

ARTICLE OPEN

Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy

Nikolay Borodinov¹, Sabine Neumayer¹, Sergei V. Kalinin¹, Olga S. Ovchinnikova¹, Rama K. Vasudevan¹ and Stephen Jesse¹

The rapid development of spectral-imaging methods in scanning probe, electron, and optical microscopy in the last decade have given rise for large multidimensional datasets. In many cases, the reduction of hyperspectral data to the lower-dimension materials-specific parameters is based on functional fitting, where an approximate form of the fitting function is known, but the parameters of the function need to be determined. However, functional fits of noisy data realized via iterative methods, such as least-square gradient descent, often yield spurious results and are very sensitive to initial guesses. Here, we demonstrate an approach for the reduction of the hyperspectral data using a deep neural network approach. A combined deep neural network/least-square approach is shown to improve the effective signal-to-noise ratio of band-excitation piezoresponse force microscopy by more than an order of magnitude, allowing characterization when very small driving signals are used or when a material's response is weak.

npj Computational Materials (2019)5:25; <https://doi.org/10.1038/s41524-019-0148-5>

INTRODUCTION

The need to recover salient information from signals with a noisy background is prevalent in virtually all measurement fields and affect an instrument's resolution, sensitivity, validity, and reliability and ultimately define its capabilities. In many cases, the noise floor can be reduced using filtering such as signal averaging¹ or lock-in^{2,3} based approaches, that have been used in radio telescopes,⁴ atomic force microscopy,⁵ and depth profilometry⁶ and many other fields. Naturally, such techniques immediately limit the temporal resolution, and can be undesirable or unsuitable when fast dynamics are probed. Furthermore, these approaches specifically constrain the physics of the system that can be probed, i.e., time-dependent systems and non-linear responses may not be well processed.

In cases where the measured signal arises from an excitation, the signal-to-noise ratio (SNR) can usually be increased by increasing the excitation of the driving signal. However, this approach can introduce at least two challenges: (1) large excitations induce system nonlinearities complicating subsequent data processing, or (2) applying large excitations can damage samples. Hence, methods that improve the ability to determine signals and the valuable information contained within signals, against noisy backgrounds, will enhance the capabilities of current instrumentation and do so without hardware or experimental modifications.

In many situations, partial information about the physics of a measurement system can be used to increase the detection and sensitivity limits. For example, in scanning probe microscopy techniques, the use of resonant amplification by the cantilever allows one to increase signal-to-noise ratios, and is in fact a central aspect in virtually all dynamic SPM measurements.^{7–9} This in turn necessitates detection methods that account for this physics. For example, dual amplitude resonant tracking methods employ amplitude-based feedback on two frequencies near a resonance, whereas band-excitation methods rely on the detection of the

response of multiple frequencies across a band centered at one or more resonance.¹⁰ This approach, in turn, requires extracting a small number of parameters describing system properties (e.g., resonance frequency, amplitude, and quality factor of a damped simple harmonic oscillator (SHO) model) from the hyperspectral data.^{11,12}

Extracting physical parameters from data falls within the larger framework of “solving inverse problems”. Some of the earliest and now traditional numerical methods for solving such types of problems involve functional fitting, and are usually conducted via standard least-squares algorithms (e.g., the Levenberg–Marquardt algorithm).^{13,14} In noisy environments, these methods can be less than ideal because the algorithms are susceptible to becoming trapped in one of many incorrect local minima, as opposed to the global minimum. This can be due to a poor choice of priors (i.e., the initialization, or the parameters ‘guess’). Therefore, improving priors is vitally important and multiple approaches exist for addressing this issue. For instance, when multiple measurements are available, multivariