inside. A small DC voltage were applied to the FTO substrate when necessary.

SKPM: For SKPM measurements, the probe scanned the surface topography using tapping mode first and then a 1V AC voltage was applied on the probe near its resonance frequency to measure the sample surface potential distribution through a DC voltage feedback loop.

Local heating: For local phase transition measurement, the nanoscale area was heated via an Anasys Instruments AN2-300 thermal probe with a resistance about 1kΩ on the MFP-3D Bio AFM. Temperature can be roughly regulated by changing the amplitude of AC voltages applied to the probe. The method was based on contact mode so that clear topography mapping can be collected from deflection signal.

Global heating: All larger scale phase transition measurements were conducted by heating up the whole airtight sample chamber controlled by an environment controller inside Cypher AFM. A feedback loop can maintain the set temperature with a precision of 0.01℃, which may lead to some noises for PFM mapping due to thermal expansion. The sample temperature was not calibrated due to the tight constraint of the chamber. After setting the target temperature each time, we waited for at least 1 min before any measurement so that there is enough time for thermal equilibrium.

2.3.2 *Ferroic domain structures*

We first present unambiguous ferroic domain structures widely observed in our crystals via SEM, topography, and polarized optic microscopy (Figure 2.9 and Figure 2.10), which is broadly aligned with recent literatures [76–78,91], yet on larger scales with richer varieties. These characteristic lamellar patterns are often seen in ferroics [55,92–94].

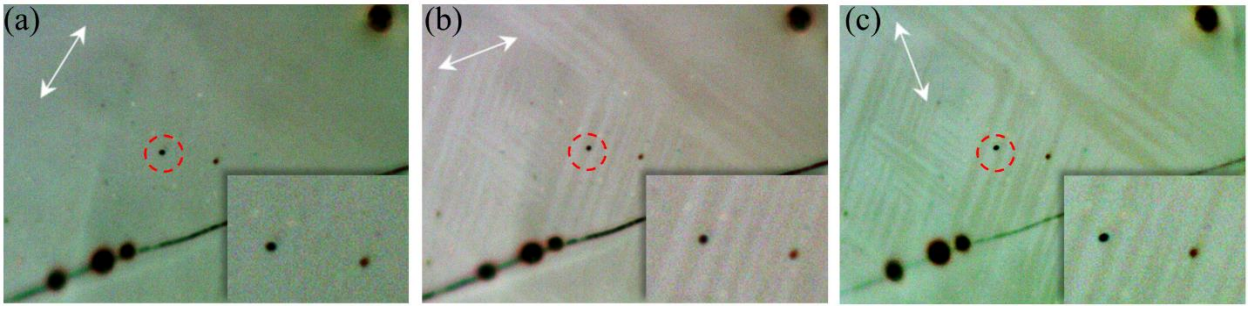

Figure 2.9. Polarized optic microscopy images showing alternating contrast of $CH_3NH_3PbI_3$ crystal when the optical polarizer is rotated (directions are marked by white arrows).

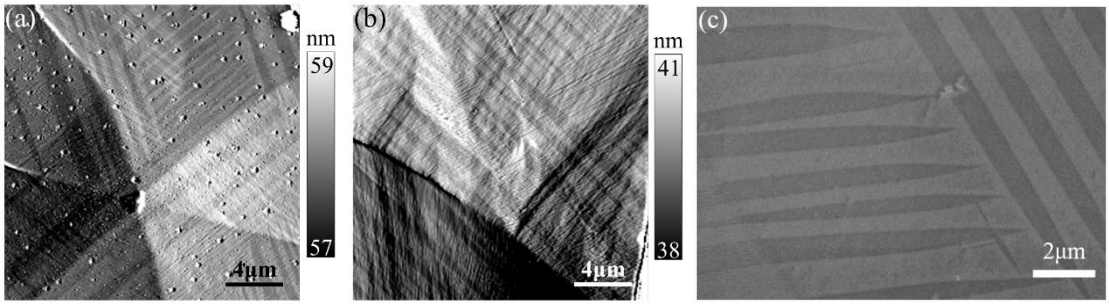

Figure 2.10. Domain patterns in $CH_3NH_3PbI_3$ crystals observed by (ab) SPM topography; (c)SEM.

Striking PFM amplitude mapping acquired through single-frequency scanning is shown in Figure 2.11(a), and the alternating domains exhibit large contrast in piezoresponse up to one order of magnitude difference, hinting their different origins. Tetragonal $CH_3NH_3PbI_3$ has three ferroic variants with equivalent transformation strains [92], regardless of it being polar I4cm or non-polar I4/mcm. Without loss of generality, we consider four possible domain walls between variant 3 and variants (1, 2), which form different angles when these walls intersect different crystalline facets [92]. Some of the expected angles on selected crystalline planes are summarized in Table 2.2, which are indeed observed in our experimental domain patterns (Figure 2.11). The most common angle between domain walls is 70.5°, suggesting that the corresponding crystalline facet is $(101)_t$ in a tetragonal lattice, consistent with $[001]_c$ growth direction observed in XRD [65]. Some of the

earlier studies attributed the PFM domains observed in $CH_3NH_3PbI_3$ to be ferroelectric [77] while others attributed them to be ferroelastic instead [76–78], and such contradictory data exist widely in the literature (Table 2.1), compelling us to examine in details the nature of ferroic domains of $CH_3NH_3PbI_3$.

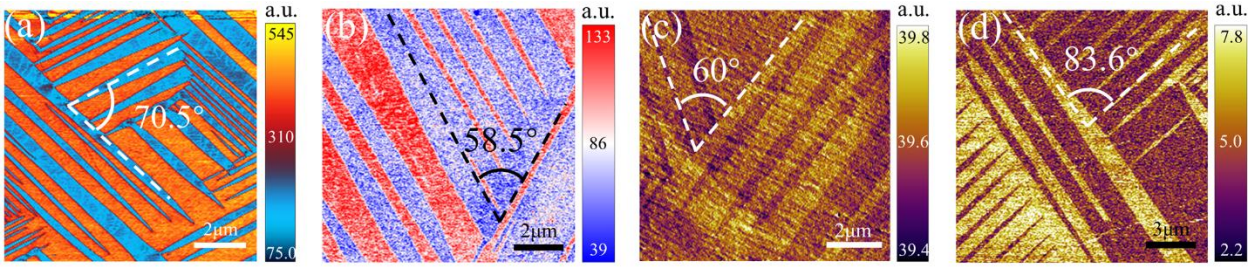

Figure 2.11. Domain patterns of $CH_3NH_3PbI_3$ crystal with angles between domain walls marked. Arbitrary units (a.u.) are used in this figure.

2.3.3 *Multi-harmonic PFM measurement*

The very presence of a sharp piezoresponse contrast in unambiguous ferroic domains suggests two distinct mechanisms in high- and low-response domains, and we argue that one is polar with true piezoelectricity, and the other is nonpolar with apparent yet weaker piezoresponse arising from electrochemical ionic activities [95,96]. The first evidence lies in lateral piezoresponse signal, one of the key differentiators of true piezoelectricity [89].

Characteristic lateral PFM mapping is indeed observed in our crystals (Figure 2.12(a)) obtained via single-frequency scan, and almost identical domain pattern also appears in vertical PFM (Figure 2.12(b)), exhibiting simultaneously high (or low) lateral and vertical piezoresponse. For a crystal in 4mm point group that a polar tetragonal $CH_3NH_3PbI_3$ belongs to, vertical piezoresponse reflects contribution from piezoelectric coefficient d_{33} with out-of-plane polar axis, while lateral piezoresponse from piezoelectric coefficient d_{15} with polar axis in-plane [12,97].

Table 2.2. Transformation strains, domain walls, crystal surfaces, and angles between walls on the crystal surfaces, all expressed in a cubic coordinate system that is parallel to the tetragonal crystalline lattice. Values highlighted in bold are found in our experiments as shown in Figure 2.11.

<table>
<tr><td colspan="2">Variant 3</td><td colspan="2">Variant 1</td><td colspan="2">Variant 2</td></tr>
<tr><td colspan="2">$\begin{bmatrix} \alpha & & \\ & \alpha & \\ & & \beta \end{bmatrix}$</td><td colspan="2">$\begin{bmatrix} \beta & & \\ & \alpha & \\ & & \alpha \end{bmatrix}$</td><td colspan="2">$\begin{bmatrix} \alpha & & \\ & \beta & \\ & & \alpha \end{bmatrix}$</td></tr>
<tr><td colspan="2">Domain walls between variant 3 and variant 1 or 2</td><td>(101)</td><td>$(10\bar{1})$</td><td>(011)</td><td>$(01\bar{1})$</td></tr>
<tr><td rowspan="8">Different crystalline surface</td><td colspan="5">Intersection of domain walls with crystalline surface and the angles between intersecting lines.</td></tr>
<tr><td rowspan="2">(101) surface, Euler angle θ=45^o</td><td>-</td><td>[010]</td><td>$[11\bar{1}]$</td><td>$[\bar{1}11]$</td></tr>
<tr><td colspan="4">54.7^o between domain walls (010) & $(11\bar{1})$ or $(\bar{1}11)$,
70.5^o between domain walls $(11\bar{1})$ & $(\bar{1}11)$</td></tr>
<tr><td rowspan="2">(111) surface, Euler angle θ=54.7^o</td><td>$[10\bar{1}]$</td><td>$[1\bar{2}1]$</td><td>$[01\bar{1}]$</td><td>$[\bar{2}11]$</td></tr>
<tr><td colspan="4">90^o between domain walls $[10\bar{1}]$ & $[1\bar{2}1]$, or $[01\bar{1}]$ & $[\bar{2}11]$,
60^o between domain walls $[10\bar{1}]$ & $[01\bar{1}]$, or $[1\bar{2}1]$ & $[\bar{2}11]$
30^o between domain walls $[10\bar{1}]$ & $[\bar{2}11]$, or $[1\bar{2}1]$ & $[01\bar{1}]$</td></tr>
<tr><td rowspan="2">(112) surface, Euler angle θ=35.3^o</td><td>$[11\bar{1}]$</td><td>$[1\bar{3}1]$</td><td>$[11\bar{1}]$</td><td>$[\bar{3}11]$</td></tr>
<tr><td colspan="4">58.5^o between domain walls $[11\bar{1}]$ & $[1\bar{3}1]$, or $[\bar{3}11]$
63.0^o between domain walls $[1\bar{3}1]$ & $[\bar{3}11]$</td></tr>
<tr><td rowspan="2">(201) surface, Euler angle θ=63.4^o</td><td colspan="2">[010]</td><td>$[\bar{1}\bar{2}2]$</td><td>$[\bar{1}22]$</td></tr>
<tr><td colspan="4">48.2.6^o between domain walls [010] & $[\bar{1}\bar{2}2]$ or $[\bar{1}22]$
83.6^o between domain walls $[\bar{1}\bar{2}2]$ & $[\bar{1}22]$</td></tr>
</table>

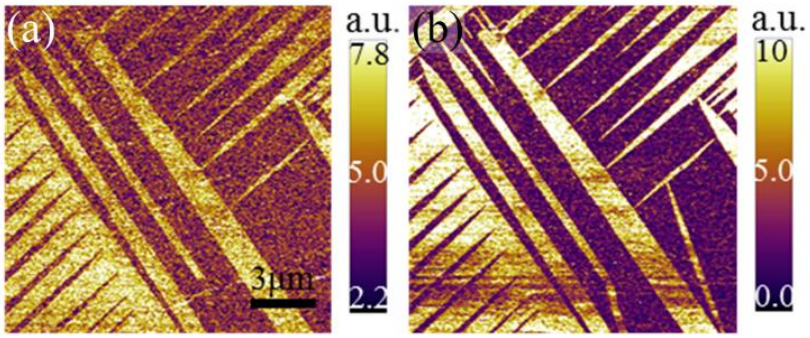

Figure 2.12. PFM amplitude mappings of (a) lateral mode and (b) vertical mode. Arbitrary units (a.u.) are used in this figure.

Thus domains with simultaneously high vertical and lateral piezoresponse can arise from $[001]_t$ polar axis inclined from the crystal surface, and the measured domain wall angle of 83.6° (Figure 2.11) suggests that the inclination angle is 63.4° (Table 2.2). However, domains with simultaneous low vertical and lateral piezoresponse in general is not compatible with 4mm point group symmetry, suggesting that they are actually nonpolar.

To confirm this interpretation, we carry out lateral PFM scan using dual amplitude resonance tracking (DART) corrected via damped harmonic oscillator (DHO) model [48]. Remarkably, DHO works well for lateral PFM in high-response polar domains but completely fails in low-response nonpolar domains, as demonstrated by black dots for all the failing points in the corrected amplitude mapping (Figure 2.13(a)). Mechanisms other than piezoelectricity such as electrostatic interactions and ionic activities are highly unlikely to induce lateral piezoresponse because of their high symmetry [89].

A separate vertical PFM scan (Figure 2.13(b)) provides strong support for alternating polar and nonpolar domains in their drastically different energy dissipation, wherein resonance tracking works well according to frequency mapping (Figure 2.13(c)). It is observed that low-response nonpolar domains exhibit much lower quality factor than high-response polar ones (Figure 2.13(d)),

and thus much higher energy dissipation [48,88,98]. This is a solid proof that low- and high-response domains do have distinct microscopic mechanisms, and the dissipative ionic processes often observed in $CH_3NH_3PbI_3$ [31,99,100] are known to result in apparent (though usually much weaker) vertical piezoresponse.

Further proof of the piezoelectricity in polar domains and nonpiezoelectric mechanisms in nonpolar ones are revealed by their distinct first and second harmonic electromechanical responses, probed as schematically shown in Figure 2.8. It is observed that first harmonic linear response dominates second harmonic quadratic ones in high-response polar domains (Figure 2.13(e)), while the trend is reversed for low-response nonpolar domains, suggesting that the former is piezoelectric and the latter is not [89]. Similar observations have been made throughout these two types of domains and at different AC excitations (Figure 2.13(f)).

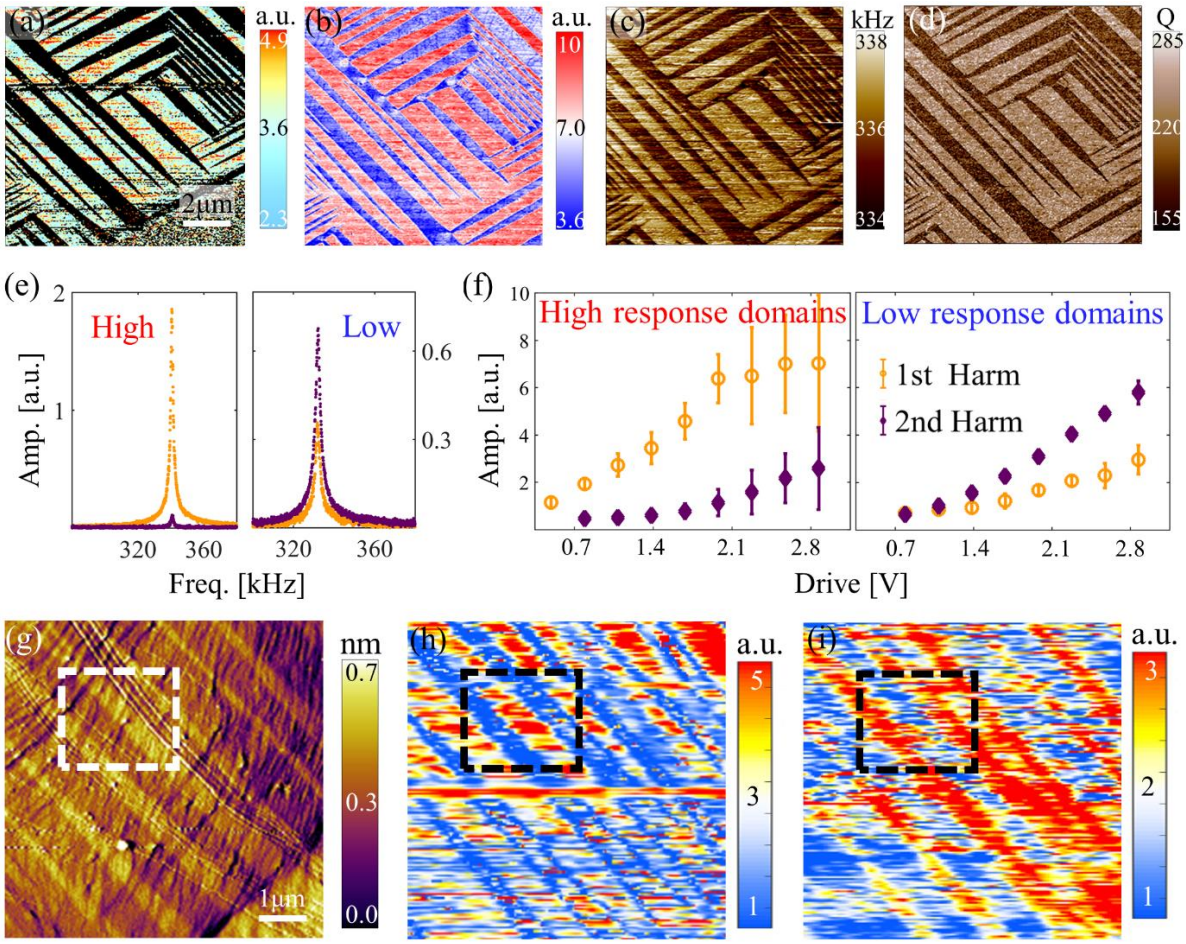

Figure 2.13. Alternating polar and nonpolar domains in $CH_3NH_3PbI_3$ crystal; (a) lateral PFM mapping showing complete failure of DHO (marked by black dots) in nonpolar domains due to their lack of true piezoelectricity; (b) vertical PFM mappings showing identical domain pattern consisting of high-response polar domains and low-response nonpolar domains; (c) resonant frequency mapping from vertical PFM acquired via DART and DHO showing elastic contrast between polar and nonpolar domains and good resonance tracking; (d) quality factor mapping showing substantially lower quality factor and thus higher energy dissipation in nonpolar domains; (e) point-wise tuning of piezoresponse versus frequency showing a point in high-response polar domain has dominant first harmonic response and negligible second harmonic one, while a point in low-response nonpolar domain has higher second harmonic response; (f) comparison of first and second harmonic responses versus AC voltages averaged over points in high- and low-response domains confirming the trend in (e); and (g) SPM topography domain pattern with (h) first and (i) second harmonic piezoresponse mappings confirming opposite contrast of first and second harmonic response.

In addition, the corresponding variations of quality factor in Figure 2.14 reveal lower quality factor and thus higher dissipation in low-response nonpolar domains, as expected. Importantly, for high-response polar domains, quality factors in first harmonic response (originating from true piezoelectricity) are higher than those of second harmonic ones (originating from nonpiezoelectric dissipative processes), indicating two distinct microscopic mechanisms (Figure 2.13(e)). Low-response nonpolar domains, on the other hand, have similar quality factors for both first and second harmonic responses at all voltages, indicating that they both arise from nonpiezoelectric dissipative processes. We have also acquired mappings of first and second harmonic piezoresponse in a different domain pattern (Figure 2.13(g)), revealing that domains with high (low) first harmonic response (Figure 2.13(h)) exhibit low (high) second harmonic response (Figure 2.13(i)), confirming without ambiguity the opposite trends and distinct microscopic mechanisms in these polar and nonpolar domains.

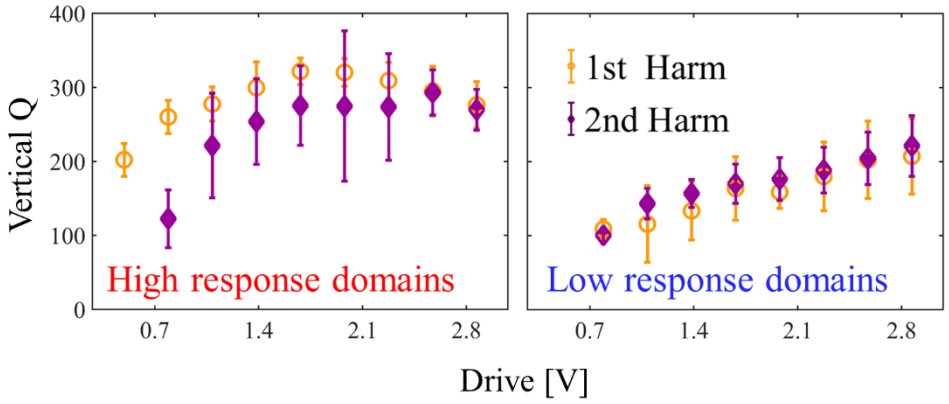

Figure 2.14. Variation of first and second harmonic quality factors versus AC voltage in polar and nonpolar domains.

2.3.4 *Local phase transition*

The alternating polar and nonpolar ferroic domains in $CH_3NH_3PbI_3$ also exhibit opposite temperature evolutions across phase transition. It is found that a thermal probe with a resistance heated by 3.5V can largely erase the SPM topography domains (Figure 2.15(a)), which reappears

after the heating voltage is reduced to 1V to decrease the temperature. This suggests reversible cubic-to-tetragonal phase transition induced by heating and cooling, as recently reported in twins from TEM observation [91]. Furthermore, mappings of the topography (Figure 2.15(b)) and PFM (Figure 2.15(c)) under global heating reveal that contrast in lamellar ferroic domains seen at 30^{o}C and 35^{o}C starts to decrease with increased temperature, and they disappear altogether at nominal temperature of 67^{o}C and 70^{o}C.

Remarkably, domain structures reemerge in both topography and PFM upon cooling, fully recovering at 35^{o}C, and similar phenomena are observed in two cycles of heating and cooling (Figure 2.16), indicating a two-way memory effect in $CH_3NH_3PbI_3$ crystal. Such memory effect is better visualized by the line scan of topography before and after heating (Figure 2.15(d)), wherein clear roof-like topography feature seen at 30^{o}C largely disappears at 67^{o}C, and then remerges at 35^{o}C with identical inclination. Much insight is gained from the evolution of piezoresponse with respect to temperature across phase transition (Figure 2.15(e)), where it is observed that piezoresponse increases upon heating in low-response nonpolar domains yet decreases in high-response polar domains, consistent with expected piezoelectric and ionic mechanisms [101], and they converge above phase transition, since nonpiezoelectric contributions remain in high-response domains beyond phase transition when the crystal becomes completely nonpolar.

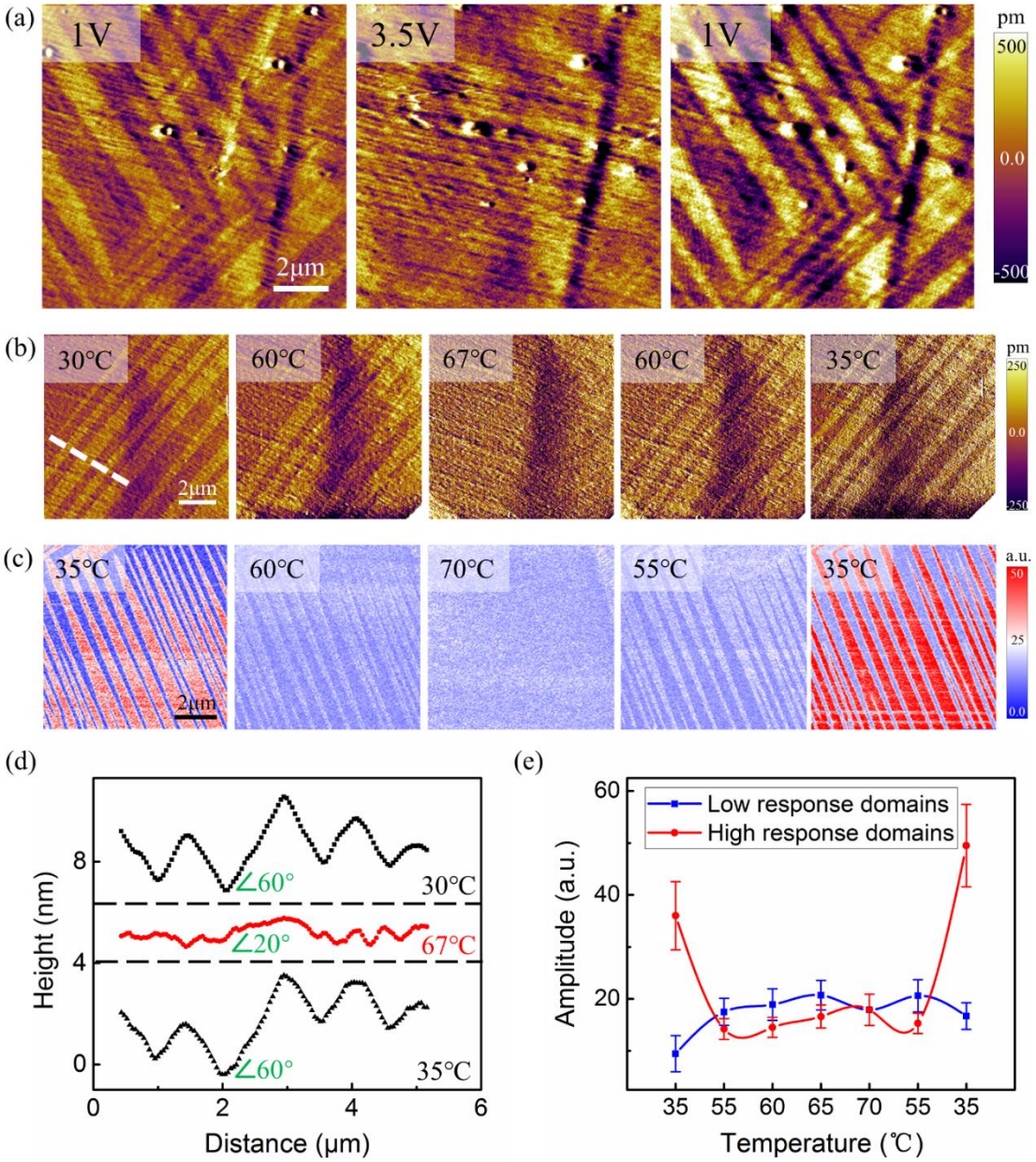

Figure 2.15. Opposite temperature variations in polar and nonpolar domains of $CH_3NH_3PbI_3$ crystal across phase transition; (a) erasing and reappearing of SPM topography domains by heating and cooling using a heated thermal probe; (b) SPM topography and (c) PFM mappings under a sequence of temperature across phase transition showing appearance and reemergence of ferroic domains; (d) topography line scan indicated in (b) before and after heating at 30°C, 67°C, and 35°C showing the disappeared topography feature at high temperature is fully recovered upon cooling; (e) piezoresponse averaged in high-response polar and low-response nonpolar domains showing opposite trend with respect to temperature, yet convergence beyond phase transition.

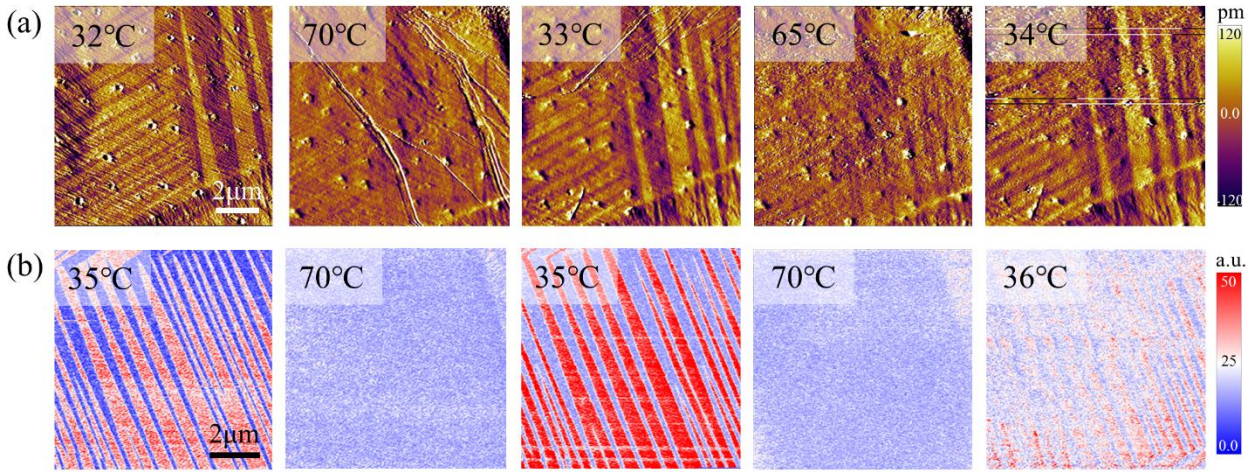

Figure 2.16. Mappings of (a) SPM topography and (b) PFM amplitude under two thermal cycles.

2.3.5 *Ferroic domains regulate photocurrent*

Finally, we show that ferroic domains correlate with photocurrent as revealed by mappings of PFM, pcAFM and SKPM. The PFM amplitude in Figure 2.17(a) shows characteristic lamellar domain pattern, which is followed closely by photocurrent distributions [77] mapped in the same area without a DC bias (Figure 2.17(b)). Furthermore, it is observed that photocurrent in high response polar domains is smaller, suggesting that polar order reduces photocurrent in $CH_3NH_3PbI_3$ on FTO/TiO_2. Such correlation between ferroic domains and electric characteristics is further supported by SKPM acquired in the same area as Figure 2.15(a-d), revealing a negative shift of surface potential in polar domains (Figure 2.17(c)), consistent with reduced photocurrent observed due to electron-collecting nature of FTO/TiO_2 substrate. Similar to topography and PFM amplitude upon heating and cooling across phase transition, memory effect is also observed in photocurrent distribution (Figure 2.17(d-f)). The domain contrast in photocurrent seen at 35^{o}C largely disappears after heating to 70^{o}C, yet reemerges after cooling to 35^{o}C. This further

demonstrates that photocurrent of $CH_3NH_3PbI_3$ crystal on FTO/TiO_2 substrates is regulated by ferroic orders.

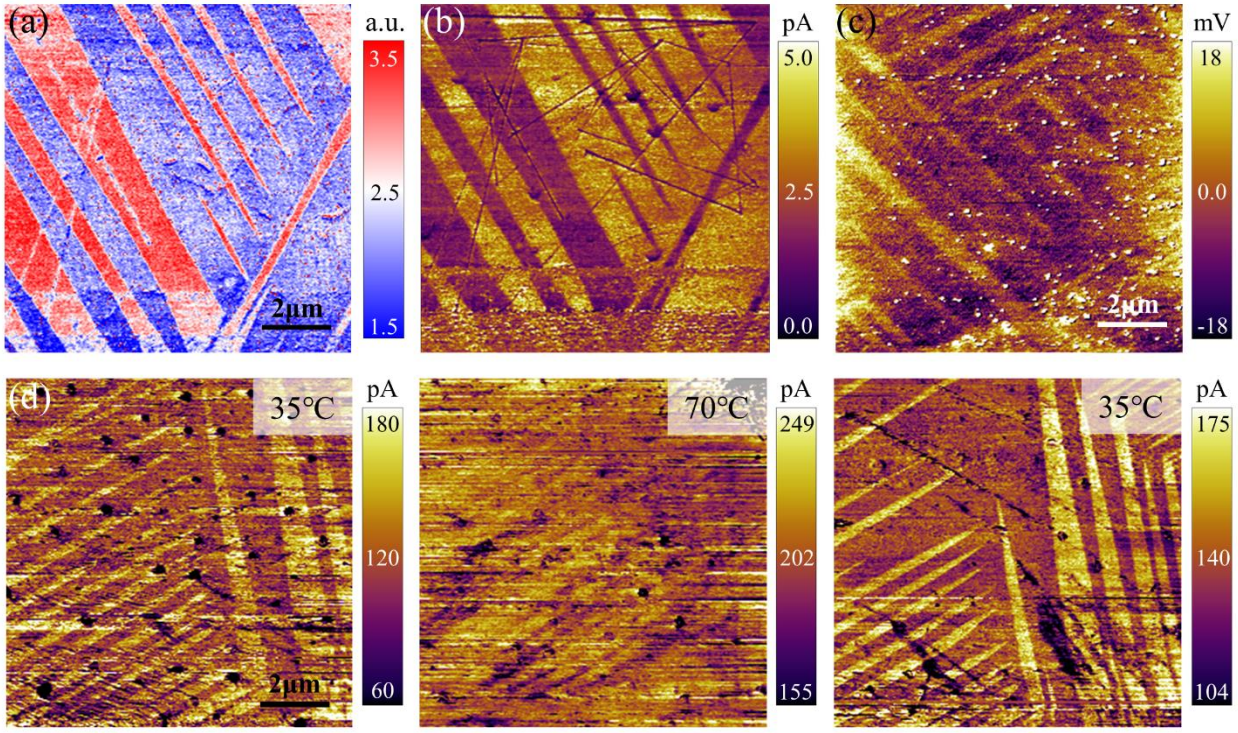

Figure 2.17. Correlation between photocurrent and ferroic domains of $CH_3NH_3PbI_3$ crystal across phase transition; (a) PFM mapping; (b) photocurrent distribution under no DC bias following ferroic domain pattern in (a) with reduced photocurrent in polar domains; (c) surface potential distribution under light follows ferroic domain pattern in Figure 2.13(a-d) with negatively shifted potential in polar domains; and (d) photocurrent distribution in a separate domain pattern at different temperatures across phase transition, showing the disappearing domain pattern at 70°C upon heating and its reemergence at 35°C after cooling.

2.4 SUMMARY AND CHALLENGES

Alternating polar and nonpolar structures in ferroic domains resolves a key puzzle of $CH_3NH_3PbI_3$, that they exhibit strong piezoresponse in polar domains, yet cannot be switched by an electric field applied either locally through a SPM tip or globally through external electrodes [76,78]. We believe this is precisely the consequence of alternating polar and nonpolar domains,

since polar domains that are normally switchable by an electric field are now locked by nonpolar ones that sandwich them. DFT calculations consistently predict that polar structure of tetragonal $CH_3NH_3PbI_3$ is energetically favored at room temperature, though the difference is rather subtle, only in the order of 10s meV [57,80,81,83,85]. Thus, such alternating polar and non-polar domains can be understood from energy landscape predicted from DFT, and such subtle differences on a fine spatial scale are extremely difficult to distinguish. Scanning probes, on the other hand, are powerful in resolving their distinct functional responses with high spatial resolution. Importantly, FTO/TiO_2 substrates are widely used in PSCs, and thus reduced photocurrent observed in polar $CH_3NH_3PbI_3$ on FTO/TiO_2 is significant for devices. Whether such reduction is caused by the intrinsic photovoltaic properties of polar domains or arises from structural effects at interface between $CH_3NH_3PbI_3$ and FTO/TiO_2, for example by band bending due to polarization, requires further investigations. Nevertheless, our studies resolve otherwise indistinguishable polar and nonpolar domains in $CH_3NH_3PbI_3$ and point a direction toward engineering ferroic domains for enhanced PCE.

From the point view of measurements, there are following technical challenges encountered in the systematic study of $CH_3NH_3PbI_3$, most of which are related to data collection and data quality. (1) Strained-based SPM measurements like PFM can be influenced by the crosstalk from topography. Even DART was also likely to fail where significant topography variation locates. Given that only two pairs of datapoints are recorded for each pixel, it is impossible to restore them once affected by unexpected interference. (2) Since environmental temperature is controlled via a feedback loop with a precision of 0.01℃, it may add extra noises to SPM mapping due to thermal expansion. Thus, SPM mapping at higher temperature requires more datapoints for better SNR, otherwise its resolution will be impaired, as seen in Figure 2.16. (3) DART is not applicable for

the second harmonic PFM. Although this measurement is solved via complex setup with a HF2LI lock-in amplifier, the image quality based on resonance tracking still has a long way to go, as shown in Figure 2.13(hi). (4) It is mentioned in Section 2.2.1 that a more efficient way to collect local IV data for a high density grid of points is desired. All of these problems from practice points to the first question that we summarized in Section 1.3, that is, how to develop innovative experimental methodologies based on SPM to acquire high quality, efficient, and physically relevant scientific data to enable deep analysis. We will try figure it out by designing a new scheme for excitation and data acquisition of SPM in next Chapter.

Chapter 3. SEQUENTIAL EXCITATION FOR HIGH QUALITY DATA

3.1 MOTIVATION

While a number of SPM techniques have been applied to probe microscopic interactions of various systems, the rough topography, hierarchical microstructure, and weak response often make an accurate determination of their nanoscale coupling nontrivial, especially in a quantitative manner. For example, we can illustrate this problem by taking a deep look on the data acquisition and analysis of PFM. As noted in Section 1.2.1, the dynamics of the interaction between a tip and sample can be described well by the following damped harmonic oscillator (DHO) model [18,88].

$$A(\omega) = \frac{A_0\omega_0^2}{\sqrt{(\omega_0^2-\omega^2)^2+(\frac{\omega_0\omega}{Q})^2}} \qquad (3.1)$$

$$\phi(\omega) = tan^{-1}[\frac{\omega_0\omega}{Q(\omega_0^2-\omega^2)}] \qquad (3.2)$$

Note that ω is the driving frequency with $A(\omega)$ and $\phi(\omega)$ as the corresponding amplitude and phase that are directly measured in experiment. As such, it is desirable to acquire the amplitude $A(\omega, x, y)$ and phase $\phi(\omega, x, y)$ over a two-dimensional (2D) space (x, y) of sample surface as well as a frequency spectrum of excitation ω, from which intrinsic electromechanical response A_0, resonant frequency ω_0, and quality factor Q can be solved by fitting above equations.

Conventional techniques such as dual amplitude resonance tracking (DART) synthesize AC waveform combining two distinct frequencies $\omega_1$and ω_2 to excite the sample [88,102]. It yields only two pairs of data for each pixel, not amenable for reliable fitting of the highly nonlinear Equations 1.1 and 1.2, and the resonance tracking is not always robust, especially for rough sample surfaces. For example, one couple of $A_1(\omega_1)$ mappings carefully acquired during trace and retrace scanning using DART on granular ceria are shown in Figure 3.1(a), which covers one grain and a

number of grain boundaries. While the trace and retrace mappings resemble each other well in most part of the scans, closer examination reveals that they differ substantially at grain boundaries. This is made clearer by the comparison of line scans shown in Figure 3.1(b), where it is evident that during trace the amplitude is reduced at the grain boundaries while under the retrace, it is enhanced. As such, the trends are completely opposite between trace and retrace at grain boundaries, while away from grain boundaries they agree with each other well. This difference highlights the failure of resonance tracking when there is significant topography variation often encountered in materials, even though the scan is carried out slowly with careful adjusted parameters to ensure the reliable tracking. For more casual scans, the problem will be more serious, as seen in Figure 3.1(c), where there are noticeable wavelet scratch-like pattern in the mapping of frequency ω_1 that extends along the scanning direction of the probe, especially at grain boundaries. Such artifacts clearly indicate the failure of resonance tracking, even within a grain where topography variation is insignificant, and such mappings are not reliable even for qualitative analysis.

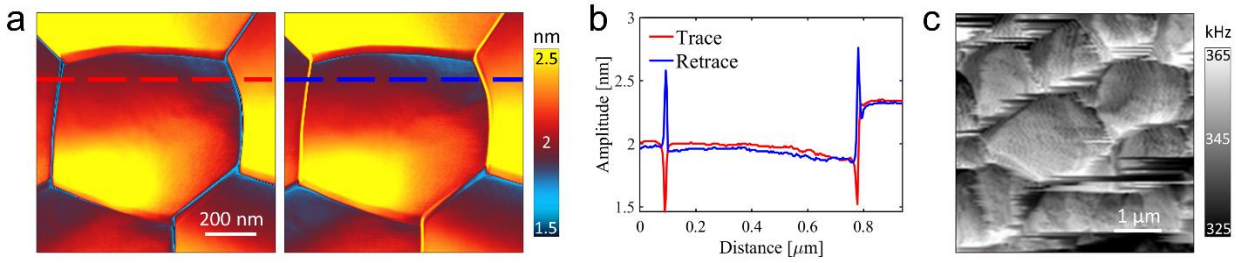

Figure 3.1. The difficulty and failure of resonance tracking; (a) amplitude mappings from trace and retrace acquired under DART; (b) comparison of corresponding line scans between trace and retrace; (c) artifacts in the mapping of excitation frequency.

Band excitation (BE) has also been developed for this problem by synthesizing a signal summing all harmonic excitation within a frequency band [103,104], so that resonance is covered

within the band and thus tracking becomes unnecessary. Nevertheless, the excitation power of BE is distributed among the band of frequencies, resulting in much reduced strength and signal-to-noise ratio (S/N) [103,104]. Newly developed general mode (G-mode) SPM records complete time- instead of frequency-domain data that is very powerful [42,105], though a substantial portion of data could be redundant and thus the subsequent analysis is not very efficient, and it requires sophisticated instrumentation that is not easily accessible. Therefore, there is still strong desire for innovative yet simple and easily accessible approach to acquire high quality (less noisy), efficient (less redundant), and physically relevant SPM data to enable deep analysis.

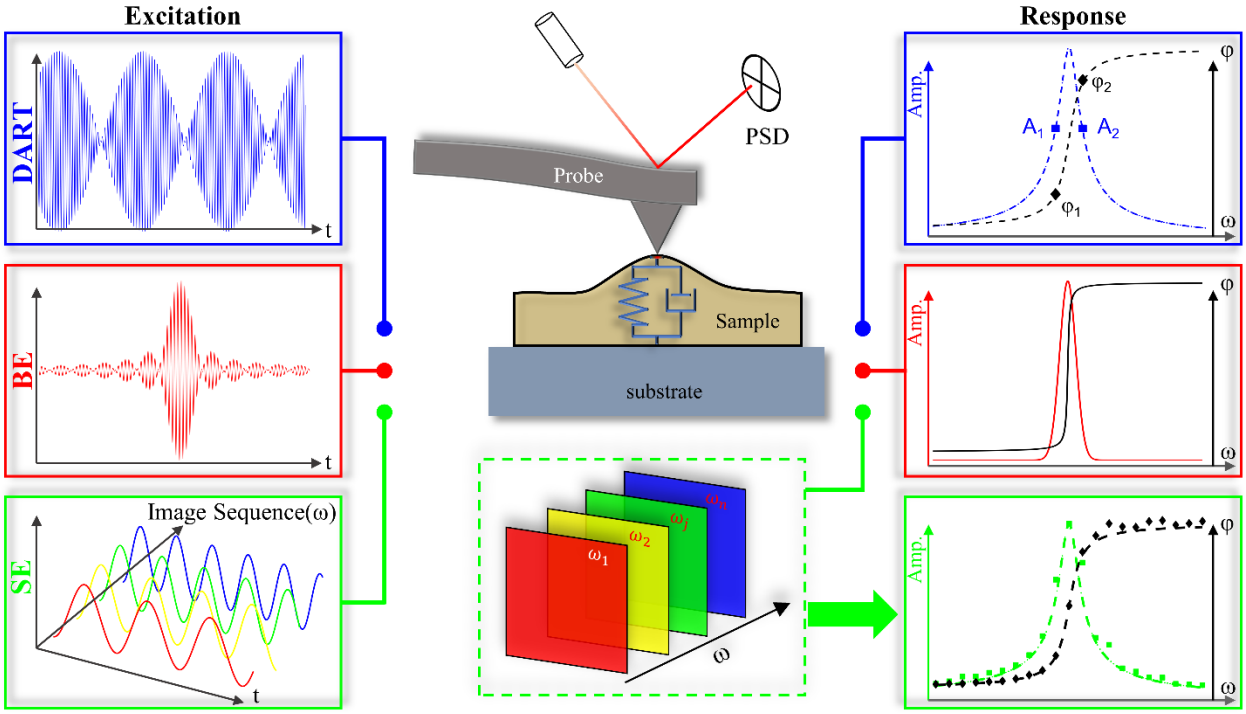

Figure 3.2. The schematics of dynamic SPM experiments based on DART, BE, and SE techniques, wherein AC waveform combining two distinct or a band of frequencies are synthesized to excite the sample under DART or BE, respectively, while a sequence of AC waveforms with different frequencies are used to excite the sample under SE.

3.2 SEQUENTIAL EXCITATION: PRINCIPLE AND IMPLEMENTATION

To this end, we develop a sequential excitation (SE) technique to acquire high quality, efficient, and physically relevant data in frequency domain [106], and the method can be easily implemented on any standard SPM without the need for any additional hardware and instrumentation. The principle of this method is to excite the sample using a sequence of AC waveforms with distinct frequencies ω_j, as shown in Figure 3.2, wherein the excitation energy is concentrated on only one frequency at a time instead of being distributed over a band of spectrum, ensuring that the signal is strong and the response is not noisy. In such a setup, each excitation frequency captures cantilever-sample resonance at selected spatial points that are unique, ensuring that the data is relevant yet not redundant. Furthermore, no resonance tracking is needed as in DART, ensuring that the measurement is robust and reliable. In a sense, such a strategy of SE is analogous to super-resolution microscopy in biology that turns specific fluorescent molecules on and off in a sequential manner for imaging [107,108], wherein we excite specific resonances of different points sequentially using distinct frequencies.

Consequently, three-dimensional (3D) data sets of amplitude $A(\omega, x, y)$ and phase $\phi(\omega, x, y)$ can be obtained from SE. Taking possible drift into account, we will take the topography mapping of the first scan as a benchmark to do the image registration for other scans. The obtained translation is applied to the amplitude and phase mappings as well to correct drift using an open-source algorithm [109]. The fit function with the nonlinear least-squares method in MATLAB is then used for DHO fitting, which returns the mappings of A_0, ω_0, Q, and R^2. R^2 is a statistical measure of how close the data are to the fitted curve. It ranges from 0 to 1. $R^2 = 1$ indicates that the fitted data f_i explains all variability in observed data y_i, while $R^2 = 0$ indicates no 'linear' relationship between them. The most general definition of R^2 is $R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$, where $SS_{tot} =$

$\sum_{i=1}^{n}(y_i - \bar{y})^2$, $SS_{res} = \sum_{i=1}^{n}(y_i - f_i)^2$. The workflow of SE is shown in Figure 3.3, while the software toolkit designed for the postprocessing of SE dataset is shown in Figure 3.4.

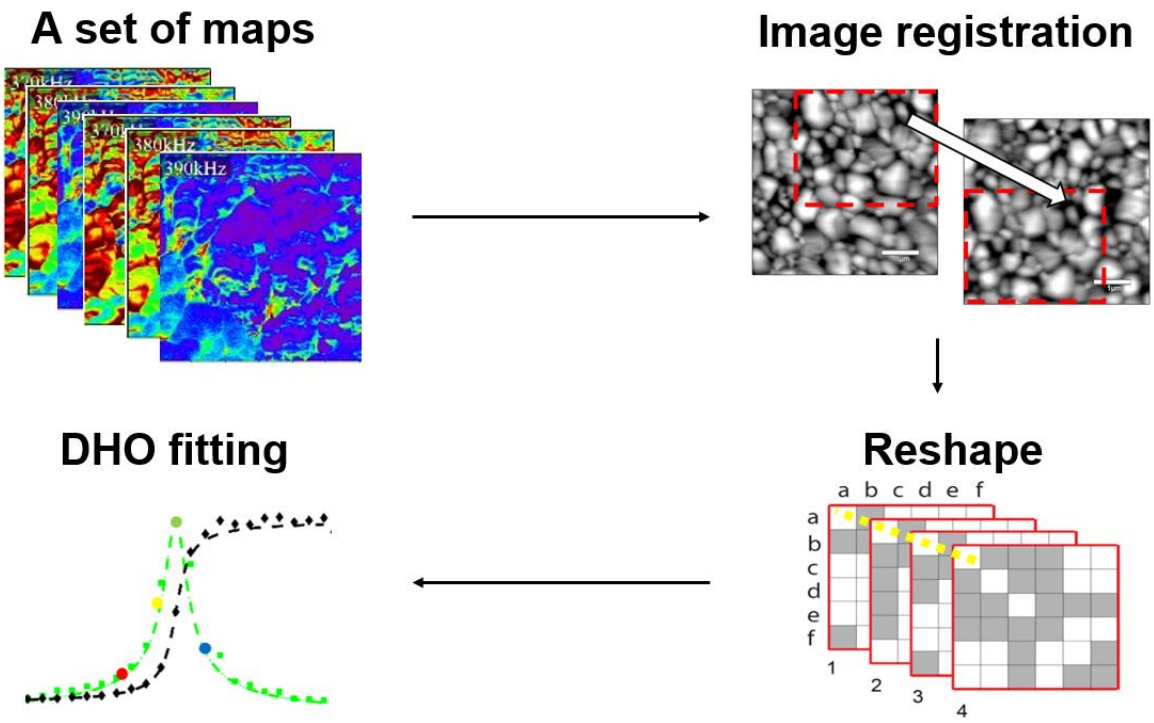

Figure 3.3. The workflow of sequential excitation.

Besides, principal component analysis (PCA) as a powerful multivariate statistical tool was also carried out [110], which is known to decompose the original dataset into orthogonal components, arrange them in decreasing order of statistical significance, and highlight and separate exotic behavior with little computation effort. Although PCA was introduced for compressing data and enhancing SNR in the beginning, of particular interest is that a clear mechanistic understanding of the unsupervised PCA acquired through sound physical principle was found, which makes it possible to speed up physical analysis by at least four orders of magnitude. In short, SE requires

no extra hardware and further instrumentation in a standard SPM, and it can be easily implemented, making it widely accessible.

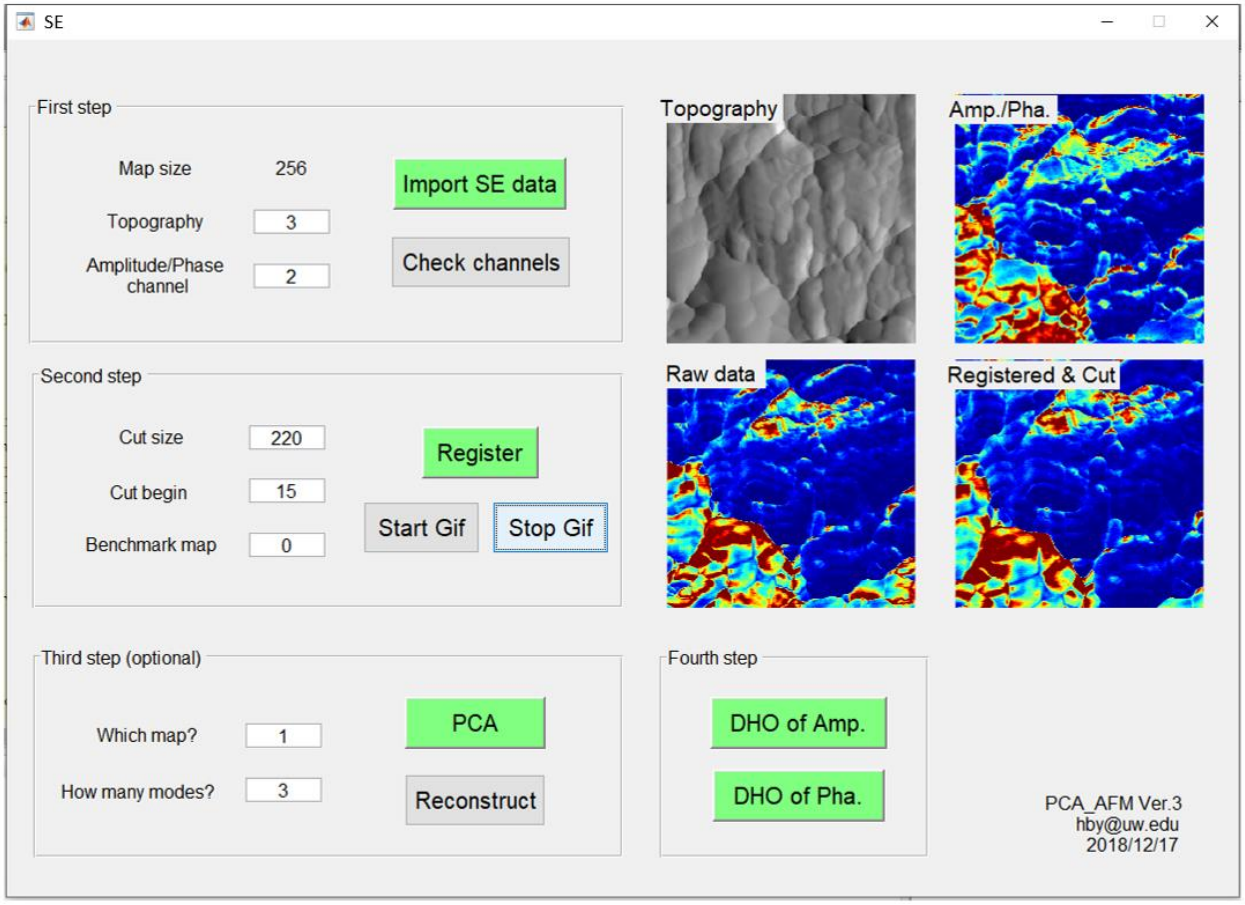

Figure 3.4. The software toolkit designed for the postprocessing of SE dataset.

3.3 SE-PFM FOR A ROUGH PZT CERAMIC

3.3.1 *Experiment*

As a demonstration, we conduct PFM on a PZT ceramic under SE, wherein a series of single frequency PFM mappings were acquired under an AC voltage of 3V and excitation frequency ranging from 320 to 400 kHz with a 2kHz increment. The scan rate was set as 1.5Hz. Therefore, a total of 41 PFM mappings were obtained by using the programmable function Macrobuilder of AR software. All raw amplitude data were then imported into MATLAB as 41 256×256 matrices.

Subpixel image registration was performed for all images by cross-correlation to correct for inevitable topography drifts during the experiment.

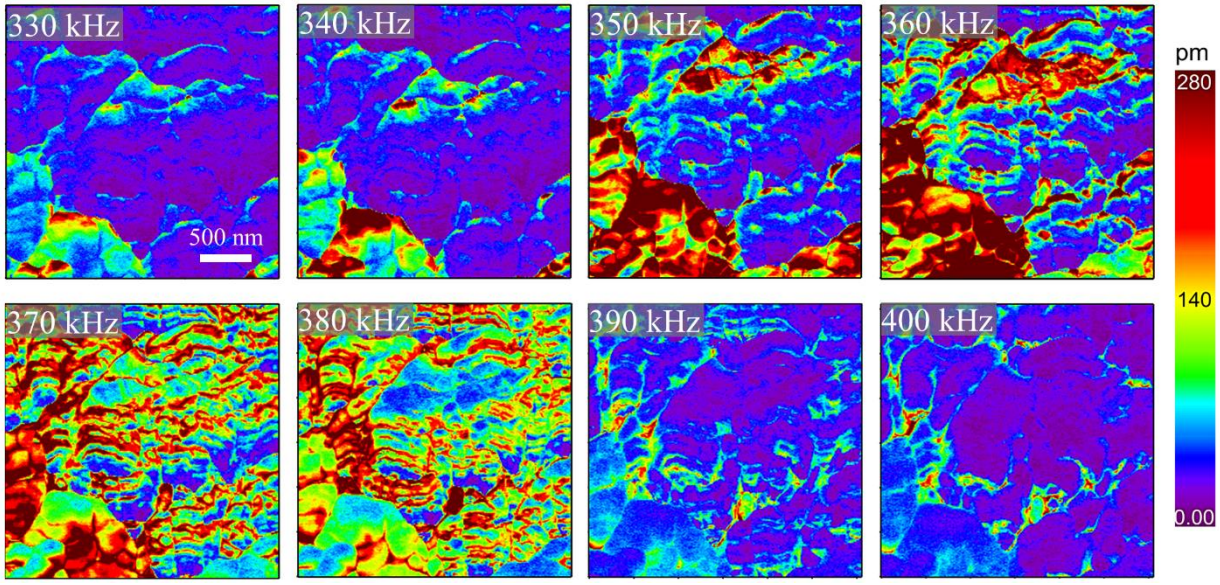

Figure 3.5. A sequence of SE-PFM amplitude mappings obtained at selected frequencies in PZT ceramic.

A sequence of its amplitude mappings $A(\omega_j, x, y)$ are shown in Figure 3.5. It is observed that PFM amplitude is very sensitive to the excitation frequency ω_j, as expected, and there are substantial amplitude changes when the excitation frequency varies. In addition, substantial spatial heterogeneity is observed within each mapping, reflecting possible variations in intrinsic piezoelectricity, elasticity, energy dissipation, or their combinations. Such crosstalk makes it difficult to determine intrinsic SPM response quantitatively, and it is necessary to deconvolute these different effects. In fact, our work was originally motivated by this very issue, which has important implications in nanoscale probing of electromechanical coupling that underpins functionalities of both synthetic materials and biology [12,111–113]. SE-PFM makes it possible to overcome such difficulties, though we must reconstruct data to determine the intrinsic response,

analogous to super-resolution microscopy. In this regard, the three-dimensional (3D) data sets of amplitude $A(\omega, x, y)$ and phase $\phi(\omega, x, y)$ obtained from SE are amenable to both physics-based DHO analysis and statistics-based PCA, making deep data analysis possible [44].

3.3.2 *Brute force analysis via DHO fitting*

We start with brute force physical analysis accomplished by fitting 3D datasets of amplitude $A(\omega, x, y)$ at each pixel (x, y) using DHO model. This is demonstrated by one representative pixel in Figure 3.6(a), yielding a resonant frequency of 347.8 kHz, quality factor of 33.23, and intrinsic electromechanical response of 15.25 pm at that particular point. Note that the sample surface is rather rough as revealed by its topography (Figure 3.6(b)), which imposes substantial difficulty for DART-PFM, yet DHO analysis can be easily applied to each pixel of SE-PFM to reconstruct the mappings of intrinsic amplitude, resonant frequency, and quality factor, as shown in Figure 3.6(c-e).

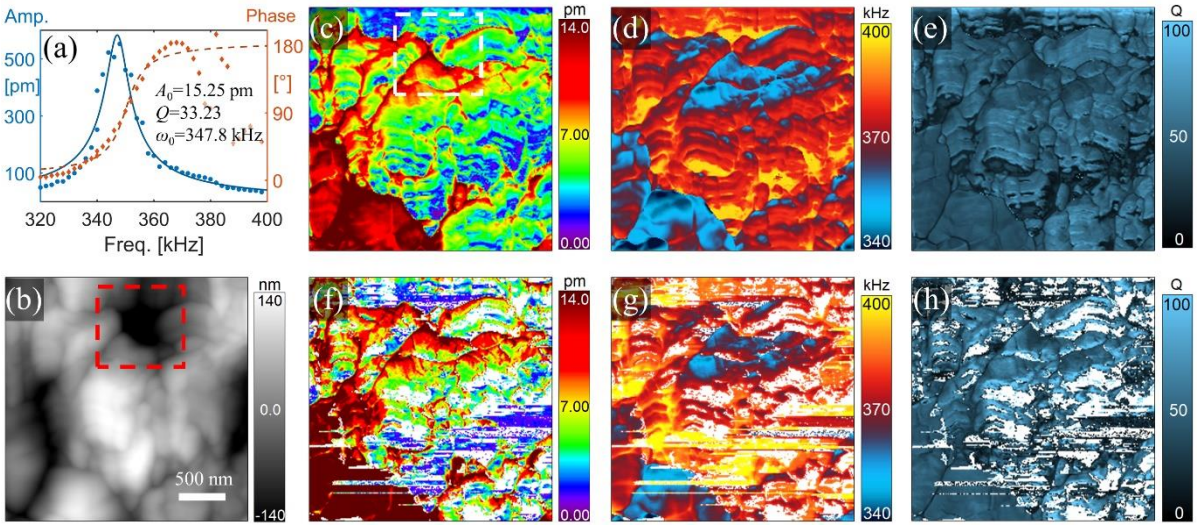

Figure 3.6. Comparison of PZT mappings acquired by SE-PFM and DART-PFM processed via DHO; (a) DHO fitting of SE-PFM spectrum data for one representative pixel; (b) rough topography mapping; and (c-e) reconstructed SE-PFM mappings of (c) intrinsic amplitude A_0, (d) resonance frequency ω_0, and (e) quality factor Q; (f-h) reconstructed DART-PFM mappings of (f) the intrinsic amplitude, (g) resonance frequency, and (h) quality factor obtained, wherein white areas show point where DHO analysis fails.

The mapping of R^2, a statistical measure known as the fitting coefficient of determination assessing how close the data are to the fitted regression line, is presented in Figure 3.7, revealing values ranging from 0.85 to 0.99 and thus a high fidelity of DHO analysis. This demonstrates the capability of SE-PFM even for highly inhomogeneous and rough samples. Such capability, however, is beyond the conventional DART-PFM, as exhibited in Figure 3.6(f-h), where it is observed that a significant percentage of points (27%) fails to yield a valid solution in DHO analysis, as highlighted by white pixels in the mappings. Such a problem also casts doubts on the points wherein DHO analysis works.

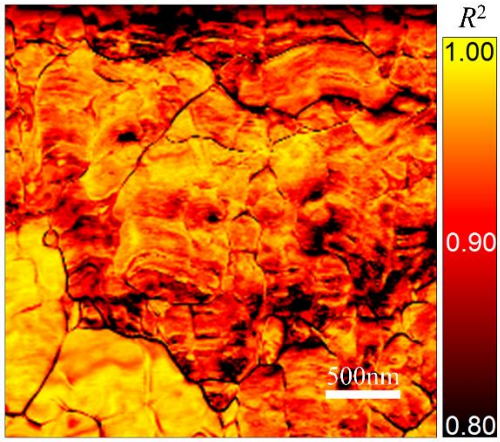

Figure 3.7. Mapping of R^2.

This highlights the advantage of SE over DART, which measures responses at only two excitation frequencies across resonance, and uses the difference between these two responses as feedback for resonance tracking. Such strategy often runs into difficulties: if the separation between two excitation frequencies is too small, they will easily fall out of resonance range during scanning and thus fail to track resonance shift; and if the separation is too large, then the responses are weak and the signal-to-noise ratios are low. For materials exhibiting substantial heterogeneity at the nanoscale, for instance near the grain boundaries [52,114,115] wherein the contact resonance frequency can shift significantly over a relatively short distance, resonance tracking and thus DHO analysis often fail [116]. This is clearly demonstrated in PZT (Figure 3.6(f-h)) that has a strong piezoelectric response yet a rough topography (Figure 3.6(b)), which is not uncommon in practice, and the excitation frequency must shift substantially during scanning (Figure 3.6(g)). An excellent ferroelectric material such as PZT still suffers from such difficulty, and the issue is only more serious for other materials with weaker electromechanical coupling. Mappings acquired from SE-PFM, on the other hand, are free of such problems.

3.3.3 *Principal component analysis*

While DHO fitting works well under SE-PFM, it is a computationally expensive process, taking an Intel Xeon E5-2695 CPU approximately 0.06 s for one pixel and thus 1.09hr for a 256×256 mapping (or 5.8 min for a parallel pool with 28 CPU workers). We thus resort to multivariate statistical tools such as PCA to speed up the analysis [110], as a sequence of images has been obtained under different excitation frequencies, ideal for PCA. Through orthogonal transformation, PCA converts a set of possibly correlated variables, in this case SE-PFM mappings under different excitation frequencies, to a set of linearly uncorrelated variables known as principal components. As a powerful unsupervised data analytic tool, PCA is widely used to compress and visualize a multi-dimensional dataset, though its physical meaning is often unclear. By carefully comparing underlying mathematics of both approaches, we are also able to draw a close connection between PCA modes and DHO expansion.

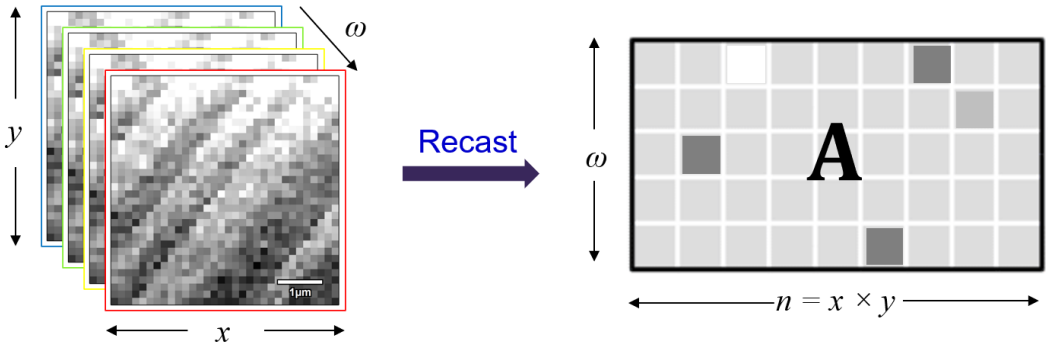

Figure 3.8. Recast of the dataset.

To this end, we first recast the 3D dataset of $\mathbf{A}(\omega, x, y)$ into a 2D matrix of $\mathbf{A}(\omega, n)$, as shown in Figure 3.8, where 2D spatial grid is collapsed into 1D. Here each row of $\mathbf{A}$ contains spatial data from a mapping scanned at a particular frequency, while each column represents spectral responses

of a particular grid point acquired under various excitation. PCA of $\mathbf{A}(\omega, n)$ can then be carried out through singular value decomposition (SVD) [117], as shown in Figure 3.9,

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^{\mathrm{T}} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u_i} \boldsymbol{w_i^T} \qquad (3.3)$$

where $\{\boldsymbol{u_i}\}$ and $\{\boldsymbol{w_i}\}$are left and right singular vectors of **A**, corresponding to the principal spectral and spatial modes sorted by their singular values $\{\sigma_i\}$. In other words, any row/column of **A** can be represented with a combination of $\{\boldsymbol{w_i}\}$ or $\{\boldsymbol{u_i}\}$, separately.

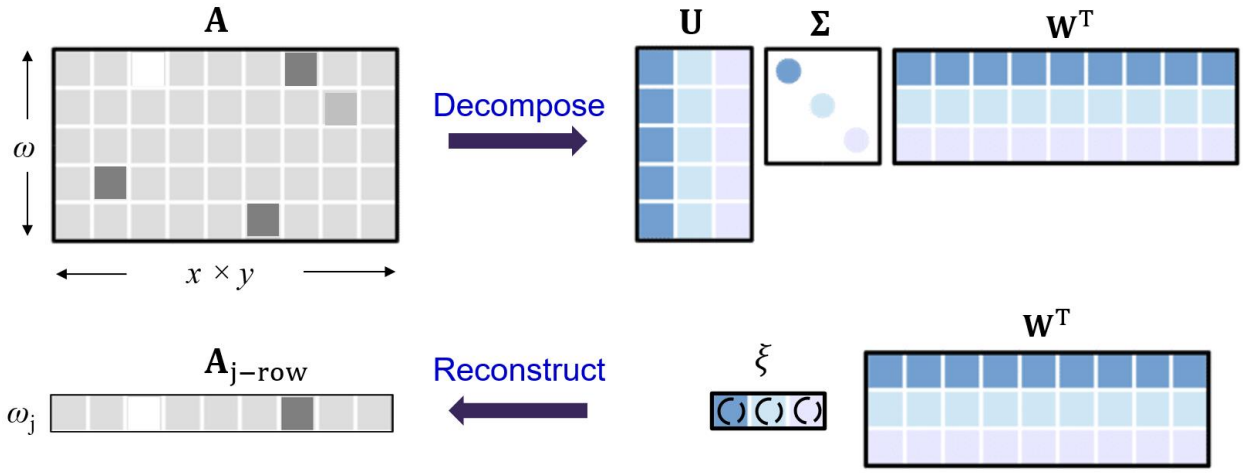

Figure 3.9. Schematics of SVD and reconstruction of the dataset.

Therefore, the j^{th} row of **A**, that is $\mathbf{A}_{j\text{-}row}$, can be regarded as a linear combination of $\boldsymbol{w_i}$ with coefficients ξ_i,

$$\mathbf{A}_{j-row} = \sum_i \xi_i(\omega_j) \cdot \boldsymbol{w_i} \qquad (3.4).$$

Intuitively, $\{\boldsymbol{w_i}\}$ are a set of orthogonal basis spanning the high dimensional space where SE data exist. For DHO model, the intrinsic property A_0, Q, ω_0 also span the space, though they are a set of nonlinear as well as correlated basis. If we can reconstruct a new series of basis like $\{\boldsymbol{\beta_i}\}$ from DHO basis and satisfy the properties of PCs, as shown in Figure 3.10, we may approximate their physical meaning.

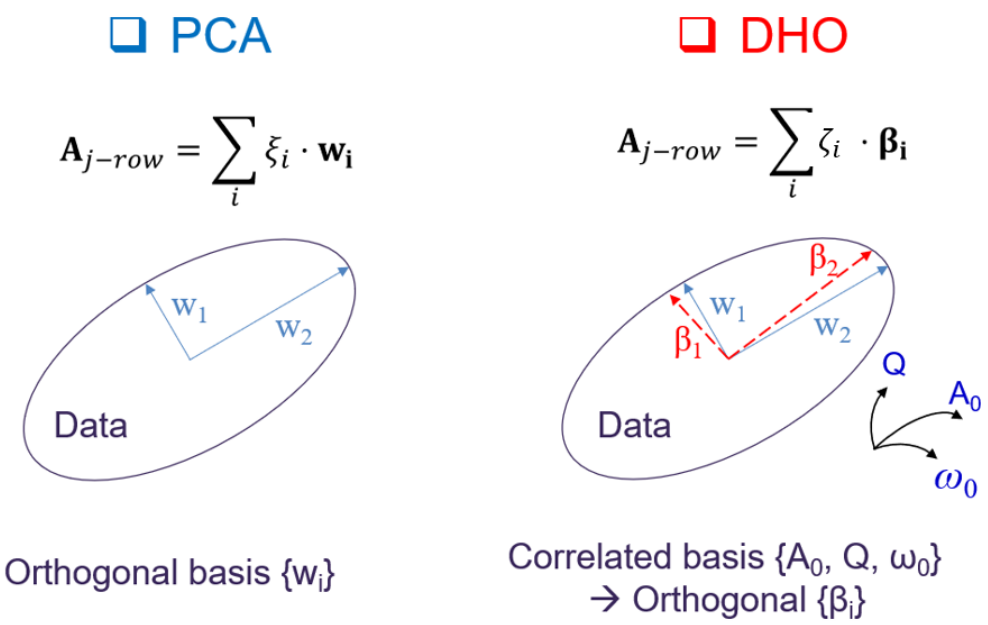

Figure 3.10. A schematic comparison of basis derived from PCA and DHO.

To this end, $\mathbf{A}_{j\text{-}row}$ can also be reformulated as follows,

$$\mathbf{A}_{j-row} = \sum_i \lambda_i(\omega_j) \cdot \boldsymbol{\alpha_i} = \sum_i \zeta_i(\omega_j) \cdot \boldsymbol{\beta_i} \qquad (3.5).$$

And we can first do Taylor expansion series of DHO model. For any given pixel scanned at a specific frequency ω_j, $A(\omega_j)$ can be reformulated as follows by plugging Equation (1.2) into (1.1),

$$\begin{aligned} A(\omega_j) &= \frac{A_0\omega_0^2}{\sqrt{\left[\frac{1}{\tan\phi(\omega_j, Q, \omega_0)}\right]^2 + 1}} \frac{Q}{\omega_0\omega_j} \\ &= \sin\phi(\omega_j, Q, \omega_0) \cdot A_0\omega_0^2 \frac{Q}{\omega_0\omega_j} \\ &= \eta(\omega_j, Q, \omega_0) \cdot A_0 Q\omega_0 \end{aligned}$$

where $\eta(\omega_j, Q, \omega_0) = sin\,\phi\,(\omega_j, Q, \omega_0)/\omega_j$. Then it can be expanded into 2D Taylor series around the spatial average $(\bar{Q}, \bar{\omega}_0)$ of the whole scanned area as,

$$
\begin{aligned}
A(\omega_j) &= A_0 Q \omega_0 \cdot \eta(\omega_j, Q, \omega_0) \\
&= A_0 Q \omega_0 \cdot [\eta_0(\omega_j, \bar{Q}, \bar{\omega}_0) + \left.\frac{\partial \eta}{\partial Q}\right|_{(\omega_j, \bar{Q}, \bar{\omega}_0)} \cdot \Delta Q + \left.\frac{\partial \eta}{\partial \omega_0}\right|_{(\omega_j, \bar{Q}, \bar{\omega}_0)} \cdot \Delta\omega_0 + \left.\frac{\partial^2 \eta}{\partial^2 \omega_0}\right|_{(\omega_j, \bar{Q}, \bar{\omega}_0)} \cdot (\Delta\omega_0)^2 + \ldots] \\
&= A_0 Q \omega_0 \cdot [\lambda_1(\omega_j) + \lambda_2(\omega_j) \cdot (Q - \bar{Q}) + \lambda_3(\omega_j) \cdot (\omega_0 - \bar{\omega}_0) + \lambda_4(\omega_j) \cdot (\omega_0 - \bar{\omega}_0)^2 + \ldots]
\end{aligned}
$$

where$\lambda_i(\omega_j)$ are coefficients that only depends on ω_j since $(\bar{Q}, \bar{\omega}_0)$ are constant. This expansion can be further generalized to the whole map as,

$$
\begin{aligned}
\mathbf{A}_{j-row} &= \mathbf{A_0 Q \omega_0} \circ [\lambda_1(\omega_j) + \lambda_2(\omega_j) \cdot (\mathbf{Q} - \bar{\mathbf{Q}}) + \lambda_3(\omega_j) \cdot (\boldsymbol{\omega_0} - \bar{\boldsymbol{\omega}}_0) + \lambda_4(\omega_j) \cdot (\boldsymbol{\omega_0} - \bar{\boldsymbol{\omega}}_0)^2 + \ldots] \\
&= \sum_i \lambda_i(\omega_j) \cdot \boldsymbol{\alpha_i}
\end{aligned}
$$

where $\{\boldsymbol{\alpha_i}\} = \mathbf{A_0 Q \omega_0} \circ [\mathbf{1}, \mathbf{Q} - \bar{\mathbf{Q}}, \boldsymbol{\omega_0} - \bar{\boldsymbol{\omega}}_0, (\boldsymbol{\omega_0} - \bar{\boldsymbol{\omega}}_0)^2, \ldots]$ and$\boldsymbol{A_0 Q \omega_0} = \boldsymbol{A_0} \circ \boldsymbol{Q} \circ \boldsymbol{\omega_0}$ is the Hadamard product of 1D vectors $\boldsymbol{A_0}$, $\boldsymbol{Q}$, and $\boldsymbol{\omega_0}$ that are reshaped from intrinsic parameter mappings $A_0(x, y)$, $Q(x, y)$, and $\omega_0(x, y)$, respectively. $\mathbf{1}$ is a 1D vector, all elements of which are 1, while $\bar{\boldsymbol{Q}} = \mathbf{1} \cdot \bar{Q}$ and $\bar{\boldsymbol{\omega}}_0 = \mathbf{1} \cdot \bar{\omega}_0$. Note that $\boldsymbol{\alpha}_1$ corresponds to $\boldsymbol{A_0 Q \omega_0}$ because it is always leading term in the 2D Taylor series, while the sequence of following $\boldsymbol{\alpha}_i$ depends on the relative variation of $\boldsymbol{Q}$ and $\boldsymbol{\omega_0}$, and the order could change for different physical systems. Since $\boldsymbol{w}_i$ is orthogonal while $\boldsymbol{\alpha_i}$ is not, to match Equation (3.4), we transform $\{\boldsymbol{\alpha_i}\}$ into a set of orthonormal basis $\{\boldsymbol{\beta_i}\}$ via Gram–Schmidt process [118], with $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1 = \boldsymbol{A_0 Q \omega_0}$, $\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - \frac{\boldsymbol{\alpha}_2 \cdot \boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|^2} \boldsymbol{\beta}_1$, and $\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_3 - \frac{\boldsymbol{\alpha}_3 \cdot \boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|^2} \boldsymbol{\beta}_1 - \frac{\boldsymbol{\alpha}_3 \cdot \boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|^2} \boldsymbol{\beta}_2$, after which $\{\boldsymbol{\beta_i}\}$ is normalized. The analogy between Equation (3.5) and Equation (3.4) is evident, suggesting that PCA components $\{\boldsymbol{w}_i\}$ are close to orthonormal basis $\{\boldsymbol{\beta_i}\}$ derived from DHO. In a completely parallel manner, the correspondence between PCA spectral eigenvectors and DHO expansion can be established by switching the row and column of $\boldsymbol{A}$, *i.e.* between PCA spectral eigenvectors $\boldsymbol{A w}_i$ and DHO spectral basis $\boldsymbol{A \beta}_i$. In particular, elements of $\boldsymbol{A w}_i$ represent the weight $\boldsymbol{\xi}_i$ that $\boldsymbol{w}_i$ takes up in each scan according to Equation (3.4). Therefore, it turns out that there is one-to-one correspondence between PCA modes and DHO

expansion basis, which is confirmed by the good agreement are shown in Figure 3.11. The structural similarity (SSIM) (99.7%, 98.5%, and 95.5%), and Pearson correlation coefficients (PCC) (90.06%, 91.15%, and 72.97%) between PCA spatial modes and DHO basis in Figure 3.11(bc) are pretty high, validating our analysis numerically.

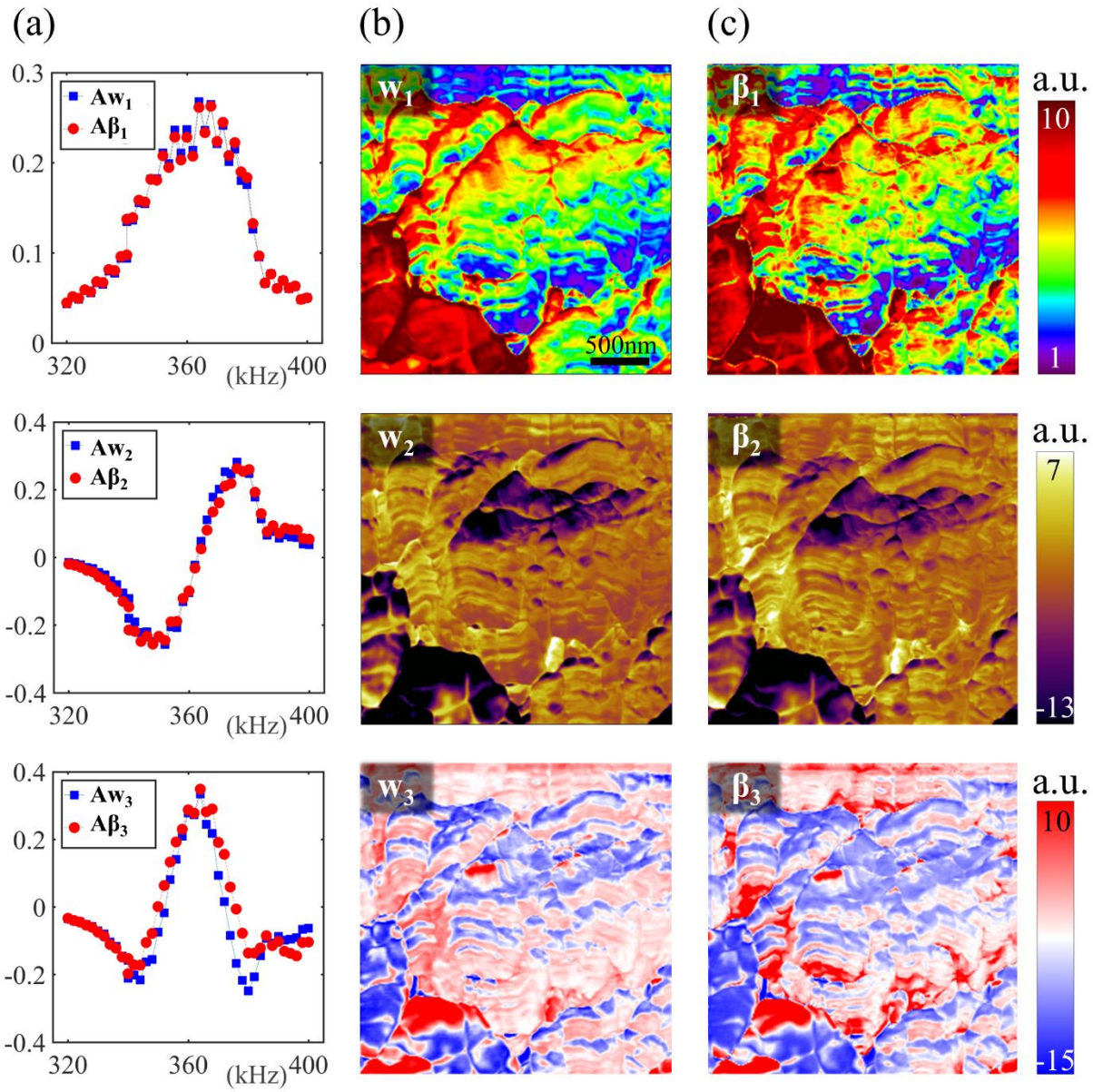

Figure 3.11. Comparison of PCA and DHO expansion for SE-PFM data of PZT; (a) first three PCA spectral eigenvectors in comparison with corresponding DHO spectral basis; (b) first three PCA spatial eigenvectors; (c) corresponding DHO spatial basis.

Here SSIM is an index for measuring the similarity between two images, which is defined as the mean of the local SSIM value map SSIM(x, y):

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_a\mu_b+c_1)(2\sigma_{ab}+c_2)}{(\mu_a^2+\mu_b^2+c_1)(\sigma_a^2+\sigma_b^2+c_2)},$$

where $\mu_a, \mu_b, \sigma_a, \sigma_b, \sigma_{ab}$ are the average, variance, and covariance of 4×4 windows a and b that are centered in the pixel (x, y) of two images, and c_1 , c_2 are two variables to stabilize the division. PCC, on the other hand, is a measurement of the linear correlation between two vectors X and Y, which are reshaped from two images, respectively. It ranges from +1 and −1, where 1 means total positive linear correlation, 0 means no linear correlation, and −1 means total negative linear correlation,

$$P_{X,Y} = \frac{\sigma_{XY}}{\varsigma_X \varsigma_Y}$$

where $\sigma_{XY}$is the covariance and $\varsigma_X, \varsigma_Y$are the standard deviation of X and Y.

Intuitively, the set of SE-PFM mappings under different frequencies contains two important information, the variation of the amplitude with respect to the spatial locations and with respect to excitation frequencies, which are interconnected in the original mappings of Figure 3.5. Under PCA, the data are transformed, such that the spatial variation is best represented by the spatial eigenvectors, and the frequency variation is reflected in the spectral eigenvectors for each PCA mode. Note that the principal components are sorted by their eigenvalues in a descending manner, with the first principal component accounting for the maximum possible variability in the data, as shown by the scree plot of variance in Figure 3.12. Note that PCA takes only 0.24 s for an Intel Xeon E5-2695 CPU to complete, which is four orders of magnitude faster than brute force fitting.

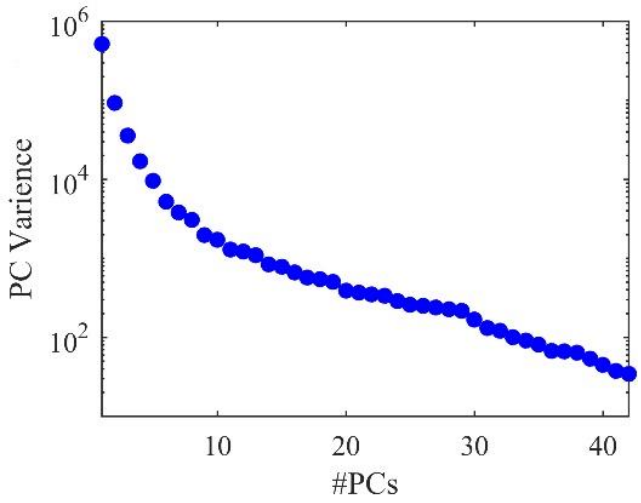

Figure 3.12. Scree plot of variance of principal components for SE-PFM results.

3.3.4 *Numerical proof*

The spatial variation of intrinsic amplitude, resonant frequency, and quality factor, key material parameters of interests in above analysis, are not known in advance. Thus, in order to unambiguously establish our physical interpretation of PCA, we construct a model three-phase system numerically with pre-determined distribution of intrinsic amplitude, resonant frequency, and quality factor for each phase, as shown in Figure 3.13(a), from which corresponding SE-PFM mappings can be computed using DHO model followed by PCA analysis. Taylor expansion of DHO can then be carried out. The comparison of the first three spectral eigenvectors is shown in Figure 3.13(b), which agree with each other well. Meanwhile, the comparison of spatial eigenvectors for PCA (Figure 3.13(c)) and DHO expansion (Figure 3.13(d)) reveals a structural similarity of over 99.9% and a Pearson correlation coefficient of over 99.8% for first three modes, suggesting that the first three PCA spatial eigenvectors are $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$ derived from $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = \mathbf{A_0} \boldsymbol{Q} \boldsymbol{\omega_0} \circ [\mathbf{1}, \mathbf{Q} - \bar{\mathbf{Q}}, \boldsymbol{\omega_0} - \bar{\boldsymbol{\omega}}_\mathbf{0}]$, respectively, since the variation of $\boldsymbol{Q}$ dominate that of $\boldsymbol{\omega_0}$ for the model system considered. This set of studies thus confirm the physical interpretation of PCA modes, which can be used to substantially speed up the analysis.

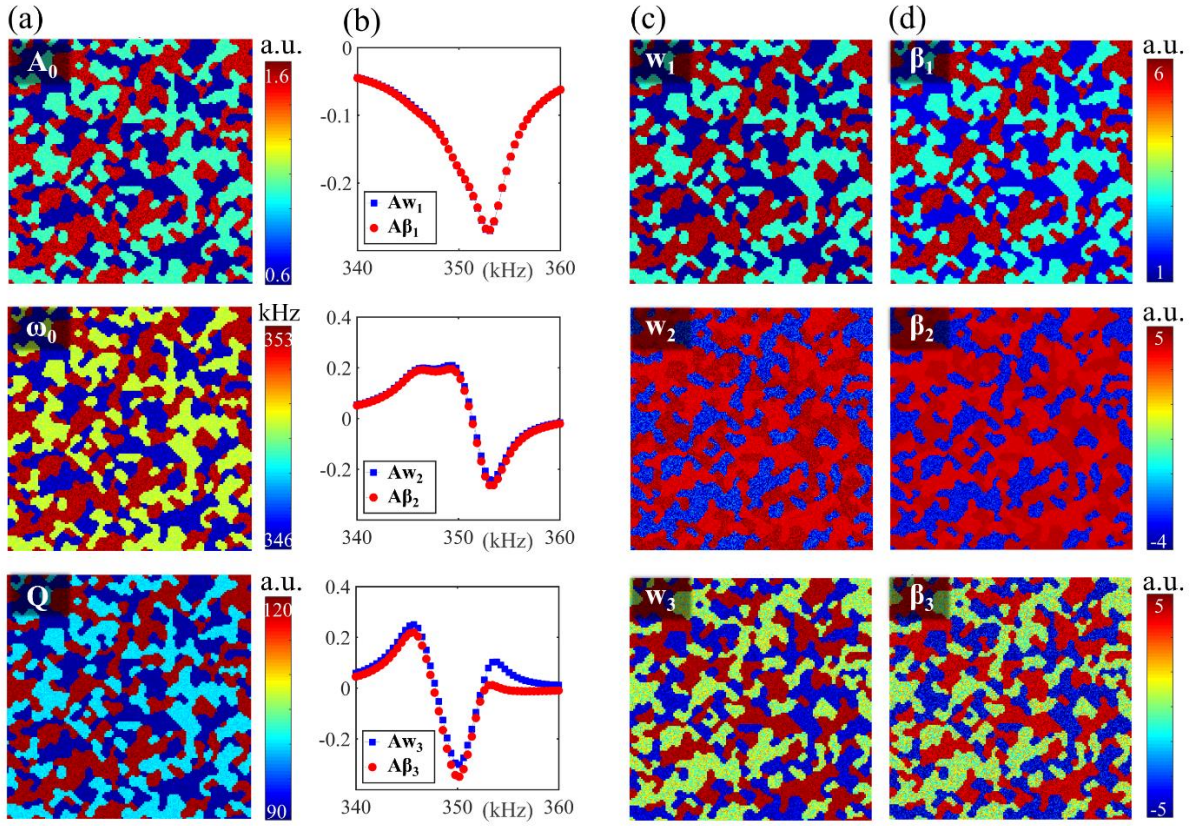

Figure 3.13. Comparison of PCA and DHO expansion for a three-phase model system with distributions of intrinsic amplitude, resonant frequency and quality factor specified in (a), from which the SE-PFM mappings can be constructed based on DHO; (b) comparison of first three spectral eigenvectors of PCA and corresponding DHO spectral basis; (c) the first three spatial eigenvectors of PCA; (d) corresponding DHO spatial basis.

3.4 OTHER APPLICATIONS

As a new way for the acquisition of high-quality data, SE can also be applied to other SPM modes such as pcAFM and SKPM for studying a wide range of materials and structures on the nanoscale.

3.4.1 *SE-pcAFM*

As introduced in Section 2.2.1, it is relatively inefficient to measure local IV curves for a map with high density grid of points. Given that single IV curve usually takes about 1s for each pixel, acquiring all curves for a 128×128 pixels map would take around 4.5hr. Such a long measurement is not practical due to drift of scanning and degradation of materials. Alternatively, we could combine SE with pcAFM, that is, mapping photocurrent distribution under a series of DC biases, from which mappings of short-circuit photocurrent, open-circuit photovoltage, and fill factor can be reconstructed with high spatial resolution [119].

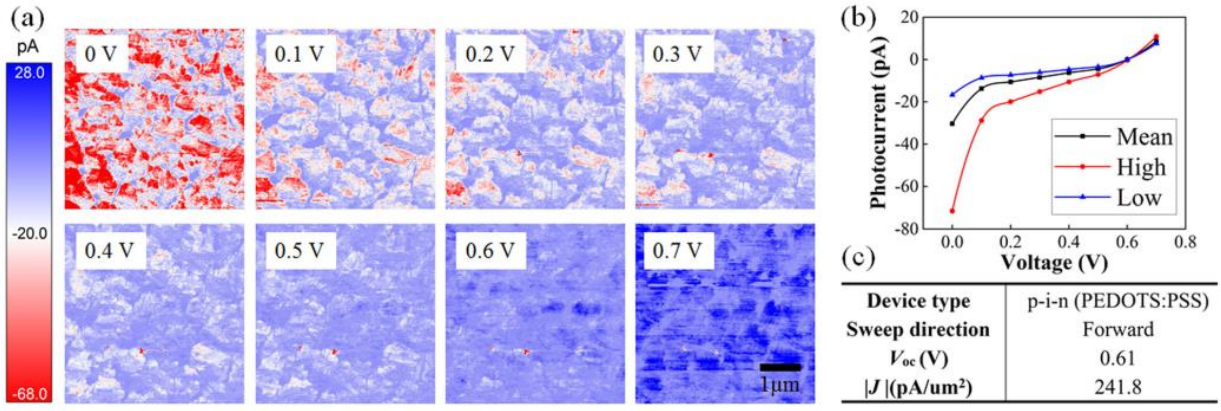

Figure 3.14. Mapping $CH_3NH_3PbI_3$ on FTO/PEDOTS:PSS by pcAFM; (a) photocurrent mapping under a series of DC biases; (b) approximated IV curves; (c) estimated open-circuit photovoltage and short-circuit photocurrent density.

Figure 3.14 shows an example of SE-pcAFM mapping of the sample, and regions of higher and lower photocurrent magnitudes are observed, reflecting structural heterogeneity of the material. Furthermore, with increased DC biases, the photocurrent gradually drops to zero, as expected, and by averaging over higher and lower current regions or the whole area scanned, representative IV curves can be obtained (Figure 3.14(b)), from which the open-circuit photovoltage and short-circuit photocurrent density can be estimated in good agreement with macroscopic measurement (Figure 3.14(c)). If the scan covers sufficiently large area that captures the representative variation of microstructural heterogeneity in the sample, then the effect of microstructure on the macroscopic performance can be quantitatively analyzed.

A more detailed example can be seen in our recent work on triple-cation mixed-halide perovskite $Cs_x(FA_yMA_{1-y})_{1-x}Pb(I_zBr_{1-x})_3$ (CsFAMA), in collaboration with groupmates Dr. Guozhan Xia and Ying Zhang. One of tasks is measuring the photovoltaic hysteresis of this material at nanoscale, which helps illustrate the effect of light-induced large enhancement of local polarization [120]. To this end, we map the local photocurrent of the sample by SE-pcAFM under a series of forward and backward biases ranging from 0 to 0.7 V at a step of 0.1 V, with selected mappings shown in Figure 3.15 along with a comparison of histogram distribution under the same forward and backward biases. It is evident that rather small difference of photocurrent exists between forward and backward biases, which is clear from the mappings (Figure 3.15(a-h)), and can be better confirmed from comparisons of the histogram distributions (Figure 3.15(i-l)). This set of data suggests that at least at the local scale as probed by pcAFM, the photovoltaic hysteresis seems small, despite strong enhancement in polarization under the light illumination.

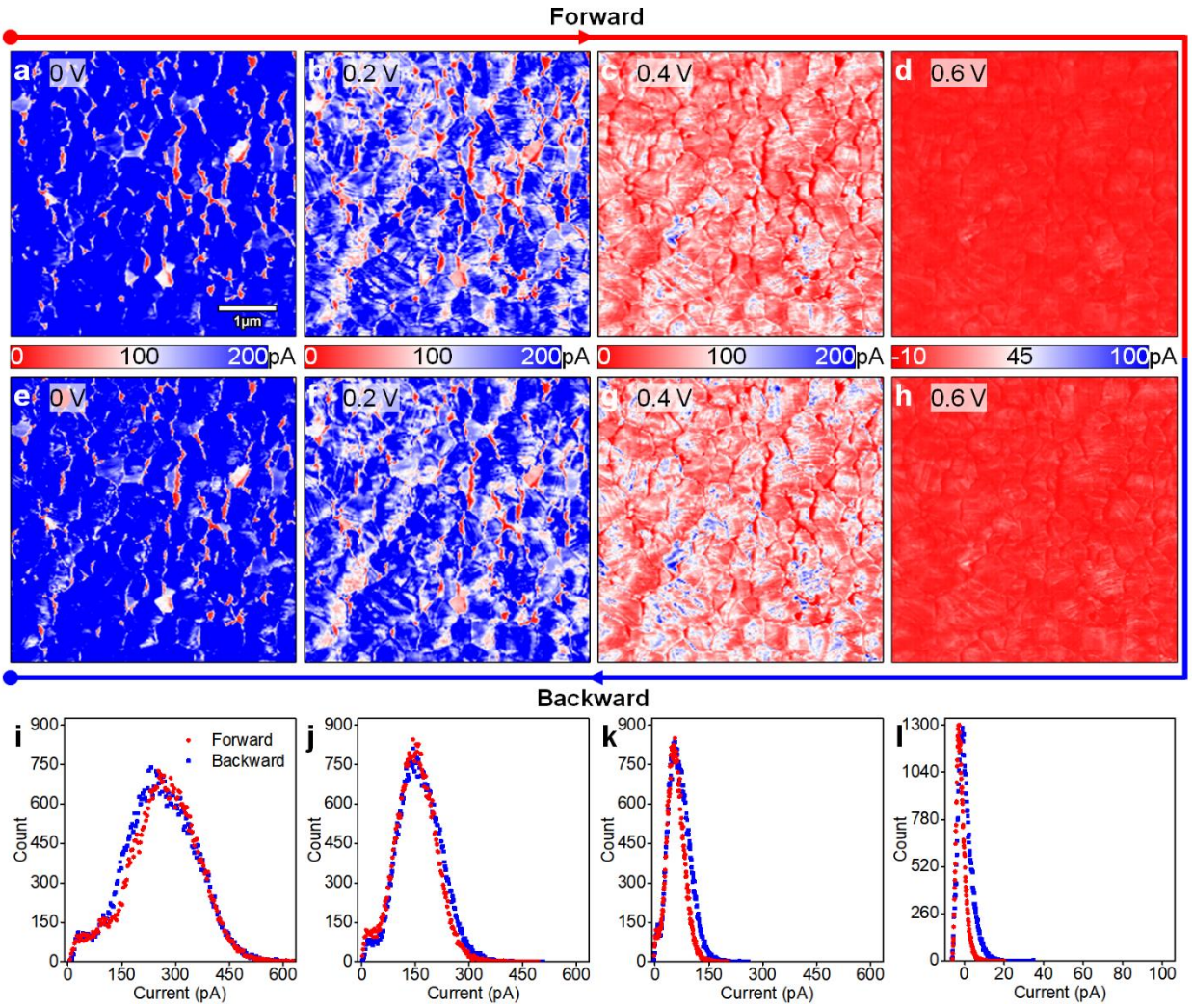

Figure 3.15. Photocurrent mappings under forward and backward biases: (a-d) photocurrent under forward biases of 0, 0.2, 0.4, and 0.6 V; (e-h) photocurrent under backward biases of 0.6, 0.4, 0.2, and 0 V; (i-l) histogram distributions of photocurrent under forward and backward biases.

In order to better quantify the local hysteresis, we carry out PCA to decouple variations of photocurrent with respect to space and bias voltage. The singular values of PCA are shown in Figure 3.16(a), suggesting that the first mode is dominant and contributes to 75.2% of the information contained in the data. The first spectral mode as shown in Figure 3.16(b), on the other hand, reflects the variation of photocurrent with respect to forward and backward biases, and

indeed very small hysteresis is seen, with the largest difference in photocurrent at the same bias measured to be 11.1 pA, which is just 6.9% of forward current at that voltage. Mode 2 contributes to 5.8% of the information according to singular value seen in Figure 3.16(a), and the hysteresis seen in Figure 3.16(c) is comparable to mode 1, with the largest photocurrent difference around 14.4 pA. While this analysis yields valuable information, the local *IV* curves in Figure 3.16(b) measured by pcAFM is quite different from the macroscopic *JV* curves measured on a PSC device shown in Figure 3.16(d). This is contributed by leaking resistance and to a less extent, contact resistance between the probe and the sample [7,121]. Taking into account the leaking resistance as schematically shown in Figure 3.16(f), we added a linear IV curve with a slope of 412 pA/V (or an equivalent resistance of 2.43 GΩ) to the first spectral mode for applied bias < 0.6 V to simulate the case that diode was nearly off. When the bias = 0.7 V, the diode is supposed to be on so that leaking current of the resistance can be ignored. With such a zeroth-order approximation, the first spectral mode of local photocurrent versus forward and backforward biases can be reconstructed, which resemble well the macroscopic *JV* curves measured on a PSC device. Again, rather small hysteresis is observed (Figure 3.16(e)), and the hysteresis index calculated from data is only 1.0%. Note that the open circuit voltage is relatively low due to the lack of hole transport layer [122].

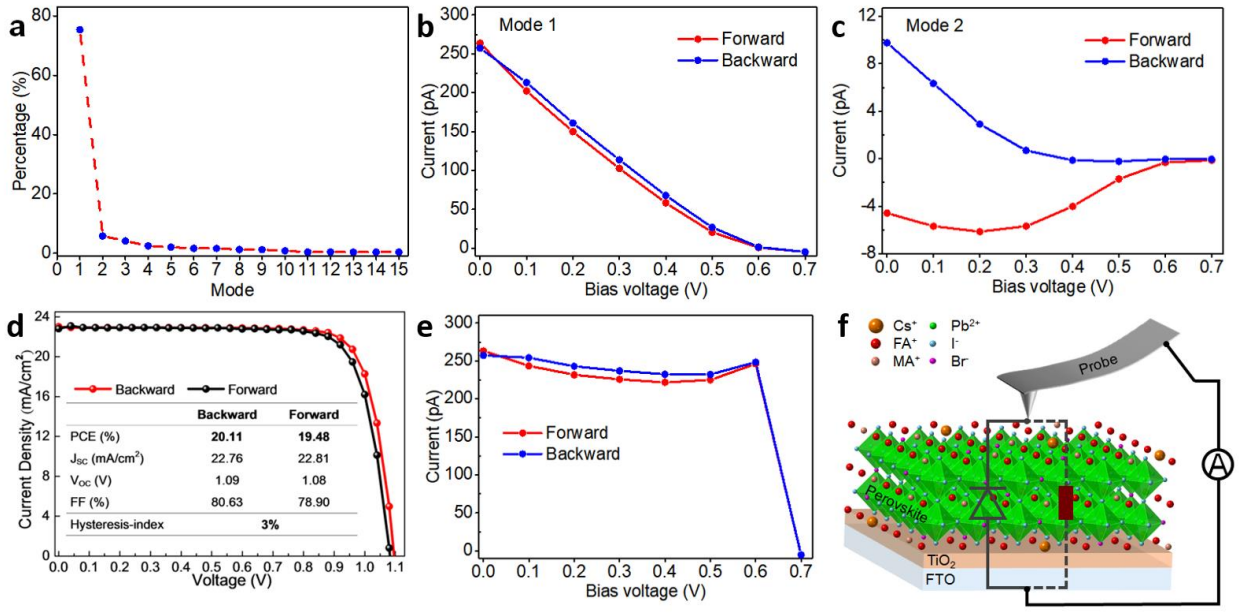

Figure 3.16. PCA analysis of photocurrents under forward and backward biases; (a) variation of singular values; (b) first mode spatial mapping; (c) variation of second spectral mode with respect to bias voltage; (d) macroscopic *JV* curve and photovoltaic properties of perovskite solar cell; (e) spectral variation with respect to bias voltage reconstructed with leading resistance, as schematically shown in (f).

3.4.2 *SE-SKPM*

In this subsection, I will also briefly introduce the application of SE on SKPM, in collaboration with groupmates Dr. Tianjie Li and Dr. Ehsan Nasr Esfahani. According to Equation (1.4), the first harmonic amplitude from electrostatic force between probe and sample surface is $Amp \sim \frac{\partial(C)}{\partial(z)} |V_{DC} - V_{sp}|$, as schematically shown in Figure 3.17.

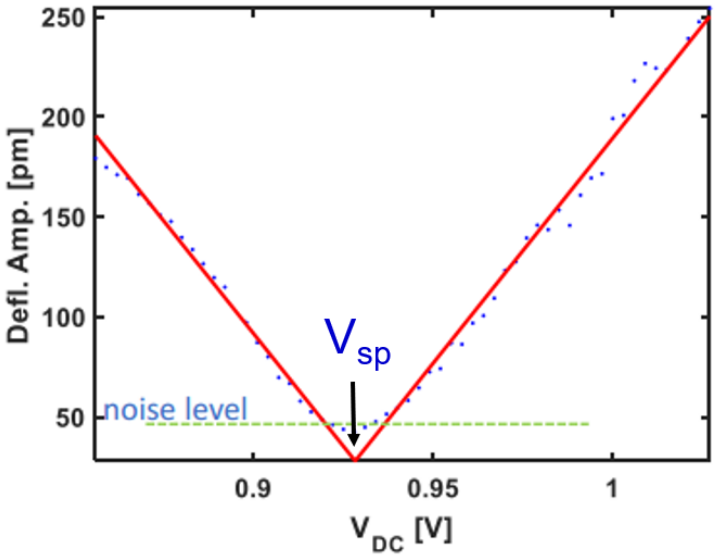

Figure 3.17. Datapoints of amplitude induced by a set of V_{DC} for single pixel.

Conventional SKPM searches V_{sp} by adjusting V_{DC} via a feedback loop until the induced amplitude is minimized. After this process, only V_{sp} is recorded for each pixel and other data points are discarded, thus losing valuable information like the slope of the polyline, which is a good indicator of relative capacitance or local permittivity of measured material. In contrast, SE-SKPM can keep the intact dataset for deep analysis by performing a set of scanning under sequentially increased V_{DC}. For example, Figure 3.18 shows SE-SKPM results of a gold-coated silicon wafer with half self-assembled monolayer (SAM) for bioscience research. The left half is the bare gold layer, while the right half is covered with NH_2-SAM, exhibiting relatively lower capacitance measured by SE-SKPM, which is beyond the capability of conventional SKPM.

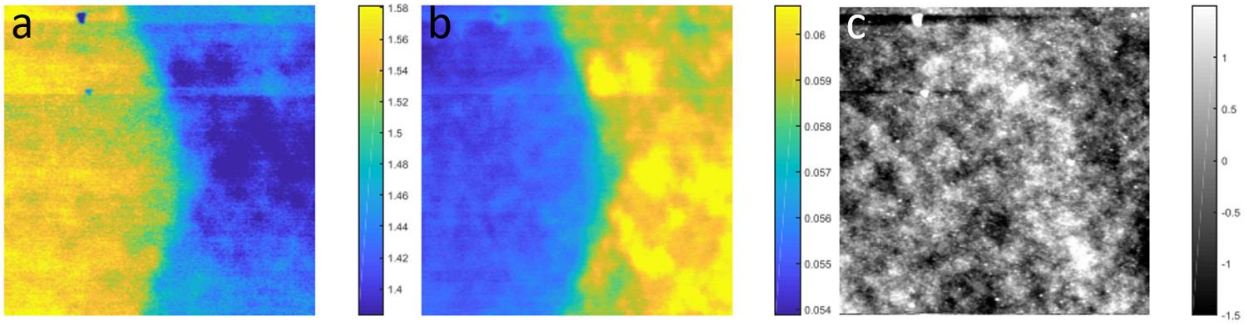

Figure 3.18. SE-SKPM maps of a gold coated silicon wafer with half NH_2-SAM: (a) slope of fitted polyline; (b) surface potential V_{sp}; (c) Topography.

3.5 CONCLUSIONS

Principal component analysis (PCA) has been widely used to compress and visualize multi-dimensional dataset, though its physical meaning is often unclear. Scanning probe microscopy (SPM) measures a wide range of sample properties through tip-sample interactions in term of cantilever dynamics, though intrinsic response is rather challenging to determine quantitatively, often interfered by various crosstalk. Sequential excitation (SE) technique we developed, in combination with dynamics-based DHO model and data analytic PCA, allows us to overcome such difficulties through deep data methodology. Of particular interest is a clear mechanistic understanding of the unsupervised PCA acquired through sound physical principle, making it possible to speed up physical analysis by at least four orders of magnitude. The method can be easily implemented in any standard SPM without the need for any additional hardware and instrumentation. While the technique is demonstrated in terms of electromechanical coupling via PFM, the dynamics involved is universal in SPM, making the method applicable to improving other SPM techniques, for example in SKPM to determine the surface potential, and in pcAFM to map photovoltaic characteristics, among others.

Chapter 4. HIGH-THROUGHPUT SE FOR BIG AND TARGETED DATA

4.1 INTRODUCTION

Although SE can obtain comprehensive information, the required continuous scanning is not appropriate for sensitive materials like $CH_3NH_3PbI_3$ as mentioned before. Even if we manage to complete several scans, the data obtained is still prone to errors due to drift and errors. Thus, we seek to upgrade it to a new SE version completing each round of sequential excitation exactly in the level of a single pixel rather than the whole map as before, requiring only fast one-scan to collect the same dataset and minimizing the damage to the sample [123].

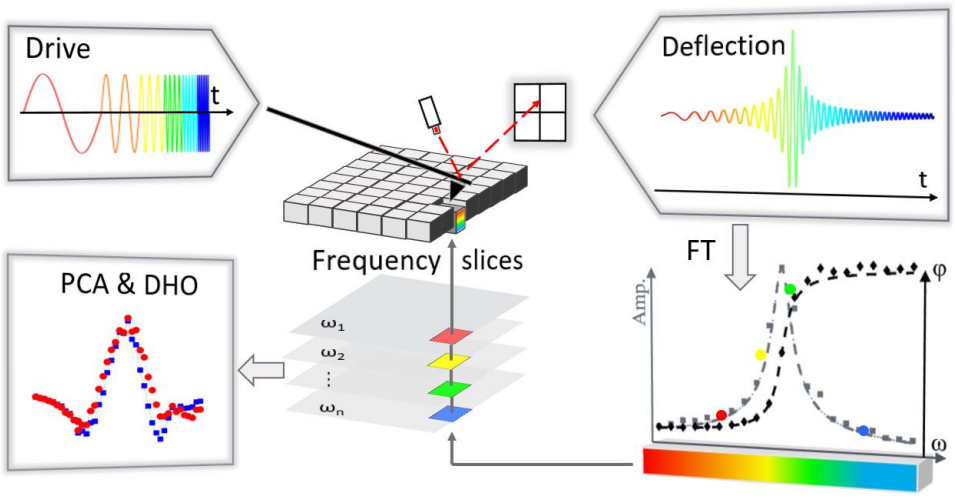

Figure 4.1. Schematic of high-throughput SE-SPM.

We name the technique high-throughput SE, as it is capable of accurately capturing full cantilever dynamics in time domain. This new method can scan at a speed comparable to DART, while keeping much higher spatial as well as spectral resolution. The approach can be applied to any SPM modes that rely on cantilever resonance for imaging. Central to our approach is a high-

resolution waveform designated as Drive in Figure 4.1 drawn in time domain, where it is seen that the frequency increases over time.

Instead of relying on signal generation built in the commercial SPM system, we will use an arbitrary waveform generator (AWG) to synthesize *m* jointed sinusoidal waves with discrete frequencies and convert them to analog signals with GHz ADC rate, which will then be fed into SPM to excite the cantilever from piezo-actuator or external lasers. The resulting deflection signal of cantilever at each pixel is recorded in time domain, and then Fourier transformed into frequency domain using the corresponding drive as a reference. This is equivalent to a digital lock-in, from which *m* pairs of amplitude and phase are obtained at each pixel over the frequency range of interest.

It is also worth mentioning that such a scenario, where the input can be any waves as desired and the raw time-domain output can be directly stored and analyzed, provides us unprecedented freedom to design various experiments for different electromechanics. For instance, as introduced in Section 2.3.3, the second harmonic measurement is a powerful tool for distinguishing local spontaneous polarization from induced Vegard strain due to different mathematical orders of their responses. However, this measurement usually requires that the tip-surface system get resonant while being excited at the half of resonance frequency, thus demanding the coordination of 4 lock-ins in total for DART. Reconfiguration of PID and software is also required and non-trivial. In contrast, the same goal can be quickly realized with high-throughput SE by halving the frequencies of input waves when doing signal synthesis and doubling them as the reference for the Fourier transform.

Here, we will demonstrate high-throughput SE using electrochemical strain microscopy (ESM) to probe defect-induced Vegard strain near grain boundaries in polycrystalline ceria. The

experiment setup of ESM is the same as that of PFM, as both of them measure the strain induced via the probe, as shown in Figure 4.2. These defects (small polarons) are thought to accumulate near grain boundaries, leading to enhanced electronic conductivity in nanocrystalline ceria [124,125]. Using SE, we can map both the linear and quadratic electrochemical strains in ceria near grain boundaries at much higher resolution and fidelity than afforded by DART. In a broader context, mapping such electromechanical coupling via dynamic strain-based scanning probe microscopy such as PFM and ESM can shed considerable insight into data storage of ferroelectrics [12,126], electrochemical conversion of batteries and solar cells [127–129], as well as voltage-gated ionic channels in biology [130].

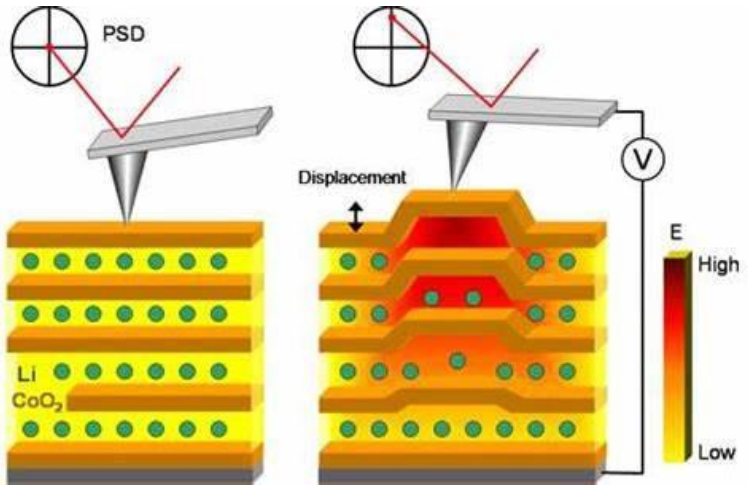

Figure 4.2. Schematic of electrochemical strain microscopy (ESM). The displacement results from the ion movement excited by applied bias via the probe. This image is reprinted from [131].

4.2 Experiment Setup

DART-ESM

This measurement is performed on a Cypher AFM with an AC amplitude of 4 V applied to a Nanosensors PPP-EFM conductive probe. The scan rate is 1.0 Hz. The dynamics of the cantilever

motion is characterized through built-in lock-in amplifiers, which physically reduces full time-domain information to limited frequency-domain data in terms of raw amplitudes and phases. The mappings of corrected amplitude, phase, resonant frequency, and quality factor were then calculated via DHO model.

SE-ESM

Both first and second harmonic resonance SE-ESM scanning were implemented using a UHF-AWG in combination with a Cypher AFM. Before scanning, the UHF-AWG has synthesized the desired waveform associated with digital markers. Once the digital trigger from SPM is received, the AWG begins to convert the waveform into analog signals with 3.5 MHz ADC rate, which will then be fed into SPM to excite the probe. At the same time, a data acquisition (DAQ) system in UHF is also triggered and start to record the full motion of the probe in time domain via deflection channel of the SPM, sampling at 3.5 MHz. As detailed in Figure 4.3, the synchronization is well tuned beforehand by adjusting the input waveform so that the excitation can work in parallel with the SPM XY scanners with delays less than 0.7 ms for each line, which is about a quarter of the scanning time for one pixel. The turn-around time between trace/retrace passes is enough for the DAQ to transfer the data to a computer for Fourier transform (FT) on-the-fly. The same sample and probe were used as DART-ESM measurement.

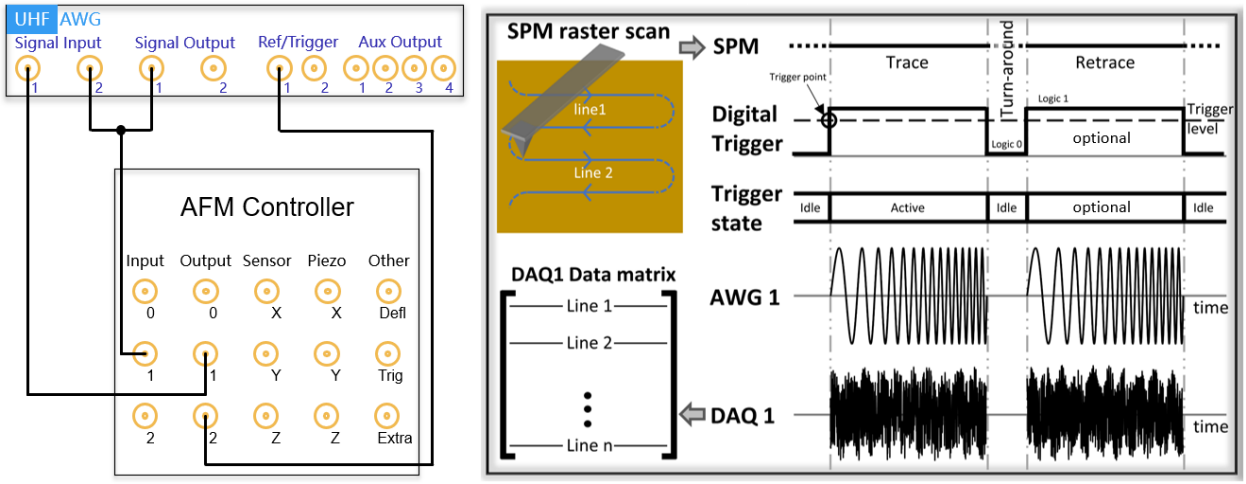

Figure 4.3. (a) Hardware implementation. (b) Signaling and synchronization.

The design of SE waveform

The unit waveform used for one pixel usually consists of 15 different segments, each of which contains multiple periods of sinusoidal waves with a specific frequency within [300 kHz, 400 kHz]. The amplitude of all sinusoidal waves is the same as DART-ESM used. Segment are sorted by the frequencies and jointed with 20 sampling points of zero amplitude, forming the unit waveform. Finally, the unit waveform is first jointed with 100 sampling points of zero amplitude, which helps the relaxation of the probe and the post-processing of data, and then repeated multiple times for one trace of scanning. Parameters may be slightly changed in real experiment to meet the requirement of synchronization mentioned above. For the first harmonic measurement, the input waveform is also used as the reference wave of FT. For the second harmonic measurement, the input waveform is repeated once and then half sampled to generate the reference wave of FT, considering that the sample response was measured at double frequency of the excitation.

4.3 DATA FITTING

We also facilitate the fitting process of spectral data by presenting them on a complex plane and extracting parameters via fast curve fitting. This idea was inspired by model testing techniques

for vibration system [132]. According to Equations (1.1) and (1.2), the solution of DHO model can be described in the complex number form as

$$z(\omega) = \frac{A_0 e^{i\phi_0}}{1-\left(\frac{\omega}{\omega_0}\right)^2+\frac{\omega}{Q\omega_0}\boldsymbol{i}} = \frac{A_0 e^{i\phi_0}}{\alpha+\beta\boldsymbol{i}} \quad (4.1),$$

where $\alpha = 1 - k^2$, $\beta = \frac{k}{Q}$, and $k = \frac{\omega}{\omega_0}$. Since ϕ_0 is the phase offset independent of excitation frequency ω, we first consider a case with $\phi_0 = 0$ for simplicity. Thus, $z(\omega) = \frac{A_0}{\alpha^2+\beta^2}(\alpha - \beta\boldsymbol{i}) = x + y\boldsymbol{i}$ with coordinates $x = \frac{A_0}{\alpha^2+\beta^2}\alpha$ and $y = \frac{A_0}{\alpha^2+\beta^2}(-\beta)$ on the complex plane. Here we will show all these datapoints (x, y) fall on a circle by proving the following equation

$$\left(x - \frac{A_0}{4}\right)^2 + \left(y + \frac{A_0 Q}{2}\right)^2 = \left(\frac{A_0 Q}{2}\right)^2 \quad (4.2).$$

Let

$$\begin{aligned} I &= \left(x - \frac{A_0}{4}\right)^2 + \left(y + \frac{A_0 Q}{2}\right)^2 \\ &= x^2 + y^2 - \frac{A_0}{2}x + A_0 Q y + A_0^2\left(\frac{1}{16} + \frac{Q^2}{4}\right) \\ &= \frac{A_0^2}{\alpha^2+\beta^2}\left(1 - \frac{\alpha}{2} - \beta Q\right) + \left(\frac{A_0}{2}\right)^2\left(\frac{1}{4} + Q^2\right) \end{aligned}$$

By replacing α and β with k and Q, the first term can be written as

$$\begin{aligned} I_1 &= \frac{A_0^2}{\alpha^2+\beta^2}\left(1 - \frac{\alpha}{2} - \beta Q\right) \\ &= \frac{A_0^2}{\alpha^2+\beta^2}\left[1 - \frac{1}{2}(1-k^2) - k\right] \\ &= \frac{A_0^2}{\alpha^2+\beta^2}\frac{1}{2}(1-k)^2 \\ &= \frac{A_0^2}{2}\frac{(1-k)^2}{(1-k^2)^2 + \left(\frac{k}{Q}\right)^2} \end{aligned}$$

It is trivial to show that $I_1 < \left(\frac{A_0}{2}\right)^2 \frac{2(1-k)^2}{(1-k^2)^2} = \left(\frac{A_0}{2}\right)^2 \frac{2}{(1+k)^2} < 2\left(\frac{A_0}{2}\right)^2$. Since $Q > 30$ for most of experiments, it is reasonable to rewrite I as follows,

$$I = I_1 + \left(\frac{A_0}{2}\right)^2 \left(\frac{1}{4} + Q^2\right) \approx \left(\frac{A_0 Q}{2}\right)^2.$$

Therefore, Equation (4.2) is proved and shows that the center of the circle is $\left(\frac{A_0}{4}, -\frac{A_0 Q}{2}\right)$ and the radius is $r = \frac{A_0 Q}{2}$. As for circle fitting, there are many mature algorithms and code to use. In this work, we adopt Pratt method that could finish the analysis of one map within only 2 seconds using the same computation power as previous SE experiments [106,133,134].

The above derivation can also be easily extended to general cases with non-zero phase offset ϕ_0, meaning all datapoints are rotated counter-clockwise with an angle of ϕ_0 about the origin, as shown in Figure 4.4. Thus, the ϕ_0 can be quickly approximated by $\tan^{-1}\left(\frac{x_0}{y_0}\right)$ once coordinates of circle center (x_0, y_0) is known, which is much more efficient and robust than direct fitting highly nonlinear Equation (1.2). Besides, the value $r = \frac{A_0 Q}{2}$ obtained from circle fitting also provides extra condition for faster DHO fitting of amplitude data.

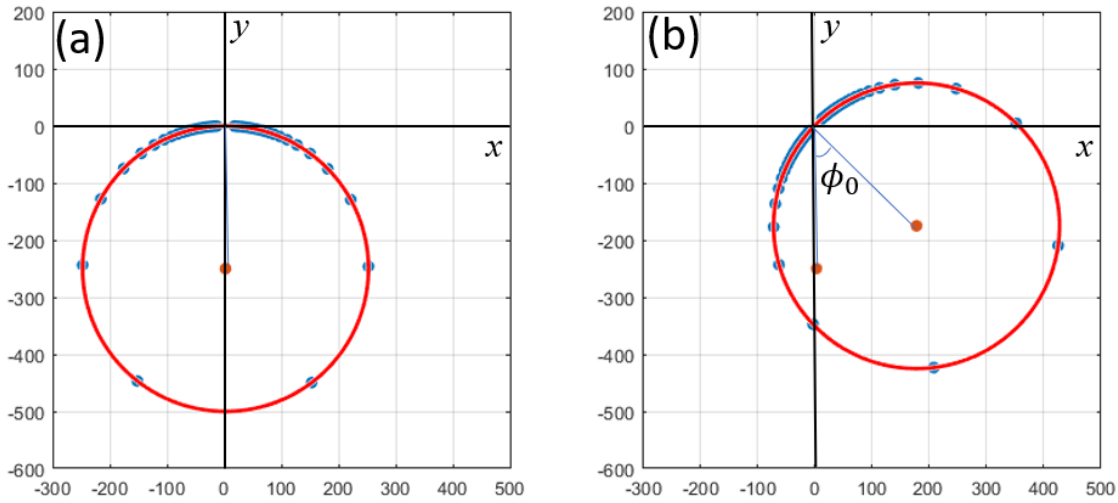

Figure 4.4. Simulated datapoints and their circle fitting results. Model parameters are $A_0 = 10, Q = 50, \omega_0 = 300$kHz and different phase offsets: (a) $\phi_0 = 0$; (b) $\phi_0 = \frac{\pi}{4}$.

4.4 RESULTS OF CERIA

As a demonstration, actual excitation and response in one of our experiments are presented in Figure 4.5. The drive consists of 15 sinusoidal waveforms with distinct frequencies ranging from 347 to 390 kHz connected in sequential manner over a time span of 2.14 ms, as seen in Figure 4.5(a), with 3 such waveforms zoomed in in Figure 4.5(b).

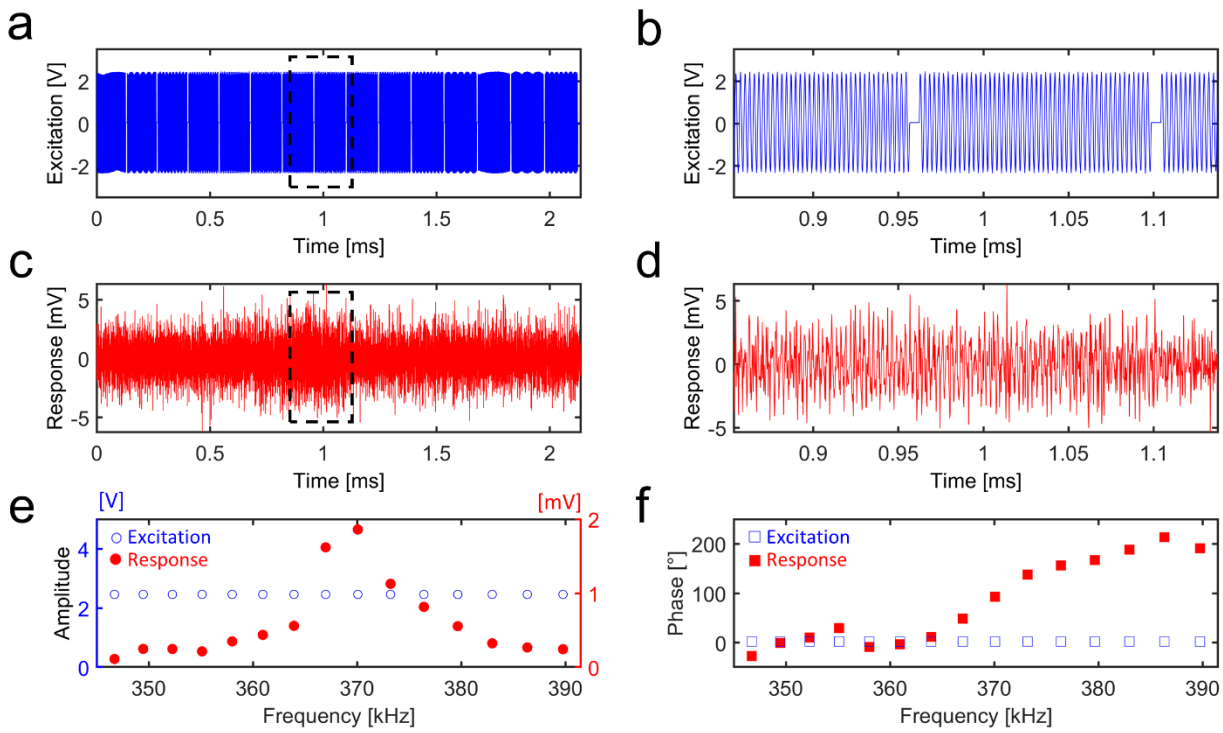

Figure 4.5. Excitation and response of SE in time- and frequency-domains; (ab) drive in time domain; (cd) response in time domain; (ef) drive and response in frequency domain.

Each waveform lasts 50 cycles, leaving sufficient time to acquire response data accurately, corresponding to a line scan rate of 0.8 Hz that is comparable to a typical DART - it takes just 5 min to complete a 256×256 pixels scan. The response is recorded in time domain, as shown in Figure 4.5(cd), which contains large volume of data and is quite noisy. Yet after Fourier transformed into frequency domain, a clear resonant peak emerges in amplitude, upon which phase

jumps by 180°, as shown in Figure 4.5(ef) along with drive signal. Such dynamics is expected from the DHO model, and this illustrates how high-throughput SE works on signal generation as well as data acquisition and processing.

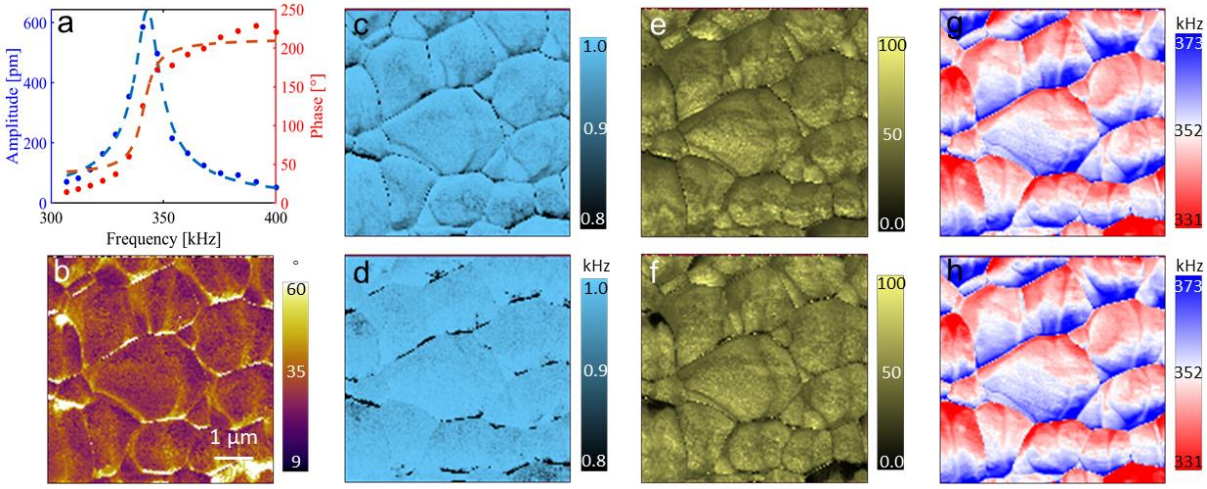

Figure 4.6. (a) DHO fitting of a representative pixel. (b) Phase map. Comparison between maps fitted with sequential amplitude data (the 1st row) and phase data (the 2nd row): (cd) R^2, (ef) Quality factor and (gh) Frequency.

We now apply high-throughput SE to probe linear and quadratic electrochemical strains of granular ceria via first and second harmonic ESM measurements [89,90], for which high fidelity mappings at grain boundaries are essential to examine the proposed accumulation of space charges in its interfacial regions. Since a sequence of $A(\omega_i)$ and $\phi(\omega_i)$ are acquired under SE at each pixel, which can be fitted as shown by Figure 4.6, we can obtain parameters intrinsic to the probed system, including intrinsic amplitude, phase, quality factor, and resonant frequency. Note that both amplitude and phase equations can be used, yielding consistent mappings as shown in Figure 4.6.

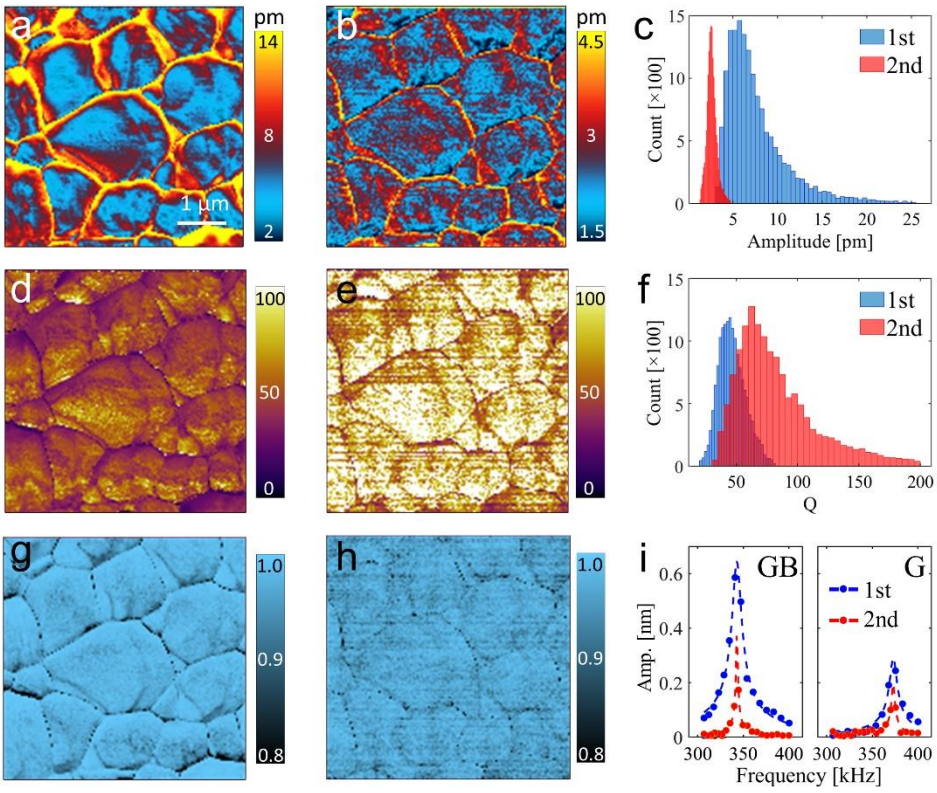

Figure 4.7. Linear and quadratic strains in granular ceria; amplitude mappings of first (a) and second (b) harmonic responses, and their histogram distribution (c); quality factor mappings of first (d) and second harmonic measurement (e), and their histogram distribution (f); mappings of R^2 of first (g) and second (h) harmonic responses indicating DHO fitting fidelity; and (i) point-wise first and second harmonic response at grain boundaries and with a grain.

Linear electrochemical strain obtained as such is shown in Figure 4.7(a), acquired via first harmonic measurement at each of the excitation frequency, while quadratic strain is shown in Figure 4.7(b), acquired via second harmonic measurement at frequencies that doubles each excitation frequency. For both mappings, it is evident that the responses are substantially enhanced at the grain boundaries, which can be seen more clearly from point-wise comparison of first and second harmonic responses at grain boundaries and within a grain in Figure 4.7(i). Furthermore, histogram distributions of first and second harmonic responses in Figure 4.7(c) reveals that the

electrochemical strain is predominantly linear, arising from Vegard strain due to fluctuations in small polaron concentration under the AC excitation, while the quadratic strain due to electrochemical dipole and thus electrostriction is also present [90]. Further insight can be learned from the mappings of quality factors associated with first (Figure 4.7(d)) and second (Figure 4.7(e)) harmonic measurements, where it is observed that second harmonic electrostriction has higher quality factor (Figure 4.7(f)), and thus smaller dissipation, as Vegard strain from small polarons is an energy-dissipative process. These observations are consistent with the proposed accumulation of space charges at grain boundaries, resulting in enhanced response at grain boundaries in both first and second harmonics.

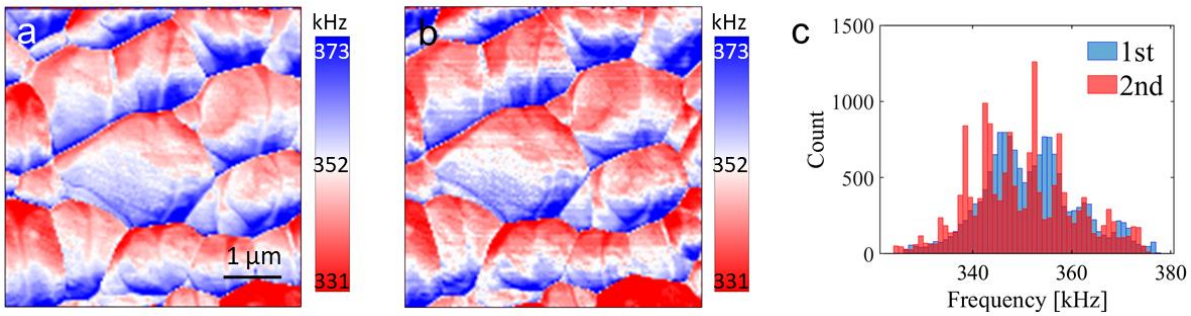

Figure 4.8. Comparison between frequency maps fitted with 1st harmonic amplitude data (a), 2nd harmonic amplitude data (b), and their histogram (c).

We also examine the mappings of resonant frequencies from first and second harmonic measurements presented in Figure 4.8, which match with each other well, demonstrating the high fidelity of the measurement. More importantly, we can examine the accuracy of the DHO fitting at each pixel in terms of R^2 coefficient, a statistical measure on how close the data points are to the fitted regression line. Mappings of R^2 coefficients are presented in Figure 4.7(gh) for first and second harmonic measurements, ranging from 0.8 to 1.0 with respective mean values of 0.97 and 0.96, respectively, demonstrating high fidelity of the fitting. Even at grain boundaries, the fitting

coefficients are mostly over 0.91. This is another advantage of SE, wherein the sequence of data enables us to accurately assess the reliability of DHO fitting for quantitative analysis. Under conventional DART, on the other hand, only two data points are available to solve the highly nonlinear DHO equations, which is not expected to be very accurate.

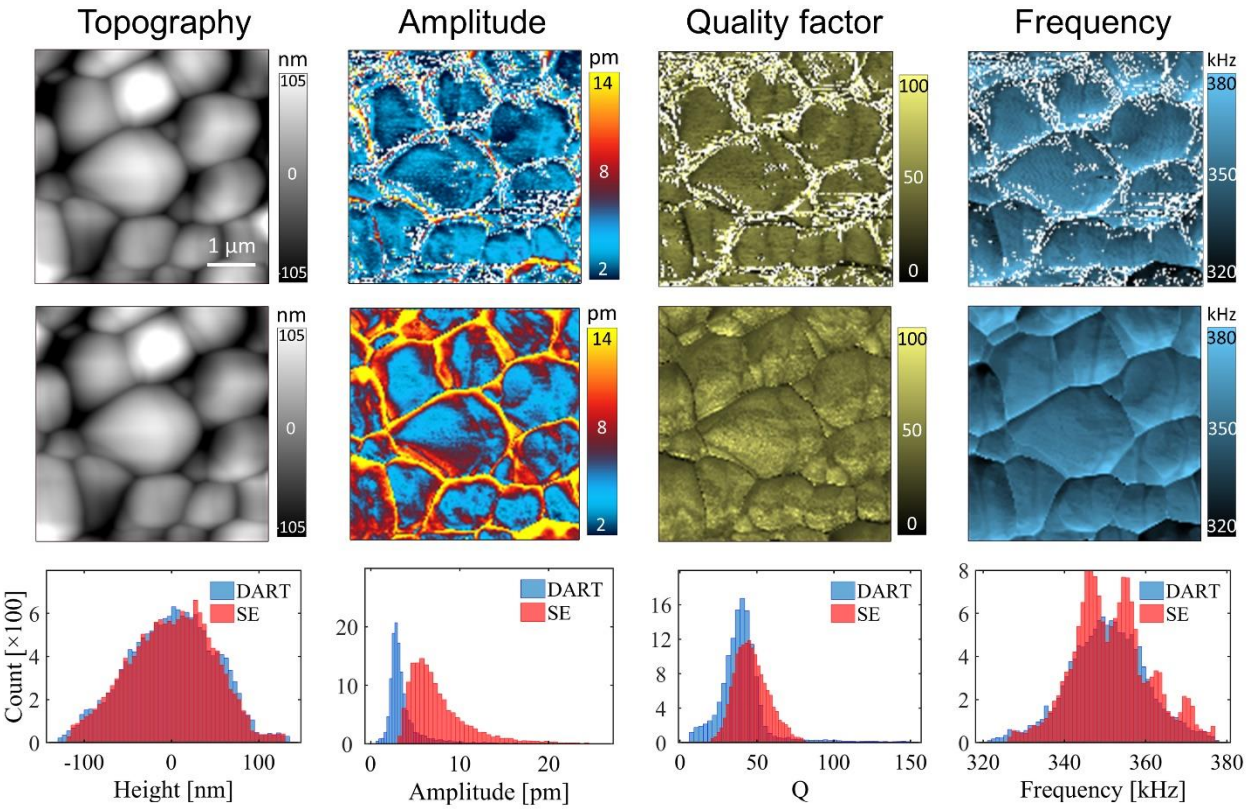

Figure 4.9. Comparison of ESM mappings of ceria acquired by DART (first row) and SE (second row), along with their histogram distribution (third row).

Indeed, we compare the mappings of intrinsic amplitude, quality factor, and resonant frequencies acquired from DART and SE in Figure 4.9, and it is evident that DART not only fails near the grain boundaries, as marked by all the white dots wherein there is no solution found for DHO, but it is also quantitatively different from SE, made evident by the comparison of histogram

distributions in the amplitude – it underestimates the amplitude response substantially. Nevertheless, it is also quite remarkable to note that DART appears to be able to track the resonant frequency pretty well for the most part of grains, and majority of the issues occur at grain boundaries.

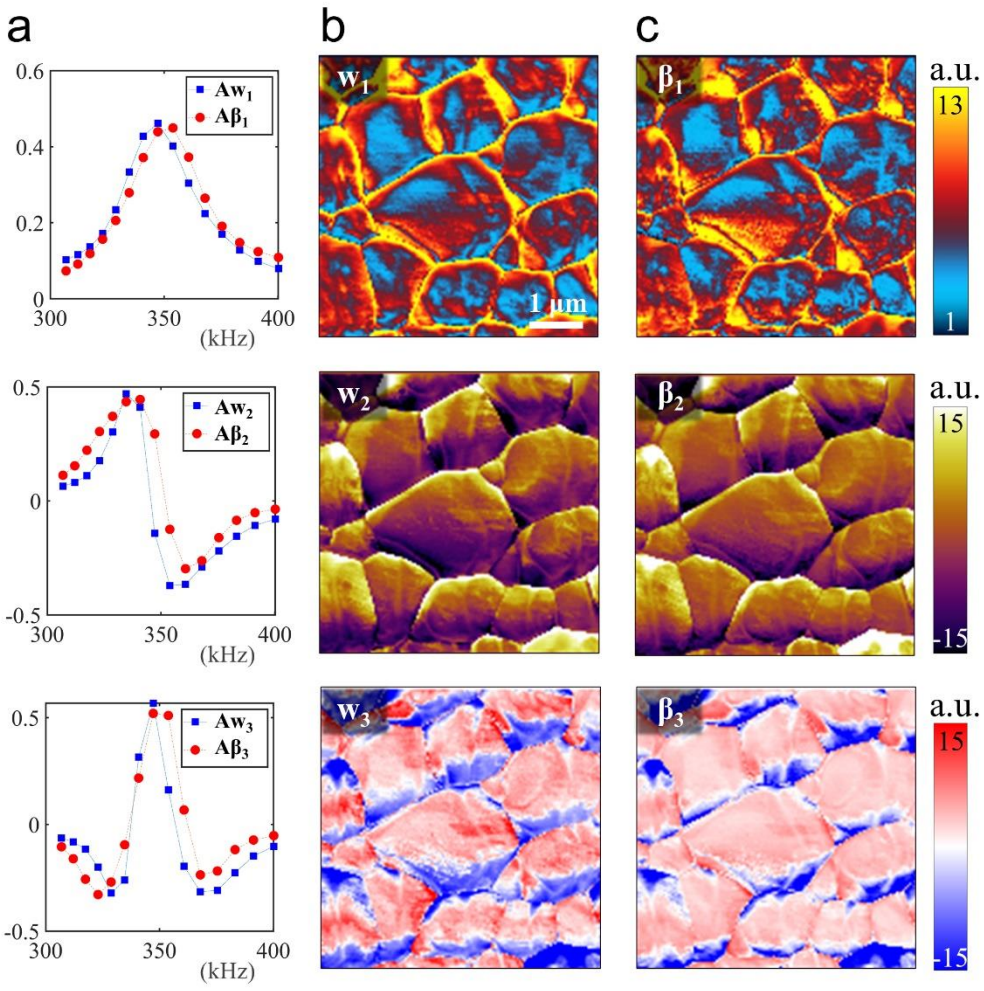

Figure 4.10. Comparison of PCA modes and DHO expansion for first harmonic ESM data of granular ceria; (a) first three PCA spectral modes in comparison with corresponding DHO spectral basis; (b) first three PCA spatial modes; (c) corresponding DHO spatial basis.

As discussed in the last chapter, the large volume of spectral information collected by SE is well suited for data analytics such as PCA. By carefully comparing underlying mathematics of

both approaches, we are able to draw a close connection between PCA modes and DHO expansion [106], rendering clear physical significance to otherwise purely statistical PCA modes. It turns out again that there is one-to-one correspondence between PCA modes and DHO expansion basis [106], which is confirmed by the good agreement as shown in Figure 4.10. The structural similarity (SSIM) (99.1%, 95.8%, and 94.3%) and Pearson correlation coefficients (PCC) (80.4%, 90.7%, and 83.9%) between PCA spatial modes and DHO basis in Figure 4.10(bc) are pretty high, validating our analysis.

4.5 Discussion and Conclusions

With ever-increasing hardware capabilities and computational powers, we are on the brink of a big data revolution for physical science [42,43,47] and SPM provides an ideal playground for the data-driven nanoscience that promises unprecedented new insight [42,106,135,136]. A good example is the recently developed G-mode SPM that relies on brute force big data analytics without pre-committing to a particular physical process [42,45], which is capable of uncovering "*unknown*" mechanisms underlying the physical systems. Here we adopt a more targeted approach, taking advantage of both physical understanding and data power. In particular, we design our excitation signals and data analysis specifically to capture cantilever resonance accurately, without acquiring too many redundant data, and ensuring that the data is clean and relevant to our physical system under the probe. As a result, our experimentation and analysis are highly efficient as well as accurate, which can be further accelerated by DHO-informed PCA analysis. Depending on particular systems under investigation, other forms of excitation signals can be designed and analyzed in a similar manner, for example by varying excitation amplitude for ferroelectric switching, and it is possible to implement a particular form of excitation on the demand. In a sense,

our approach is well aligned with the movement from big data to deep data, *i.e.*, from data mining, correlation analysis, and unsupervised classification to causative data analytics that fuse physical understanding into big data [38,40,41,44,46,47]. For this purpose, innovative experimental and/or computational methodologies to acquire high quality (less noisy), efficient (less redundant), and physically relevant scientific data is essential, and this work is an attempt along this direction.

The power of our approach is best illustrated at the grain boundaries of ceria, wherein enhanced electrochemical strain is evident, attributed to their accumulation of space charges. Since SE does not depend on resonance tracking, it outperforms conventional approach as expected when it comes to large spatial variation. Moreover, SE is able to justify fitting results as well as the validity of raw data by taking advantages of multiple datapoints and statistics, while DART lacks this ability and is more vulnerable to various measurement errors with only two datapoints. As shown in Figure 4.7, even though maps probed by DART on grain looks relatively smooth, it may still underestimate response. Furthermore, I also conduct SE-ESM in collaboration with groupmates Dr. Junxi Yu and Dr. Shanshan Duan on a solid electrolytes $Li_{1.3}Al_{0.3}Ti_{1.7}(PO_4)_3$ at a high temperature (115℃), which is usually challenging for conventional resonance tracking due to continuous deformation of sample originating from environmental temperature variation. But the R^2 map of SE approach in Figure 4.11 is overall greater than 0.9, implying it is still very trustworthy in a temperature-controlled environment, given that fitting is based on multiple datapoints. Therefore, we believe our high-throughput SE provides a powerful tool to resolve spatial variation of such cases with nanometer resolution.

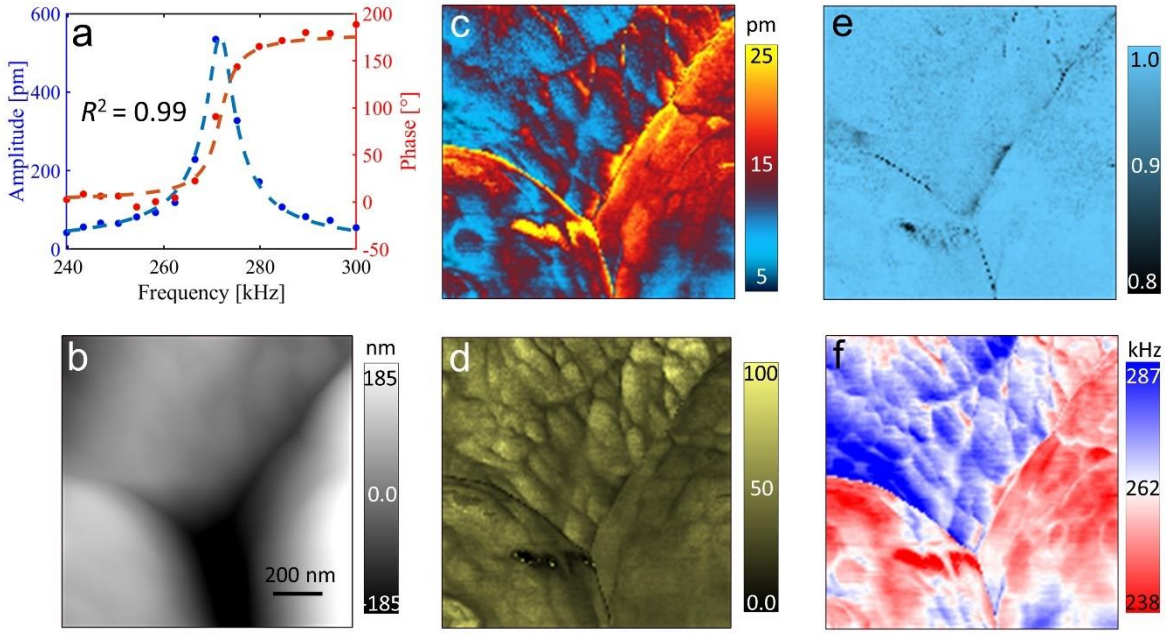

Figure 4.11. SE-ESM mappings of LATP at 115°C: (a) DHO fitting of a representative pixel, (b) Topography, (c) Amplitude, (d) Quality factor, (e) R^2 and (f) Frequency.

In conclusion, we have developed a high-throughput SPM that captures electrochemical strain of ceria in just one scan, having scanning speed comparable to conventional DART, yet with much enhanced spatial resolution and high quantitative fidelity. This enables us to image accumulation of space charges across grain boundaries of ceria with nanometer resolution. The ideas can be applied to a variety of SPM modes for studying a wide range of materials and structures at the nanoscale, especially at the interfaces, and it embodies the spirit of deep data wherein targeted data acquisition and physics-informed data analytics prove to be powerful.

Chapter 5. AN ARTIFICIAL INTELLIGENCE SPM

5.1 INTRODUCTION

Our ultimate goal is to realize an artificial intelligence SPM (AI-SPM) based on machine learning and ever-growing training data. Artificial intelligence (AI) and machine learning have promised to revolutionize the way we live and work. In ancient game of Go, AI has shed unprecedented new insights that have not been recognized by mankind over several thousand years [137]; in medicine, AI has offered diagnosis that rivals the best human doctors [41,138]; and in physics as well as materials sciences, AI has enabled accelerated discovery of new substances, compounds, and mechanisms [37,40,139,140]. One of particularly promising areas for AI is image analysis [141,142], where it far outperforms human beings in pattern recognitions, capable of discerning subtle features that are elusive to the naked eyes. For example, Figure 5.1 shows a demonstration where grain boundaries of perovskites are identified using machine learning. According to such information from AI, grain boundaries can then be zoomed in and critically examined with a higher spatial resolution without user intervene.

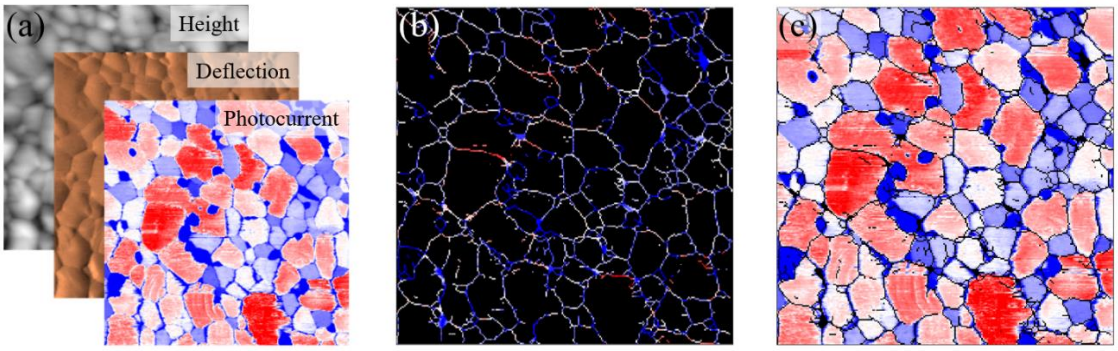

Figure 5.1. (a) Maps of perovskite solar cells obtained with pcAFM mode; (b) grain boundaries and (c) grains extracted by machine learning algorithms.

Although SPM is a powerful tool in probing, elucidating, and manipulating materials and structures at the nanoscale, operating it is tedious at times for experienced users. Important yet subtle information is very often overlooked by novices when conducting experiments. Insights are only realized during the post data processing afterward, which is often too late - it is virtually impossible to get back to the critical locations again for further probing, such as defects, heterogeneities, and interfaces, where the most interesting physics occur. The sample could be decomposed for example, or the elusive critical points could be lost like a needle in a forest.

This is an ideal scenario for AI-SPM we propose, which is capable of not only recognition and classification, but can also follow up with additional probing upon critical features for further insight, saving all the trouble for human users. This vision is schematically shown in Figure 5.2, consisting of an AI-SPM that feeds scanning data to a machine learning algorithm in real time. The algorithm is pre-trained with data accumulated in our lab in the past 10 years for material classification and feature recognition. Based on a particular class of materials recognized by the AI, additional features that are relevant to the underlying system will be identified on the fly, such as domain walls (DWs) or grain boundaries (GBs), among others. Through control algorithms, the probe will get back to the identified critical feature in real time and carry out further experimentation appropriate for the probed system on the fly, yielding additional data for analyzing the underlying physical processes. Key to such a vision is a highly efficient machine learning algorithm capable of high-fidelity pixel-by-pixel recognition instead of relying on data from full mapping, so that real time classification and control are made possible during scan. We believe such a strategy is applicable to a wide range of instrumentations and experimentations.

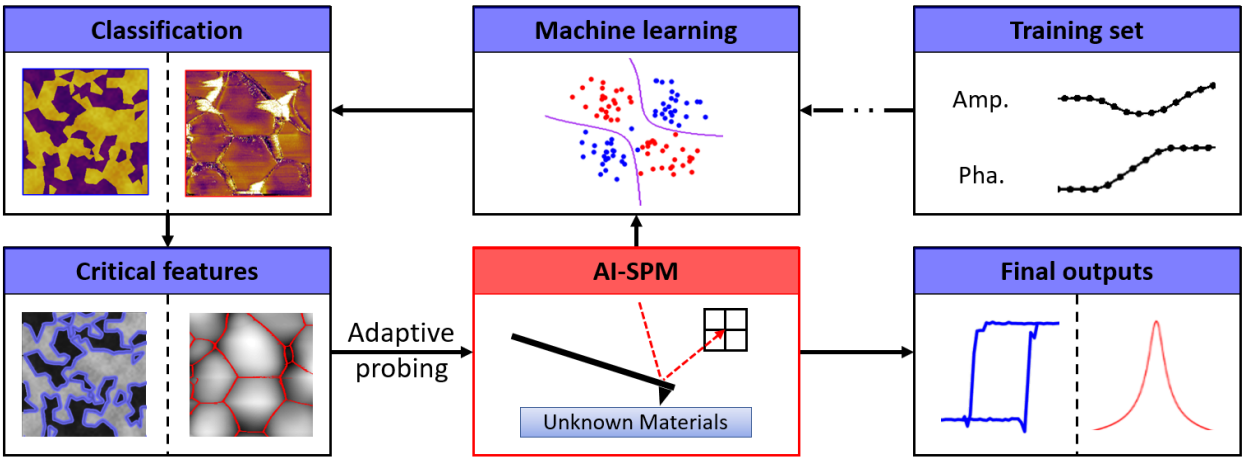

Figure 5.2. The concept of AI-SPM that feeds scanning data to machine learning in real time and classifies samples under probing accordingly, with appropriate features identified. Additional experiments are then carried out on the fly near critical spots for additional data and further insight, all without human interference.

5.2 Support Vector Machine

The machine learning algorithm adopted for our AI-SPM is called support vector machine (SVM). It is widely used in industry and academia [143], with its applications ranging from handwritten digit recognition for postal automation services [144], E-mail spam filtering [145], and accelerating discovery of new piezoelectric materials [40]. SVM can be easily trained with a set of labeled examples, each of which consists of a fixed number of features $\boldsymbol{x} = (x_1, x_2 \ldots, x_n)$ and a label y illustrating whether it belongs to one of the two categories (y =+1 or -1). As $\boldsymbol{x}$ can be regarded as datapoints in a n-dimensional space, we want to find an (n-1)-dimensional hyperplane that divides these points into their respective classes. This hyperplane can be represented as $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = 0$, where $\vec{\boldsymbol{w}}$ is the normal vector to the hyperplane and b is the intercept. Thus, the distance from a datapoint $\boldsymbol{x}$ to the hyperplane is $r = \frac{|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b|}{\|\boldsymbol{w}\|}$ and the corresponding prediction function is

$$f(\boldsymbol{x}_i) = \begin{cases} +1, & \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b > 0 \\ -1, & \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b < 0 \end{cases}.$$

How to find appropriate hyperplane for the classification? If the training dataset is linearly separable, as schematically shown in Figure 5.3, we can search for two parallel hyperplanes (parallel lines in 2D space) which not only separates the data but also keep a distance as large as possible between them. The region bounded by these two hyperplanes is usually called "margin", and the hyperplane lying halfway between them is the maximum-margin hyperplane. In other words, training a linear SVM classifier aims to find the value of $\boldsymbol{w}$ and b that make this margin as wide as possible. In fact, the max-margin hyperplane is completely determined by datapoints that lie nearest to it, so they are so-called support vectors [143].

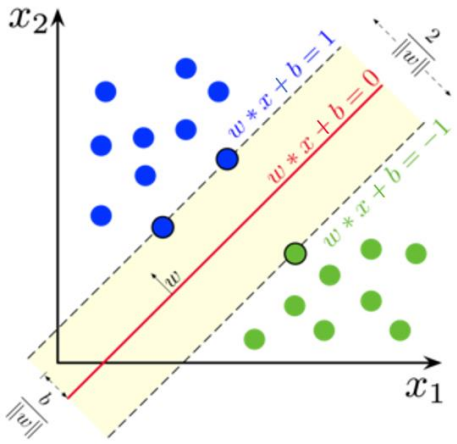

Figure 5.3. Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors. This image is reprinted from [146].

Without loss of generality, we can describe these hyperplanes as $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = \pm 1$ (dashed lines in Figure 5.3) and the distance between them is $d = \frac{2}{\|\boldsymbol{w}\|}$. Therefore, finding maximum-margin requires maximizing $\frac{1}{\|\boldsymbol{w}\|}$ or minimizing $\|\boldsymbol{w}\|$ and it can be described by this optimization problem:

$$\min_{\boldsymbol{w},b} \frac{1}{2}\|\boldsymbol{w}\|^2$$

$$s.t.\ y_i(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i + b) \geq 1, i = 1,2,\dots,m.$$

We prefer minimizing $\|\boldsymbol{w}\|^2$ to minimizing $\|\boldsymbol{w}\|$, because the former will give the same result and has a simple derivative at $\boldsymbol{w} = \boldsymbol{0}$. The above problem is known as a Quadratic Programming (QP) problem with linear constraints [147]. Many off-the-shelf solvers are available to solve it using a variety of techniques that are outside the scope of this dissertation. Here we just briefly introduce how SVM works for linear classification, while other methods such as soft margin and kernel tricks can be used to extend SVM to cases in which the data are not linearly separable [148]. In following sections, we will show how to integrate SVM into the AI-SPM system and evaluate its performance on real data.

5.3 TRAINING OF THE AI-SPM

To demonstrate the concept of AI-SPM, we consider dynamic strain-based scanning probe microscopy (ds-SPM) that is widely used to probe electromechanical coupling at the nanoscale, including piezoresponse force microscopy (PFM) and electrochemical strain microscopy (ESM) [12,97,127–129,149], both of which excite samples through a charged conductive probe and measure the corresponding local deformation. The electromechanical coupling is ubiquitous in both natural materials, synthetic devices, and biological systems, such as ferroelectric materials [12,126], lithium ion batteries [127–129], and voltage gated ion channels [130], underpinning a wide range of functionalities in information processing, energy conversion, and biological processes. Despite their vast different microscopic mechanisms, these electromechanical couplings often exhibit themselves in ds-SPM as apparent piezoresponse [95], and it is quite challenging for users to discern their dominating microscopic origin. A couple of examples are

shown in Figure 5.4, wherein the amplitude and phase mappings of dynamic strain for a typical ferroelectric lead zirconate titanate (PZT) and electrochemical $LiVO_3$ (LVO) are compared. While they have quite different microscopic mechanisms, the mappings closely resemble each other except for some subtle difference: the 180^o phase reversal at the interface with much reduced amplitude as observed in PZT is a signature of ferroelectric domain wall, which is not present in LVO. The question is whether we can train a machine learning algorithm that first differentiates ferroelectric domains from non-ferroelectric mappings, and then responds with additional probing necessary at critical locations for further analysis, for example identify DWs in ferroelectrics and GB in electrochemical materials, after which detailed probing relevant to the particular system can be carried out across these important materials interfaces.

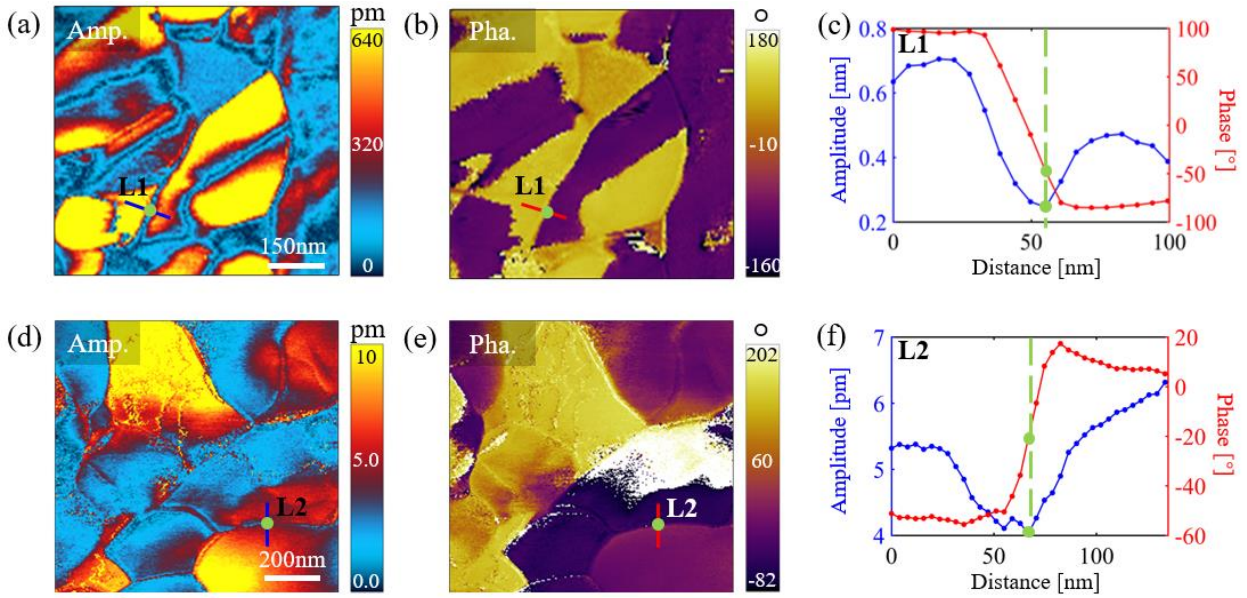

Figure 5.4. Comparing the amplitude and phase mappings of dynamic strain measured by ds-SPM for (a-c) ferroelectric PZT and (d-f) electrochemical LVO, along with their respective line profiles.

5.3.1 *Principle of method*

180^o domains are commonly presented in ferroelectric materials and they usually exhibit much reduced piezoresponse on DWs with 180^o phase contrast, as revealed by Figure 5.4(c). On the other hand, amplitude and phase behaviors of non-ferroelectric solids, such as electrochemical materials, are usually not well defined as revealed by Figure 5.4 (f), wherein the phase contrast is smaller than 180^o.

According to these observations, we employ a SVM model to develop a physics-based classifier that is capable of extracting ferroelectric DWs pixel by pixel from the inputted PFM mappings, thereby helping distinguishing ferroelectric materials from electrochemical ones, for which a different algorithm is introduced to extract GBs from SPM topography mappings, as detailed in Figure 5.7. Note that while popular convolutional neural network (CNN) has achieved remarkable success in the field of image recognition, it can only categorize a whole map but not capable of delineating the exact DWs or GBs of interest. Fully convolutional networks derived from CNN is capable of identifying lattice atoms in raw scanning transmission electron microscopy (STEM) data and should be applicable to our problem, but it requires extensive GPUs to facilitate the training process as well as a large amount of training data with DWs or GBs accurately labelled at the level of pixel. Such sophistications are unnecessary for our particular application, since SVM-based AI algorithm needs only a small dataset that can be trained in less than 10s on an ordinary PC, making it widely accessible. More importantly, this SVM-based algorithm is much more efficient than CNN, critical for classification and control on the fly. On the other hand, we also note that due to subtleties and complexities often exhibited in ferroelectric and electrochemical materials, careful classification based on machine learning beyond simple

rule-based analysis is necessary for our AI-SPM system, and we will come back to this point later with illustrating examples.

As schematically shown in Figure 5.5, we first prepare a training dataset for our SVM model, wherein amplitude and phase variations across a morphology interface are used as indicators to classify whether the interface is a ferroelectric DW or not. For each pixel P_0 (marked as green star) on PFM maps of ferroelectric $LiNbO_3$ in Figure 5.5, 6 adjacent pixels (marked as white dots) are picked from a line centered at P_0 and parallel to its phase gradient. The distance from these 6 pixels to P_0 are fixed and their respective amplitude and phase, 14 features in total, are sufficient to represent the local variance across P_0. Such 14 features with a label of the pixel (DW or not) is then fed into the SVM model.

Figure 5.5. Schematics of training dataset using PFM mappings of $LiNbO_3$ and the corresponding classification of ferroelectric DWs.

Since each map contains 256×256 pixels and thus generates almost the same amount of training data (with the exception on the map border), it turns out that only 5 pairs of ds-SPM maps are sufficient to train the SVM model, making it highly efficient compared to CNN that has to use a whole map as one training example. More details about our training set is presented in the next subsection. SVM first represents these training data as points in a hyperspace, the dimensionality of which depends on the number of features n. It then tries to find the maximum-margin hyperplane $\boldsymbol{w}^T \cdot \boldsymbol{x} + b = 0$ to separate these points into two categories, as schematically shown in Figure 5.5.

The points that are most close to the hyperplane from both sides are called the support vectors, as marked on Figure 5.5, and the larger distance from these two points to the hyperplane (so-called functional margin), the better performance of the trained classifier model. Finally, testing data from new images that are denoised by a local median filtering with a 3-by-3-pixel window can be mapped into the same space, and then classified according to which side of the hyperplane they fall into, revealing whether it is on a DW or not. Note that the pixel-by-pixel recognition of SVM makes it possible to adaptively adjust experimental parameters during scanning, while for CNN that relies on full mapping for recognition, such real time adjustment is not possible. This process is repeated for all pixels except those on the border of maps, creating a binary mask with DWs marked as True. Since DWs are continuous lines on the map, the length of longest line on the binary mask is used to judge whether there are ferroelectric DWs or not.

5.3.2 *Detailed implementation*

Training data. Although real maps with manually highlighted DWs can be utilized as training dataset, we also succeed in training a model based on simulated maps without tedious labelling work. Specifically, a binary mask of random DWs is first generated and then rendered to mimic the pattern of real maps with respect of morphology, as shown in Figure 5.6. Considering that the histograms of phase maps are always concentrated around several specific angles, delivering very sparse information, we randomly change the phase offset when picking the 14 features to build a robust model that can work for other realistic cases with various phase offsets. We also intentionally vary the scaling of amplitude features and add white noises to them for the same reason.

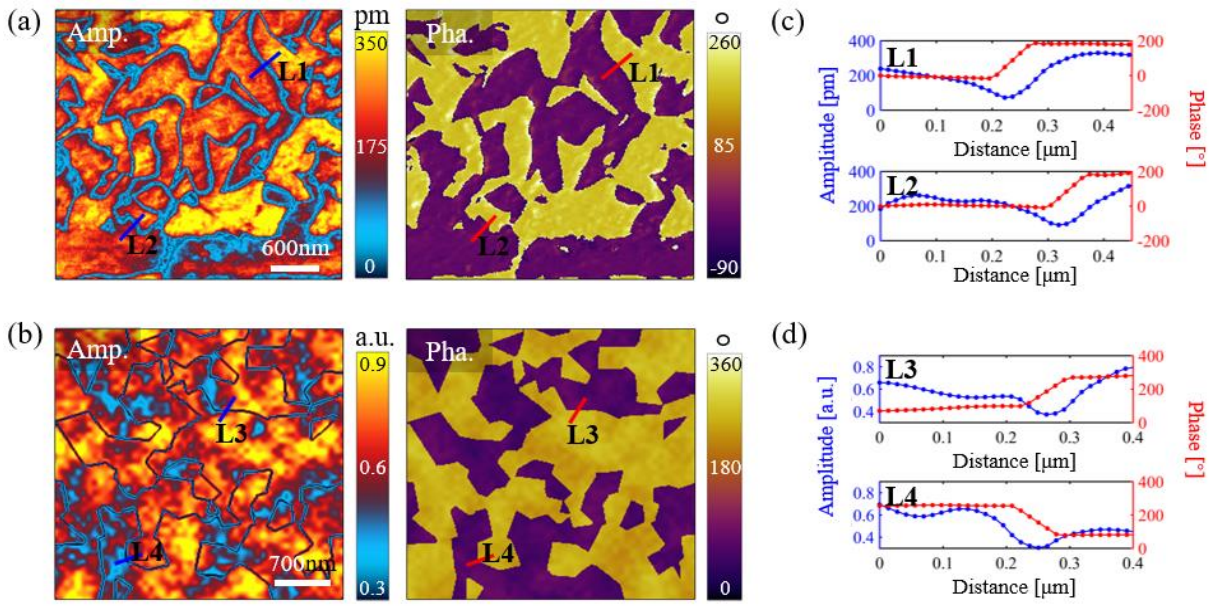

Figure 5.6. (a) Real PFM maps of a PMN-PT sample; (b) simulated PFM maps of a FE sample. Corresponding amplitude and phase scans of: (c) lines L1 and L2, (d) lines L3 and L4.

Training of SVM. Functions "fitcsvm" and "predict" in MATLAB are employed to train SVM models and classify new examples, separately. Gaussian kernel is used to implicitly map input dataset into high-dimensional feature spaces, enabling an efficient non-linear classification. Since amplitude and phase values are two set of independent features, we use them separately to train two different SVM models – Amp model and Pha model. When doing prediction, the algorithm will assign a pixel as point of DWs if and only if both Amp and Pha models determine it to be DW based on input features.

AI-SPM system. The AI-SPM experimentation is performed with a commercial Asylum Research MFP-3D AFM. The system first conducts a DART PFM mapping to survey possible DWs with the pretrained SVM model. If DWs are found, the probe will move to locate in the middle of the longest DW identified. Then, the system will zoom in on this area with a scan size half of previous one to double-check those DWs. Finally, the middle point on the DW as well as

other 4 points across it will be marked so that a set of switching spectroscopy PFM (SS-PFM) tests can be completed on each of these points. If no DWs are found, the material will be further assessed for its apparent piezoresponse mechanism. The system will highlight GBs in a height map by using the method discussed in Figure 5.7 and then zoom in on a specific GB to finish a few first and second harmonic comparison experiments. The decision-making process of location here is similar to that of DWs. Corresponding video can be seen on this reference [150]. The AI is implemented with MATLAB code and can automatically send commands to SPM after analyzing the scanning data on the fly.

Extraction of grain boundaries (GBs). As shown in the upright inset of Figure 5.7(a), the first-order derivative of height map usually is not continuous at GBs, leading to an overwhelming second-order derivative values there compared to those of nearby area. By setting a threshold value at 90% of the maximum second-order derivative of the map, a binary mask of GBs can be produced as shown in Figure 5.7(d). This method works well in general, but may run into difficulty in identifying exact location of GB when there is large topography variation, which calls for more advanced algorithms.

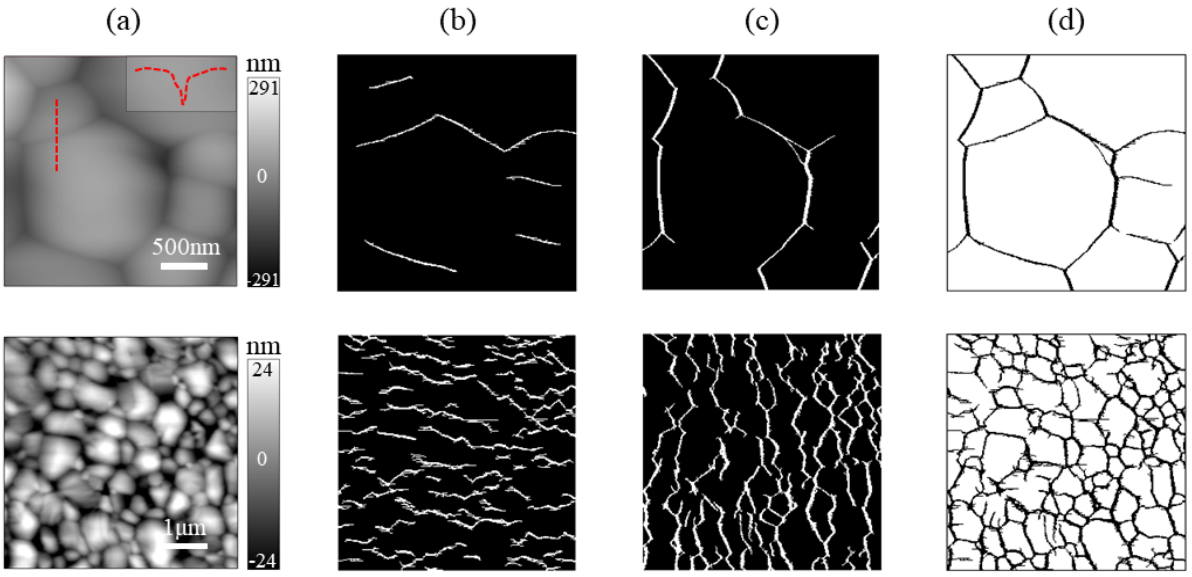

Figure 5.7. GBs of Ceria (1st row) and $CH_3NH_3PbI_3$ (2nd row) identified by the second-order derivative method. (a) Topography and a height profile as an upright inset; (b) second-order derivative along y direction; (c) second-order derivative along x direction; (d) GBs generated from the union of (b-c).

Canny edge detector. It is a technique to extract edges from different vision objects in images [151]. In this work, we used edge function in MATLAB with canny method and default setting.

Multi-harmonic measurements. For each measured point, a set of AC voltages from 1.5 to 7.5 V were applied with an increment of 0.5 V. At each voltage step, the sample is excited around f_0 first and then around $f_0/2$, thereby generating two set of tuning data around f_0. The corresponding first and second harmonic amplitude can be extracted by fitting the raw data with the DHO model.

5.4 Results and Discussion

One positive example is shown in Figure 5.8(a), wherein 180° DWs in molecular crystal of diisopropylammonium bromide has been correctly identified by the SVM, as marked by the red lines and overlaid on the topography. The negative example of Figure 5.8(b), on the other hand,

shows ESM maps of a LiV_3O_8 sample for which GBs are identified instead. In fact, we have surveyed 7174 ds-SPM mappings accumulated in our lab in the past 10 years, and the normalized confusion matrix in Figure 5.8(c) shows that 97.3% of 475 maps predicted having 180° DWs are correctly classified, while 99.6% of 6699 maps predicted having no 180° DWs are properly identified as well.

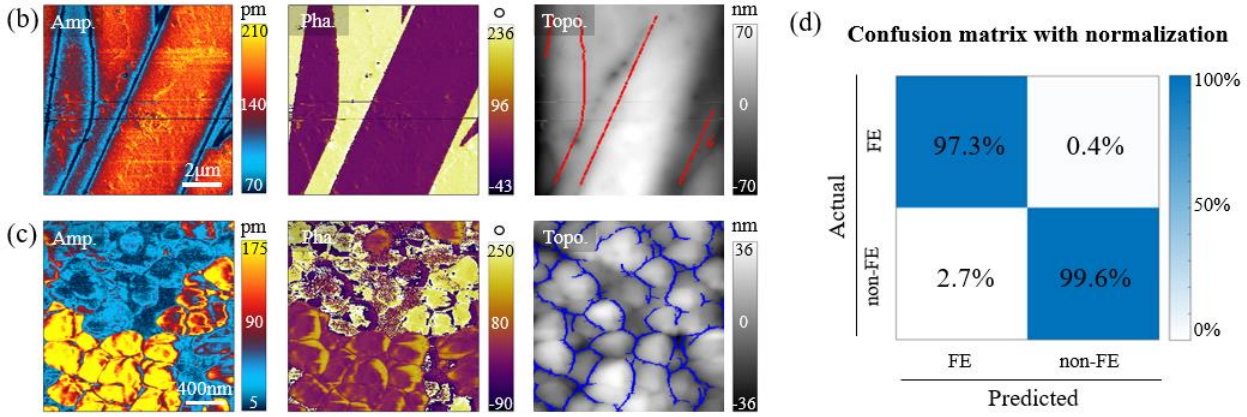

Figure 5.8. Performance of SVM algorithm for AI-SPM; (a) PFM mapping of ferroelectric diisopropylammonium bromide, with 180° DWs identified and overlaid on topography; (b) ESM mapping of electrochemical LiV_3O_8, with GBs identified and overlaid on topography; (c) normalized confusion matrix of classification performance on 7174 ds-SPM maps of various materials.

Additional examples are presented in Figure 5.9. These results confirmed that the SVM-based AI algorithm is capable of classification and feature identification of ferroelectric materials with 180° DWs. On the other hand, popular Canny edge detector often fails, as detailed in Figure 5.10. This is because phase contrast at ferroelectric DWs in real materials often deviates from idealized 180°, and it is often interfered by topography features as well. Since edge detector are highly sensitive to the gradient of a map, slight phase distortion due to scanning disturbance, impurities, topography variation, or artificial pattern may cause false identification, as made evident in Figure

5.10, while our machine learning algorithm does not suffer from such problems. This demonstrates the need for machine learning-based AI instead of simple rule-based analysis in the classification and identification.

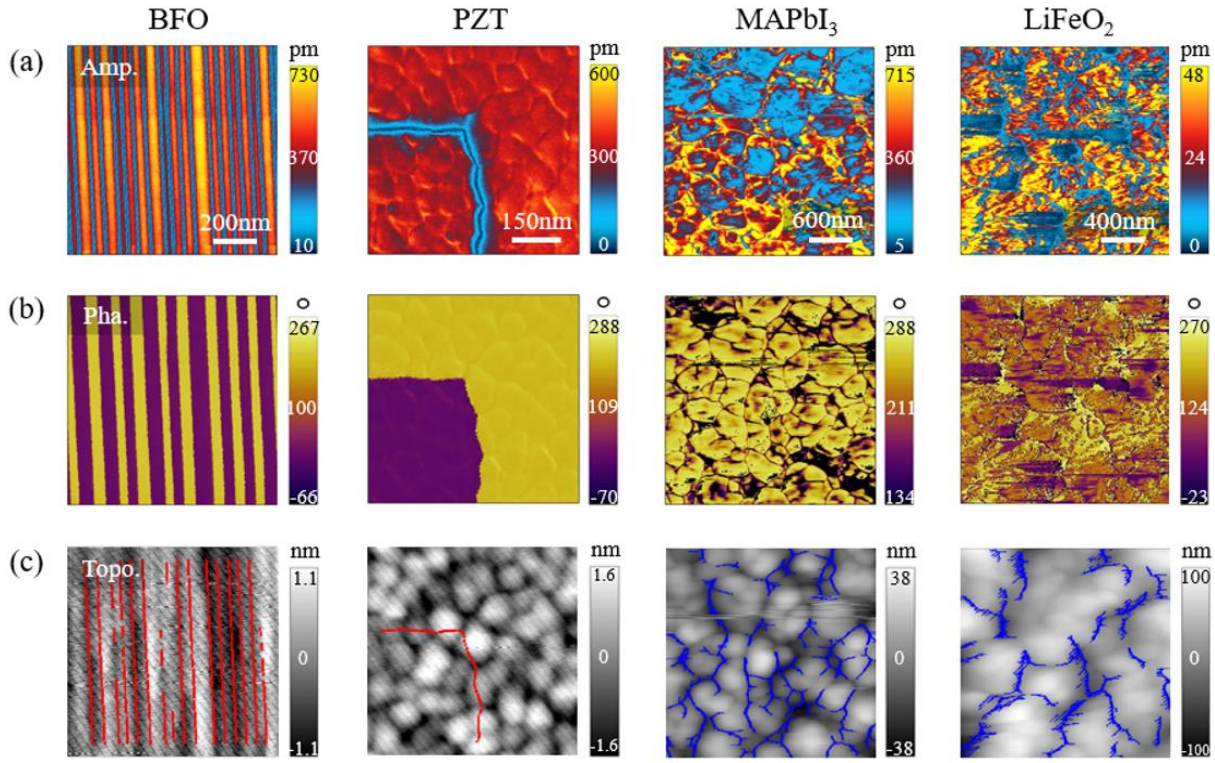

Figure 5.9. Other typical ferroelectric and electrochemical mappings determined by the AI.

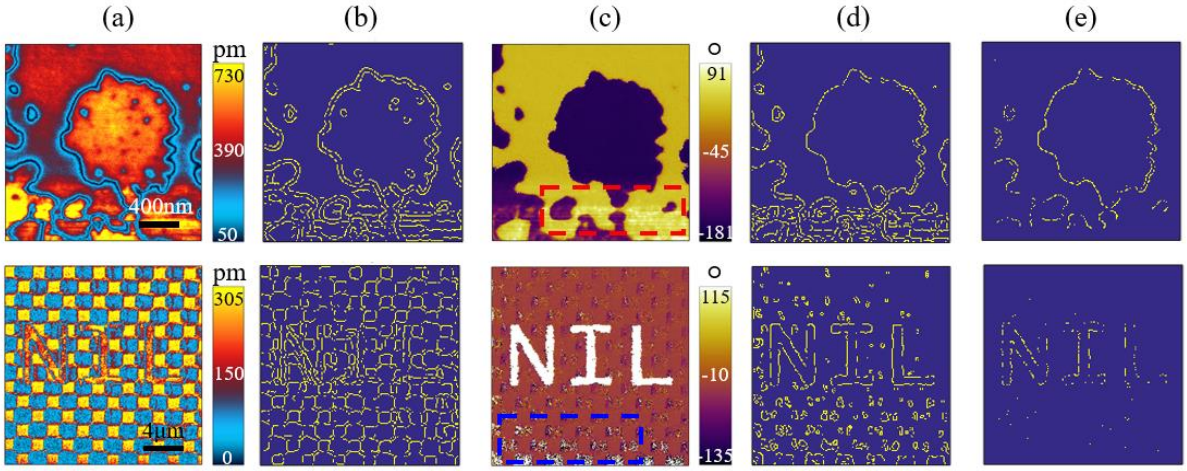

Figure 5.10. Failure of edge-detector algorithm. Comparison between the DWs identified by Canny edge detector and AI algorithms. Real PFM maps are columns (a) Amplitude and (c) Phase, the corresponding edges detected by Canny are shown in columns (b) and (d), separately; (e) DWs identified by AI algorithms.

Our ultimate goal is to demonstrate an AI-SPM platform that is capable of not only classifying ferroelectric materials with 180^o DWs in real time, but also adopt adaptive experimentation on the fly to probe the characteristics and mechanisms of apparent piezoresponse in details at critical material interface, i.e. DWs in ferroelectrics and GBs in electrochemical materials. Such additional probing is necessary to confirm the classification without ambiguity, a common practice by human users. To this end, two blind experiments have been conducted on ferroelectric PMN-Pt single crystal and electrochemical Ceria ceramic, both of which unknown to the AI-SPM in advance, and the screenshot of the complete experimental processes are recorded. As is shown in Mov. S1 on this reference [150], PMN-Pt was determined to be ferroelectric having 180^o DWs during scanning, and representative amplitude, phase, and topography (overlaid with identified DWs) mappings are shown in Figure 5.11(ab). After such preliminary classification a “ferroelectric routine” was triggered, with the scanning probe zoomed in on an identified DW and carrying out switching

spectroscopy PFM (SS-PFM) experiments [152] on a line of points across DW, yielding hysteresis and butterfly loops of Figure 5.11(c) characteristic for ferroelectric materials, and thus confirm the ferroelectric classification without ambiguity. When Ceria was tested, as shown in Mov. S2 on the reference [150], the AI-SPM found no 180^{o} DWs from its amplitude and phase mappings in Figure 5.11(d), and thus a "non-ferroelectric routine" was triggered to identify GBs overlaid on topography in Figure 5.11(e), after which the scanning probe was zoomed in on an identified GB and carried out measurements of first and second harmonic piezoresponse across GBs [89,90].

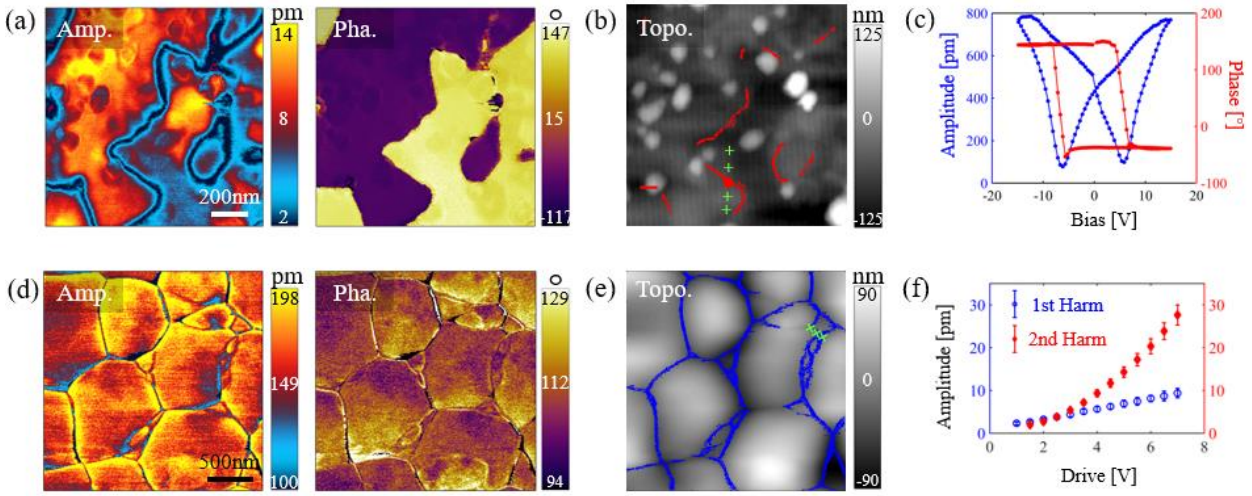

Figure 5.11. Demonstration of AI-SPM for two "unknown" samples that are determined to be ferroelectric (a-c) and electrochemical (d-f); (ad) mappings of amplitude and phase; (be) DWs and GBs recognized; and (cf) ferroelectric hysteresis and butterfly loops on DW and comparison of first and second harmonic piezoresponse on GB, all measured on the fly.

As seen in Figure 5.11(f), second harmonic piezoresponse dominates the first harmonic one in Ceria, characteristic of electrochemical materials, and thus confirm its non-ferroelectric nature without ambiguity. As is clear from the movies, both experiments have been conducted in artificial intelligence manner without human users' interference, demonstrating the capability of our AI-

SPM system. Note that the probed sample could be neither ferroelectric nor electrochemical, which can be revealed by the characteristics of first and second harmonic piezoresponse across grain boundaries.

5.5 CONCLUSIONS

What we demonstrate here is a simple yet powerful prototype artificial intelligence SPM that is trained to carry out complicated scientific experiments from beginning to end, all on its own, and it is just matter of time to incorporate more profound physical processes and more sophisticated deep learning algorithm to expand its power. Indeed, our current implementation is based on interfacial profile of vertical piezoresponse, which could be complemented by additional lateral probing for possible in-plane polarization, as well as pre-probing lithography for single-domain ferroelectrics. Classifying atypical polar materials such as surface polar phases [96], mixed polar/nonpolar-phases [153], and highly fragmental, curved and fine-scale domains in relaxors [154] could impose additional challenges, but there is no fundamental difficulty. We also ignored possible change of surface condition and wearing of probe during scanning, which could be addressed by auto-tuning the probe parameters with the help of advanced convolutional neural network. We acknowledge that the DWs and GBs recognized by the current AI implementation is not really elusive per se, but the strategy points toward a viable direction toward more subtle features that is easy to be overlooked. We expect that similar strategy can also be developed for a wide range of scientific instruments from transmission electron microscope to X-ray diffractometer, as well as a broader physical machineries and systems that heavily rely on human experience to operate at the moment. It is also conceivable that an eco-system can emerge from such vision that all the AI-SPM are interconnected to share and strengthen training data, machine learning algorithm, as well as control, so that user experiences and know-how are no longer limited

in a particular lab, but readily spread over the network, and we have made our algorithms publicly accessible to facilitate such movement [155]. More importantly, a general intelligent machine may evolve from such eco-system, which is capable of all SPM experiments and analysis on its own, revolutionizing the way we do SPM experiments.

Chapter 6. CONCLUSIONS AND FUTURE WORK

6.1 CONCLUSIONS

The work discussed in this dissertation mainly focus on development of scanning probe microscopy methodologies with respect to data acquisition, data analysis and data reuse for the study of advanced functional materials.

The needs for a data-driven SPM is well illustrated in Chapter 2 by halide perovskite $CH_3NH_3PbI_3$ which is widely pursued for the next generation solar cells. The microscopic processes in $CH_3NH_3PbI_3$ are highly complex, involving generation and transport of photo-carriers, ionic migration and defects, and possible polarization domain and switching, which all interact with light, heat, and electric field on various time and length scales. In order to understand the effects of such microscopic processes on photovoltaic conversion, detailed local distributions have been mapped out, including microstructure morphologies, surface potential, photocurrent, piezoresponse, chemical degradation, and ionic activities. We find that single crystalline $CH_3NH_3PbI_3$ films possess ferroic domains with alternating instead of pure polar and nonpolar orders, and that polar domains exhibit reduced photocurrent in contrast to previous theoretical expectations. Although we manage to obtain as much information as possible using a conventional SPM, the data collected is still prone to errors due to drifting and elusive due to its complexity and the measurements are tedious. These issues are not limited to perovskite solar cells, and are actually quite common for advanced materials and other complex systems.

Thus, we seek to design a data-driven SPM integrating big data analytics to overcome above difficulties. As is introduced in Chapter 3, we first developed a sequential excitation scanning probe microscopy (SE-SPM) technique that excites a rough PZT ceramic using a sequence of AC waveforms with distinct frequencies ω, wherein the excitation energy is concentrated on only one

frequency at a time instead of being distributed over a band of spectrum. No resonance tracking is needed as in DART, ensuring that the measurement is robust and reliable. Besides, this approach requires no extra hardware and instrumentation, making it widely accessible for any standard SPM. Although SE can obtain multi-dimensional dataset, the required continuous scanning is not appropriate for sensitive materials like $CH_3NH_3PbI_3$. As mentioned in Chapter 4, we upgrade it to the high-throughput SE which is capable of completing each round of sequential excitation exactly in the level of a single pixel rather than the whole map, requiring only fast one-scan to collect the same dataset and minimizing the damage to the sample. This new method can scan at a speed comparable to DART, while keeping much higher spatial as well as spectral resolution due to a record of full time-domain responses. The obtained 3D data sets of amplitude $A(\omega, x, y)$ and phase $\phi(\omega, x, y)$ are used to extract intrinsic properties A_0, ω_0, and Q of local area via the DHO model. Multivariate statistical tool PCA is also utilized to facilitate data processing and a mechanistic interpretation of the unsupervised PCA is acquired through DHO model.

In Chapter 5, we propose a data-driven AI-SPM experimentation based on machine learning and ever-growing training data, as operating SPM is tedious at times for experienced users. Support vector machine is first introduced and then adopted as a core algorithm for the AI-SPM system. Compared with convolutional neural network, SVM requires less computational power and is more efficient for pixel-by-pixel recognition, which is critical for autonomously adaptive experimentation on the fly. The SVM-based AI is examined by 7174 ds-SPM mappings accumulated in our lab and exhibits a true positive rate of 97.3% and a high true negative rate of 99.6%. We also conduct two blind experiments on ferroelectric PMN-Pt single crystal and electrochemical Ceria ceramic, both of which is unknown to the AI-SPM in advance. The system succeeds in labeling materials to particular classes, identifying relevant features such as domain

walls or grain boundaries, and carrying out SS-PFM or multi-harmonic measurements on these critical features in real time.

6.2 FUTURE WORK

Multi-modality imaging offered by SPM has shed considerable insight into local photovoltaic characteristics of perovskite solar cells. But each of the techniques only focuses on a particular aspect of the local heterogeneity and inhomogeneous response, such as cAFM for photocurrent, SKPM for surface potential, and PFM for mechanical displacement. Any of them alone would be insufficient to investigate the dominant mechanisms inside materials, thus multiple techniques must work in concert on the same scanned area. However, looking for the same microscopic area after switching among different SPM modes is like finding a needle in a forest. Therefore, we are exploring the possibility of integrating two or more modes into one single experimental setup, for example combing SE-PFM and SE-pcAFM. As both two modes utilize conductive probes to apply bias and difference only exists on the acquired information, capability and flexibility of measurement can be greatly enhanced once this goal is achieved, for example direct observation as well as analysis on the coupling between photocurrent and polarization switching is possible without annoying area searching. Currently, technique issues such as implementations of appropriate external current amplifier and electromagnetic shielding are being solved.

Besides, it is worth to apply high-throughput SE to pcAFM. In Chapter 3, in order to achieve local IV curves of high spatial grids, sequence of images have to be acquired under different DC biases. The yielded dataset has turned out to be useful for the estimation of microscopic photovoltaic characteristics such as the open-circuit photovoltage and short-circuit photocurrent density, but a faster test is always desired for sensitive materials and users. This would require

further development of hardware instrumentation and analytic software, as the time constant of the measured circuit and compensation capacitor need to be well tuned.

AI-SPM discussed in Chapter 5 has shown promising applications in classifying and determining material properties automatically. However, it is limited to specific modes in relatively stable measurement conditions. Considering that the quality of mapping is readily impaired due to artifacts on the sample, abrasion of probe, and thermal disturbance, we plan to strengthen the AI-SPM system for more general operation. On the one hand, we can train the AI to detect and avoid those situations in real-time by employing advanced neural network frameworks for the recognition of unexpected impurities and defects. On the other hand, when the condition is not serious so that it cannot be identified by AI, we can try another deep learning technique Generative Adversarial Network (GAN) for post-processing, which has proven to be powerful in image inpainting. Therefore, mappings with good quality are likely to be obtained from the AI-SPM.

We believe these directions are valuable for the deep study of many advanced functional materials, and will help to speed up their commercial applications in the future.

BIBLIOGRAPHY

[1] Gocha A. Smart Materials Make Smartphones. Am Ceram Soc Bull 2018;97:8–23.

[2] Hong CH, Shin JH, Ju BK, Kim KH, Park NM, Kim BS, et al. Index-matched indium tin oxide electrodes for capacitive touch screen panel applications. J Nanosci Nanotechnol 2013. https://doi.org/10.1166/jnn.2013.7814.

[3] Kao S-H, Tsao M-F, Chiang C-W, Chao M-C. A miniature TCXO for GPS/GNSS application. Proc. 2014 Symp. Piezoelectricity, Acoust. Waves, Device Appl., IEEE; 2014, p. 138–41.

[4] Walton D. Specialty glass: A new design element in consumer electronics. Electron Des 2010.

[5] Young R. OLEDs expected to gain ground as LCD investment slows. Inf Disp (1975) 2019. https://doi.org/10.1002/msid.1035.

[6] Versus. https://versus.com/en/apple-a13-bionic-vs-apple-a4

[7] Bonnell DA, Kalinin S V. Scanning Probe Microscopy for Energy Research. Am Ceram Soc Bull 2013;10.1142/86:40–1.

[8] Berhe TA, Su W-N, Chen C-H, Pan C-J, Cheng J-H, Chen H-M, et al. Organometal halide perovskite solar cells: degradation and stability. Energy Environ Sci 2016;9:323–56. https://doi.org/10.1039/C5EE02733K.

[9] Uddin A, Upama MB, Yi H, Duan L. Encapsulation of organic and perovskite solar cells: A review. Coatings 2019. https://doi.org/10.3390/coatings9020065.

[10] Binnig G, Quate CF, Gerber C. Atomic force microscope. Phys Rev Lett 1986;56:930. https://doi.org/10.1103/PhysRevLett.56.930.

[11] Kalinin S V., Balke N. Local electrochemical functionality in energy storage materials and devices by scanning probe microscopies: Status and perspectives. Adv Mater 2010. https://doi.org/10.1002/adma.201001190.

[12] Li J, Li J-F, Yu Q, Chen QN, Xie S. Strain-based scanning probe microscopies for functional materials, biological structures, and electrochemical systems. J Mater 2015;1:3–21. https://doi.org/10.1016/j.jmat.2015.03.001.

[13] A schematic illustration of SPM https://myscope.training/legacy/spm/background/.

[14] Inc. AR. Imaging and Spectroscopy Applications Guide 2013.

[15] Piezoelectric sensorics: force, strain, pressure, acceleration and acoustic emission sensors, materials and amplifiers. Choice Rev Online 2002. https://doi.org/10.5860/choice.40-0924.

[16] Güthner P, Dransfeld K. Local poling of ferroelectric polymers by scanning force microscopy. Appl Phys Lett 1992;61:1137–9. https://doi.org/10.1063/1.107693.

[17] Kalinin S V., Rodriguez BJ, Jesse S, Shin J, Baddorf AP, Gupta P, et al. Vector Piezoresponse Force Microscopy. Microsc Microanal 2006;12:206–20. https://doi.org/10.1017/S1431927606060156.

[18] French AP. Vibrations and waves. CRC press; 1971.

[19] Fujihira M. Kelvin probe force microscopy of molecular surfaces. 1999. https://doi.org/10.1146/annurev.matsci.29.1.353.

[20] Nonnenmacher M, O'Boyle MP, Wickramasinghe HK. Kelvin probe force microscopy. Appl Phys Lett 1991. https://doi.org/10.1063/1.105227.

[21] Murrell MP, Welland ME, O'Shea SJ, Wong TMH, Barnes JR, McKinnon AW, et al.

Spatially resolved electrical measurements of SiO2 gate oxides using atomic force microscopy. Appl Phys Lett 1993. https://doi.org/10.1063/1.108579.

[22] Lanza M. Conductive Atomic Force Microscopy: Applications in Nanomaterials. John Wiley & Sons; 2017.

[23] Schematic figure showing the setup of a C-AFM https://www.ma-tek.com/en-global/services/index/Pro_category_02/Pro_08.

[24] Kojima A, Teshima K, Shirai Y, Miyasaka T. Organometal halide perovskites as visible-light sensitizers for photovoltaic cells. J Am Chem Soc 2009;131:6050–1. https://doi.org/10.1021/ja809598r.

[25] Lee MM, Teuscher J, Miyasaka T, Murakami TN, Snaith HJ. Efficient hybrid solar cells based on meso-superstructured organometal halide perovskites. Science (80-) 2012;338:643–7. https://doi.org/10.1126/science.1228604 [doi].

[26] Burschka J, Pellet N, Moon S-JJ, Humphry-Baker R, Gao P, Nazeeruddin MK, et al. Sequential deposition as a route to high-performance perovskite-sensitized solar cells. Nature 2013;499:316–9. https://doi.org/10.1038/nature12340.

[27] Huanping Z, Qi C, Gang L, Song L, Tze-bing S, Hsin-Sheng D, et al. Interface engineering of highly efficient perovskite solar cells. Science (80-) 2014;345:542–6. https://doi.org/10.1126/science.1254050.

[28] Chen W, Wu Y, Yue Y, Liu J, Zhang W, Yang X, et al. Efficient and stable large-area perovskite solar cells with inorganic charge extraction layers. Science (80-) 2015;350:944–8. https://doi.org/10.1126/science.aad1015.

[29] Dong Q, Fang Y, Shao Y, Mulligan P, Qiu J, Cao L, et al. Solar cells. Electron-hole diffusion lengths >175 μm in solution-grown CH3NH3PbI3 single crystals. Science (80-) 2015;347:967–70. https://doi.org/10.1126/science.aaa5760.

[30] Yang WS, Park B-W, Jung EH, Jeon NJ, Kim YC, Lee DU, et al. Iodide management in formamidinium-lead-halide–based perovskite layers for efficient solar cells. Science (80-) 2017;356:1376–9. https://doi.org/10.1126/science.aan2301.

[31] Eames C, Frost JM, Barnes PRF, O'Regan BC, Walsh A, Islam MS. Ionic transport in hybrid lead iodide perovskite solar cells. Nat Commun 2015;6:7497. https://doi.org/10.1038/ncomms8497.

[32] Marchioro A, Teuscher J, Friedrich D, Kunst M, Van De Krol R, Moehl T, et al. Unravelling the mechanism of photoinduced charge transfer processes in lead iodide perovskite solar cells. Nat Photonics 2014;8:250–5. https://doi.org/10.1038/nphoton.2013.374.

[33] Sherkar TS, Momblona C, Gil-Escrig L, Ávila J, Sessolo M, Bolink HJ, et al. Recombination in Perovskite Solar Cells: Significance of Grain Boundaries, Interface Traps, and Defect Ions. ACS Energy Lett 2017;2:1214–22. https://doi.org/10.1021/acsenergylett.7b00236.

[34] Wang P, Zhao J, Wei L, Zhu Q, Xie S, Liu J, et al. Photo-induced Ferroelectric Switching in Perovskite CH3NH3PbI3 Films. Nanoscale 2017;9:3806–17. https://doi.org/10.1039/C6NR09310H.

[35] Pazoki M, Jacobsson TJ, Kullgren J, Johansson EMJ, Hagfeldt A, Boschloo G, et al. Photoinduced Stark Effects and Mechanism of Ion Displacement in Perovskite Solar Cell Materials. ACS Nano 2017;11:2823–34. https://doi.org/10.1021/acsnano.6b07916.

[36] Agrawal A, Choudhary A. Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. APL Mater 2016;4:53208.

https://doi.org/10.1063/1.4946894.
[37] Kalinin S V., Sumpter BG, Archibald RK. Big-deep-smart data in imaging for guiding materials design. Nat Mater 2015;14:973–80. https://doi.org/10.1038/nmat4395.
[38] Munevar S. Unlocking Big Data for better health. Nat Biotechnol 2017;35:684.
[39] Collins L, Belianinov A, Proksch R, Zuo T, Zhang Y, Liaw PK, et al. G-mode magnetic force microscopy: Separating magnetic and electrostatic interactions using big data analytics. Appl Phys Lett 2016;108:193103. https://doi.org/10.1063/1.4948601.
[40] Accelerated Discovery of Large Electrostrains in BaTiO3 - Based Piezoelectrics Using Active Learning 2018. https://doi.org/10.1002/adma.201702884.
[41] Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 2018;172:1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010.
[42] Belianinov A, Kalinin S V., Jesse S. Complete information acquisition in dynamic force microscopy. Nat Commun 2015;6:6550. https://doi.org/10.1038/ncomms7550.
[43] Hill J, Mulholland G, Persson K, Seshadri R, Wolverton C, Meredig B. Materials science with large-scale data and informatics: Unlocking new opportunities. MRS Bull 2016;41:399–409. https://doi.org/DOI: 10.1557/mrs.2016.93.
[44] Kalinin S V, Strelcov E, Belianinov A, Somnath S, Vasudevan RK, Lingerfelt EJ, et al. Big, Deep, and Smart Data in Scanning Probe Microscopy. ACS Nano 2016;10:9068–86. https://doi.org/10.1021/acsnano.6b04212.
[45] Collins L, Ahmad M, Wu T, Hu B, Kalinin S V, Jesse S, et al. Breaking the time barrier in Kelvin probe force microscopy: fast free force reconstruction using the G-Mode platform. ACS Nano 2017;11:8717–29. https://doi.org/10.1021/acsnano.7b02114.
[46] Liu Y, Zhao T, Ju W, Shi S. Materials discovery and design using machine learning. J Mater 2017;3:159–77. https://doi.org/10.1016/j.jmat.2017.08.002.
[47] Byers J. The physics of data. Nat Phys 2017;13:718.
[48] Rodriguez BJ, Callahan C, Kalinin S V, Proksch R. Dual-frequency resonance-tracking atomic force microscopy. Nanotechnology 2007;18:475504. https://doi.org/10.1088/0957-4484/18/47/475504.
[49] Green MA, Dunlop ED, Hohl-Ebinger J, Yoshita M, Kopidakis N, Ho-Baillie AWY. Solar cell efficiency tables (Version 55). Prog Photovoltaics Res Appl 2019;28.
[50] Best Research-Cell Efficiency Chart https://www.nrel.gov/pv/cell-efficiency.html.
[51] Kutes Y, Zhou Y, Bosse JL, Steffes J, Padture NP, Huey BD. Mapping the Photoresponse of CH3NH3PbI3 Hybrid Perovskite Thin Films at the Nanoscale. Nano Lett 2016;16:3434–41. https://doi.org/10.1021/acs.nanolett.5b04157.
[52] Shao Y, Fang Y, Li T, Wang Q, Dong Q, Deng Y, et al. Grain boundary dominated ion migration in polycrystalline organic–inorganic halide perovskite films. Energy Environ Sci 2016;501:395–8. https://doi.org/10.1039/C6EE00413J.
[53] Dymshits A, Henning A, Segev G, Rosenwaks Y, Etgar L. The electronic structure of metal oxide/organo metal halide perovskite junctions in perovskite based solar cells. Sci Rep 2015;5:8704. https://doi.org/10.1038/srep08704.
[54] Zhu H, Huang B, Wu S, Xiong Z, Li J, Chen W. Facile Surface Modification upon CH3NH3PbI3 Films Leading to Simultaneously Improved Efficiency and Stability of Inverted Perovskite Solar Cells. J Mater Chem A 2018;6. https://doi.org/10.1039/C8TA00267C.
[55] Tagantsev AK, Cross LE, Fousek J. Domains in ferroic crystals and thin films. Springer;

2010.
[56] Frost JM, Butler KT, Brivio F, Hendon CH, Van Schilfgaarde M, Walsh A. Atomistic origins of high-performance in hybrid halide perovskite solar cells. Nano Lett 2014;14:2584–90. https://doi.org/10.1021/nl500390f.
[57] Liu S, Zheng F, Koocher NZ, Takenaka H, Wang F, Rappe AM. Ferroelectric domain wall induced band gap reduction and charge separation in organometal halide perovskites. J Phys Chem Lett 2015;6:693–9. https://doi.org/10.1021/jz502666j.
[58] Sherkar TS, Jan Anton Koster L. Can ferroelectric polarization explain the high performance of hybrid halide perovskite solar cells? Phys Chem Chem Phys 2016;18:331–8. https://doi.org/10.1039/c5cp07117h.
[59] Salje EKH. Ferroelastic Materials. Annu Rev Mater Res 2012. https://doi.org/10.1146/annurev-matsci-070511-155022.
[60] Stoumpos CC, Malliakas CD, Kanatzidis MG. Semiconducting Tin and Lead Iodide Perovskites with Organic Cations: Phase Transitions, High Mobilities, and Near-Infrared Photoluminescent Properties. Inorg Chem 2013;52:9019–38. https://doi.org/dx.doi.org/10.1021/ic401215x.
[61] Dang Y, Liu Y, Sun Y, Yuan D, Liu X, Lu W, et al. Bulk crystal growth of hybrid perovskite material $CH_3NH_3PbI_3$. CrystEngComm 2015;17:665–70. https://doi.org/10.1039/C4CE02106A.
[62] Baikie T, Barrow NS, Fang Y, Keenan PJ, Slater PR, Piltz RO, et al. A combined single crystal neutron/X-ray diffraction and solid-state nuclear magnetic resonance study of the hybrid perovskites $CH_3NH_3PbX_3$ (X = I, Br and Cl). J Mater Chem A 2015;3:9298–307. https://doi.org/10.1039/C5TA01125F.
[63] Baikie T, Fang Y, Kadro JM, Schreyer M, Wei F, Mhaisalkar SG, et al. Synthesis and crystal chemistry of the hybrid perovskite (CH3NH3)PbI3 for solid-state sensitised solar cell applications. J Mater Chem A 2013;1:5628–41. https://doi.org/10.1039/c3ta10518k.
[64] Weller MT, Weber OJ, Henry PF, Di Pumpo AM, Hansen TC. Complete structure and cation orientation in the perovskite photovoltaic methylammonium lead iodide between 100 and 352 K. Chem Commun 2015;51:4180–3. https://doi.org/10.1039/C4CC09944C.
[65] Fang HH, Raissa R, Abdu-Aguye M, Adjokatse S, Blake GR, Even J, et al. Photophysics of organic-inorganic hybrid lead iodide perovskite single crystals. Adv Funct Mater 2015;25:2378–85. https://doi.org/10.1002/adfm.201404421.
[66] Whitfield PS, Herron N, Guise WE, Page K, Cheng YQ, Milas I, et al. Structures, Phase Transitions and Tricritical Behavior of the Hybrid Perovskite Methyl Ammonium Lead Iodide. Sci Rep 2016;6:35685. https://doi.org/10.1038/srep35685.
[67] Sewvandi GA, Kodera K, Ma H, Nakanishi S, Feng Q. Antiferroelectric Nature of CH3NH3PbI3-xClx Perovskite and Its Implication for Charge Separation in Perovskite Solar Cells. Sci Rep 2016;6:30680. https://doi.org/10.1038/srep30680.
[68] Rakita Y, Bar-Elli O, Meirzadeh E, Kaslasi H, Peleg Y, Hodes G, et al. Tetragonal CH3NH3PbI3 is ferroelectric. Proc Natl Acad Sci 2017;114:E5504–12. https://doi.org/10.1073/pnas.1702429114.
[69] Sharada G, Mahale P, Kore BP, Mukherjee S, Pavan MS, De C, et al. Is CH3NH3PbI3 Polar? J Phys Chem Lett 2016;7:2412–9. https://doi.org/10.1021/acs.jpclett.6b00803.
[70] Beilsten-Edmands J, Eperon GE, Johnson RD, Snaith HJ, Radaelli PG. Non-ferroelectric nature of the conductance hysteresis in CH3NH3PbI3 perovskite-based photovoltaic devices. Appl Phys Lett 2015;106:173502. https://doi.org/10.1063/1.4919109.

[71] Hoque MNF, Yang M, Li Z, Islam N, Pan X, Zhu K, et al. Polarization and Dielectric Study of Methylammonium Lead Iodide Thin Film to Reveal its Nonferroelectric Nature under Solar Cell Operating Conditions. ACS Energy Lett 2016;1:142–9. https://doi.org/10.1021/acsenergylett.6b00093.

[72] Kim H, Kim SK, Kim BJ, Shin K, Gupta MK, Jung HS, et al. Ferroelectric Polarization in CH3NH3PbI3 Perovskite. J Phys Chem Lett 2015;6:1729–35. https://doi.org/10.1021/acs.jpclett.5b00695.

[73] Kutes Y, Ye L, Zhou Y, Pang S, Huey BD, Padture NP. Direct Observation of Ferroelectric Domains in Solution-Processed. J Phys Chem Lett 2014;5:3335–9. https://doi.org/10.1021/jz501697b.

[74] Chen B, Shi J, Zheng X, Zhou Y, Zhu K, Priya S. Ferroelectric solar cells based on inorganic–organic hybrid perovskites. J Mater Chem A 2015;3:7699–705. https://doi.org/10.1039/C5TA01325A.

[75] Kim Y-J, Dang T-V, Choi H-J, Park B-J, Eom J-H, Song H-A, et al. Piezoelectric properties of CH3NH3PbI3 perovskite thin films and their applications in piezoelectric generators. J Mater Chem A 2016;4:756–63. https://doi.org/10.1039/C5TA09662F.

[76] Hermes IM, Bretschneider SA, Bergmann VW, Li D, Klasen A, Mars J, et al. Ferroelastic Fingerprints in Methylammonium Lead Iodide Perovskite. J Phys Chem C 2016;120:5724–31. https://doi.org/10.1021/acs.jpcc.5b11469.

[77] Röhm H, Leonhard T, Hoffmann MJ, Colsmann A. Ferroelectric domains in methylammonium lead iodide perovskite thin-films. Energy Environ Sci 2017;10:950–5. https://doi.org/10.1039/C7EE00420F.

[78] Strelcov E, Dong Q, Li T, Chae J, Shao Y, Deng Y, et al. CH 3 NH 3 PbI 3 perovskites : Ferroelasticity revealed. Sci Adv 2017;3:e1602165. https://doi.org/10.1126/sciadv.1602165.

[79] Coll M, Gomez A, Mas-Marza E, Almora O, Garcia-Belmonte G, Campoy-Quiles M, et al. Polarization switching and light-enhanced piezoelectricity in lead halide perovskites. J Phys Chem Lett 2015;6:1408–13. https://doi.org/10.1021/acs.jpclett.5b00502.

[80] Fan Z, Xiao J, Sun K, Chen L, Hu Y, Ouyang J, et al. Ferroelectricity of CH3NH3PbI3 perovskite. J Phys Chem Lett 2015;6:1155–61. https://doi.org/10.1021/acs.jpclett.5b00389.

[81] Frost JM, Butler KT, Brivio F, Hendon CH, Van Schilfgaarde M, Walsh A. Atomistic origins of high-performance in hybrid halide perovskite solar cells. Nano Lett 2014;14:2584–90. https://doi.org/10.1021/nl500390f.

[82] Quarti C, Mosconi E, De Angelis F. Interplay of orientational order and electronic structure in methylammonium lead iodide: Implications for solar cell operation. Chem Mater 2014;26:6557–69. https://doi.org/10.1021/cm5032046.

[83] Stroppa A, Quarti C, De Angelis F, Picozzi S. Ferroelectric Polarization of $CH_3NH_3PbI_3$: A Detailed Study Based on Density Functional Theory and Symmetry Mode Analysis. J Phys Chem Lett 2015;6:2223–31. https://doi.org/10.1021/acs.jpclett.5b00542.

[84] Zheng F, Takenaka H, Wang F, Koocher NZ, Rappe AM. First-Principles Calculation of the Bulk Photovoltaic Effect in $CH_3NH_3PbI_3$ and $CH_3NH_3PbI_{3-x}Cl_x$. J Phys Chem Lett 2015;6:31–7. https://doi.org/10.1021/jz502109e.

[85] Yin W-J, Yang J-H, Kang J, Yan Y, Wei S-H. Halide perovskite materials for solar cells: a theoretical review. J Mater Chem A 2015;3:8926–42. https://doi.org/10.1039/C4TA05033A.

[86] Stroppa A, Quarti C, De Angelis F, Picozzi S. Ferroelectric Polarization of CH3NH3PbI3: A Detailed Study Based on Density Functional Theory and Symmetry Mode Analysis. J Phys Chem Lett 2015:2223–31. https://doi.org/10.1021/acs.jpclett.5b00542.
[87] Zhao J, Kong G, Chen S, Li Q, Huang B, Liu Z, et al. Single crystalline CH 3 NH 3 PbI 3 self-grown on FTO/TiO 2 substrate for high efficiency perovskite solar cells. Sci Bull 2017;62:1173–6. https://doi.org/10.1016/j.scib.2017.08.022.
[88] Gannepalli A, Yablon DG, Tsou AH, Proksch R. Mapping nanoscale elasticity and dissipation using dual frequency contact resonance AFM. Nanotechnology 2011;22:355705. https://doi.org/10.1088/0957-4484/22/35/355705.
[89] Nataly Chen Q, Ou Y, Ma F, Li J, Chen QN, Ou Y, et al. Mechanisms of electromechanical coupling in strain based scanning probe microscopy. Appl Phys Lett 2014;242907:1–5. https://doi.org/10.1063/1.4884422.
[90] Yu J, Esfahani EN, Zhu Q, Shan D, Jia T, Xie S, et al. Quadratic electromechanical strain in silicon investigated by scanning probe microscopy. J Appl Phys 2018;123:155104.
[91] Rothmann MU, Li W, Zhu Y, Bach U, Spiccia L, Etheridge J, et al. Direct observation of intrinsic twin domains in tetragonal CH3NH3PbI3. Nat Commun 2017;8:14547. https://doi.org/10.1038/ncomms14547.
[92] Bhattacharya K. Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect. vol. 2. Oxford University Press; 2003.
[93] Eerenstein W, Mathur ND, Scott JF. Multiferroic and magnetoelectric materials. Nature 2006;442:759–65.
[94] Salje EK. Phase Transitions in Ferroelastic and Co-elastic Crystals. Cambridge University Press; 1991.
[95] Vasudevan RK, Balke N, Maksymovych P, Jesse S, Kalinin S V. Ferroelectric or non-ferroelectric: Why so many materials exhibit “ferroelectricity” on the nanoscale. Appl Phys Rev 2017;4:021302. https://doi.org/10.1063/1.4979015.
[96] Yang SM, Morozovska AN, Kumar R, Eliseev EA, Cao Y, Mazet L, et al. Mixed electrochemical–ferroelectric states in nanoscale ferroelectrics. Nat Phys 2017;13:812–818. https://doi.org/10.1038/NPHYS4103.
[97] Bonnell DA, Kalinin SV, Kholkin AL, Gruverman A. Piezoresponse Force Microscopy: A Window into Electromechanical Behavior at the Nanoscale. MRS Bull 2009;34:648–57. https://doi.org/10.1557/mrs2009.176.
[98] Xie S, Gannepalli A, Chen QN, Liu Y, Zhou Y, Proksch R, et al. High resolution quantitative piezoresponse force microscopy of $BiFeO_3$ nanofibers with dramatically enhanced sensitivity. Nanoscale 2012;4:408–13. https://doi.org/10.1039/C1NR11099C.
[99] Yuan Y, Huang J. Ion Migration in Organometal Trihalide Perovskite and Its Impact on Photovoltaic Efficiency and Stability. Acc Chem Res 2016;49:286–93. https://doi.org/10.1021/acs.accounts.5b00420.
[100] Li C, Tscheuschner S, Paulus F, Hopkinson PE, Kie?ling J, K?hler A, et al. Iodine Migration and its Effect on Hysteresis in Perovskite Solar Cells. Adv Mater 2016;28:2446–54. https://doi.org/10.1002/adma.201503832.
[101] Kumar A, Jesse S, Morozovska AN, Eliseev E, Tebano A, Yang N, et al. Variable temperature electrochemical strain microscopy of Sm-doped ceria. Nanotechnology 2013;24:145401. https://doi.org/10.1088/0957-4484/24/14/145401.
[102] Xie S, Gannepalli A, Chen QN, Liu Y, Zhou Y, Proksch R, et al. High resolution quantitative piezoresponse force microscopy of BiFeO3 nanofibers with dramatically

enhanced sensitivity. Nanoscale 2012;4:408–13. https://doi.org/10.1039/C1NR11099C.
[103] Li T, Chen L, Zeng K. In situ studies of nanoscale electromechanical behavior of nacre under flexural stresses using band excitation PFM. Acta Biomater 2013;9:5903–12. https://doi.org/https://doi.org/10.1016/j.actbio.2013.01.003.
[104] Rodriguez SJ and SVK and RP and APB and BJ. The band excitation method in scanning probe microscopy for rapid mapping of energy dissipation on the nanoscale. Nanotechnology 2007;18:435503.
[105] Collins L, Belianinov A, Somnath S, Balke N, Kalinin S V., Jesse S. Full data acquisition in Kelvin Probe Force Microscopy: Mapping dynamic electric phenomena in real space. Sci Rep 2016;6:30557. https://doi.org/10.1038/srep30557.
[106] Huang B, Esfahani EN, Li J. Mapping intrinsic electromechanical responses at the nanoscale via sequential excitation scanning probe microscopy empowered by deep data. Natl Sci Rev 2018;6:55–63. https://doi.org/10.1093/nsr/nwy096.
[107] Sotomayor M, Schulten K, Evans E, Bell GI, Davis MM, Evans E, et al. Far-Field Optical Nanoscopy 2007;2:1153–9. https://doi.org/10.1126/science.1137395.
[108] Balzarotti F, Eilers Y, Gwosch KC, Gynnå AH, Westphal V, Stefani FD, et al. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. Science (80-) 2017;355:606–12. https://doi.org/10.1126/science.aak9913.
[109] Guizar-Sicairos M, Thurman ST, Fienup JR. Efficient subpixel image registration algorithms. Opt Lett 2008;33:156–8.
[110] Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst 1987;2:37–52.
[111] Liberman MC, Gao J, He DZZ, Wu X, Jia S, Zuo J. Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. Nature 2002;419:300.
[112] Sachs F, Brownell WE, Petrov AG. Membrane electromechanics in biology, with a focus on hearing. MRS Bull 2009;34:665–70.
[113] Plonsey R, Barr RC. Bioelectricity: a quantitative approach. Springer Science & Business Media; 2007.
[114] Rodriguez BJ, Choudhury S, Chu YH, Bhattacharyya A, Jesse S, Seal K, et al. Unraveling Deterministic Mesoscopic Polarization Switching Mechanisms: Spatially Resolved Studies of a Tilt Grain Boundary in Bismuth Ferrite. Adv Funct Mater 2009;19:2053–63. https://doi.org/10.1002/adfm.200900100.
[115] Eshghinejad A, Nasr Esfahani E, Wang P, Xie S, Geary TC, Adler SB, et al. Scanning thermo-ionic microscopy for probing local electrochemistry at the nanoscale. J Appl Phys 2016;119:205110. https://doi.org/10.1063/1.4949473.
[116] Bradler S, Schirmeisen A, Roling B. Piezoresponse force and electrochemical strain microscopy in dual AC resonance tracking mode: Analysis of tracking errors. J Appl Phys 2018;123:035106. https://doi.org/10.1063/1.5004472.
[117] Golub GH, Reinsch C. Singular value decomposition and least squares solutions. Numer Math 1970;14:403–20.
[118] Greub WH. Linear algebra. vol. 23. Springer Science & Business Media; 2012.
[119] Li J, Huang B, Esfahani EN, Wei L, Yao J, Zhao J. Touching is believing : interrogating organometal halide perovskite solar cells at the nanoscale via scanning probe microscopy. Npj Quantum Mater 2016;2:1–19. https://doi.org/10.1038/s41535-017-0061-4.
[120] Xia G, Huang B, Zhang Y, Zhao X, Wang C, Jia C, et al. Nanoscale Insights into Photovoltaic Hysteresis in Triple-Cation Mixed-Halide Perovskite: Resolving the Role of

Polarization and Ionic Migration. Adv Mater 2019;31:1–9. https://doi.org/10.1002/adma.201902870.
[121] Nelson J. The Physics of Solar Cells. 2003. https://doi.org/10.1142/p276.
[122] Edri E, Kirmayer S, Cahen D, Hodes G. High Open-Circuit Voltage Solar Cells Based on Organic–Inorganic Lead Bromide Perovskite. J Phys Chem Lett ;4:897–902.
[123] Huang B, Esfahani EN, Yu J, Gerwe BS, Adler SB, Li J. High-Throughput sequential excitation for nanoscale mapping of electrochemical strain in granular ceria. Nanoscale 2019;11:23188–96. https://doi.org/10.1039/c9nr07438d.
[124] Chen QN, Adler SB, Li J. Imaging space charge regions in Sm-doped ceria using electrochemical strain microscopy. Appl Phys Lett 2014;105:1–5. https://doi.org/10.1063/1.4901102.
[125] Kim S, Maier J. On the conductivity mechanism of nanocrystalline ceria. J Electrochem Soc 2002;149:J73–83.
[126] Kalinin S V, Gruverman A. Scanning probe microscopy: electrical and electromechanical phenomena at the nanoscale. vol. 1. Springer Science & Business Media; 2007.
[127] Zhu J, Lu L, Zeng K. Nanoscale Mapping of Lithium-Ion Diffusion in a Cathode within an All-Solid-State Lithium-Ion Battery by Advanced Scanning Probe Microscopy Techniques. ACS Nano 2013;7:1666–75. https://doi.org/10.1021/nn305648j.
[128] Luchkin SY, Romanyuk K, Ivanov M, Kholkin AL. Li transport in fresh and aged LiMn2O4 cathodes via electrochemical strain microscopy. J Appl Phys 2015;118:72016.
[129] Nataly Chen Q, Liu Y, Liu Y, Xie S, Cao G, Li J. Delineating local electromigration for nanoscale probing of lithium ion intercalation and extraction by electrochemical strain microscopy. Appl Phys Lett 2012;101:63901.
[130] Armstrong CM, Hille B. Voltage-gated ion channels and electrical excitability. Neuron 1998;20:371–80.
[131] Kalinin S, Balke N, Jesse S, Tselev A, Kumar A, Arruda TM, et al. Li-ion dynamics and reactivity on the nanoscale. Mater Today 2011;14:548–58. https://doi.org/10.1016/S1369-7021(11)70280-2.
[132] Agilent Technologies. Fundamentals of Modal Testing. Appl Note 2000:1–56.
[133] Nikolai Chernov (2020). Circle Fit (Pratt method) (https://www.mathworks.com/matlabcentral/fileexchange/22643-circle-fit-pratt-method), MATLAB Central File Exchange. Retrieved April 21, 2020.
[134] Pratt V. Direct least-squares fitting of algebraic surfaces. Proc. 14th Annu. Conf. Comput. Graph. Interact. Tech. SIGGRAPH 1987, 1987. https://doi.org/10.1145/37401.37420.
[135] Trivedi H, Shvartsman V V., Medeiros MSA, Pullar RC, Lupascu DC. Sequential piezoresponse force microscopy and the “small-data” problem. Npj Comput Mater 2018;4:1–7. https://doi.org/10.1038/s41524-018-0084-9.
[136] Giridharagopal R, Precht JT, Jariwala S, Collins L, Jesse S, Kalinin S V., et al. Time-Resolved Electrical Scanning Probe Microscopy of Layered Perovskites Reveals Spatial Variations in Photoinduced Ionic and Electronic Carrier Motion. ACS Nano 2019;13:2812–21. https://doi.org/10.1021/acsnano.8b08390.
[137] Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of go without human knowledge. Nature 2017;550:354.
[138] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115.
[139] Li L, Yang Y, Zhang D, Ye Z, Jesse S, Kalinin S V, et al. Machine learning – enabled

identification of material phase transitions based on experimental data : Exploring collective dynamics in ferroelectric relaxors. Sci Adv 2018;4:eaap8672. https://doi.org/10.1126/sciadv.aap8672.
[140] Bartók AP, De S, Poelking C, Bernstein N, Kermode JR, Csányi G, et al. Machine learning unifies the modeling of materials and molecules. Sci Adv 2017;3:e1701816.
[141] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436.
[142] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. Adv Neural Inf Process Syst 2012:1097–105.
[143] Steinwart I, Christmann A. Support vector machines. Springer Science & Business Media; 2008.
[144] Decoste D, Schölkopf B. Training invariant support vector machines. Mach Learn 2002;46:161–90.
[145] Drucker H, Wu D, Vapnik VN. Support vector machines for spam categorization. IEEE Trans Neural Networks 1999;10:1048–54.
[146] Larhmam. SVM margin https://en.wikipedia.org/wiki/Support-vector_machine#/media/File:SVM_margin.png.
[147] Boyd S, Vandenberghe L. Convex Optimization. 2004. https://doi.org/10.1017/cbo9780511804441.
[148] Chih-Wei Hsu, Chih-Chung Chang and C-JL. A Practical Guide to Support Vector Classification. BJU Int 2008.
[149] Morozovska AN, Eliseev EA, Balke N, Kalinin S V. Local probing of ionic diffusion by electrochemical strain microscopy: Spatial resolution and signal formation mechanisms. J Appl Phys 2010;108:53712.
[150] Huang B, Li Z, Li J. An artificial intelligence atomic force microscope enabled by machine learning. Nanoscale 2018;10:21320–6. https://doi.org/10.1039/c8nr06734a.
[151] Canny J. A computational approach to edge detection. Readings Comput Vis 1987:184–203.
[152] Jesse S, Baddorf AP, Kalinin S V. Switching spectroscopy piezoresponse force microscopy of ferroelectric materials. Appl Phys Lett 2006;88:1–4. https://doi.org/10.1063/1.2172216.
[153] Huang B, Kong G, Esfahani EN, Chen S, Li Q, Yu J, et al. Ferroic domains regulate photocurrent in single-crystalline CH 3 NH 3 PbI 3 films self-grown on FTO/TiO2 substrate. Npj Quantum Mater 2018;3:30.
[154] Li Q, Liu Y, Withers RL, Wan Y, Li Z, Xu Z. Piezoresponse force microscopy studies on the domain structures and local switching behavior of Pb(In1/2Nb1/2)O3-Pb (Mg1/3Nb2/3)O3-PbTiO3 single crystals. J Appl Phys 2012;112:52006.
[155] https://github.com/Boyuan-H/AI-AFM

VITA

EDUCATION

University of Washington **Seattle, WA** **Jun. 2020**
- Ph.D. in Mechanical Engineering with a 3.8/4.0 GPA.
- UW CEI graduate fellowship.

Nanjing University **Nanjing, China** **Jun. 2016**
- Bachelor of Science in Physics.

RESEARCH PROJECTS

Developing data-driven scanning probe microscopy (SPM) **Feb. 2018 – Dec. 2019**
- Developed a microscope-based data infrastructure for materials discoveries, integrating an arbitrary wave generator and a data acquisition system to record dynamic response in time domain;
- Accelerated data analysis by orders of magnitude with domain-knowledge informed principal component analysis.

Inventing an artificial intelligence SPM (AI-SPM) **Dec. 2017 – Nov. 2018**
- Developed an AI-SPM for pattern recognition and feature identification in ferroelectric materials and electrochemical systems;
- The AI-SPM can respond to classification via adaptive experimentation with additional probing at critical domain walls and grain boundaries, all in real-time on the fly without human interference.

Studying perovskite solar cells via SPM **Oct. 2016 – Jun. 2018**
- Found $CH_3NH_3PbI_3$ crystals self-grown on FTO/TiO_2 substrate consist of ferroic domains with alternating polar and nonpolar orders, and polar domains possess reduced photocurrent.

Monte Carlo simulation for an antiferroelectric system **Oct. 2015 – Jun. 2016**
- Simulated the structural evolution of an antiferroelectric system on Linux clusters, based on the Landau phenomenological theory.

PUBLICATION

- J. Yu*, **B. Huang***, et al. " Resolving local dynamics of dual ions at the nanoscale in electrochemically active materials." ***Nano Energy***, 2020.
- G. Xia*, **B. Huang***, et al. "Nanoscale Insights into Photovoltaic Hysteresis in Triple‐Cation Mixed‐Halide Perovskite: Resolving the Role of Polarization and Ionic Migration." ***Advanced Materials***, 2019.
- **B. Huang**, et al. " High-throughput sequential excitation for nanoscale mapping of electrochemical strain in granular ceria." ***Nanoscale***, 2019.
- J. Peng*, **B. Huang***, et al. "Resolving Fine Electromechanical Structure of Collagen Fibrils via Sequential Excitation Piezoresponse Force Microscopy." ***Nanotechnology***, 2019.
- **B. Huang**, et al. "Artificial intelligent atomic force microscope enabled by machine learning." ***Nanoscale***, 2018.
- **B. Huang**, et al. "Mapping intrinsic electromechanical responses at the nanoscale via sequential excitation scanning probe microscopy empowered by deep data." ***National Science Review***, 2018.
- **B. Huang**, et al. "Ferroic domains regulate photocurrent in single-crystalline $CH_3NH_3PbI_3$ films self-grown on FTO/TiO_2 substrate." ***npj Quantum Materials***, 2018.
- H. Zhu, **B. Huang**, et al. "Facile surface modification of $CH_3NH_3PbI_3$ films leading to simultaneously improved efficiency and stability of inverted perovskite solar cells." ***Journal of Materials Chemistry A***, 2018.
- J. Li, **B. Huang**, et al. "Touching is believing: interrogating halide perovskite solar cells at the nanoscale via scanning probe microscopy." ***npj Quantum Materials***, 2017.
- **B. Huang**, et al. "Antiferroelectric polarization switching and dynamic scaling of energy storage: A Monte Carlo simulation." ***Journal of Applied Physics***, 2016.

***** Equal contribution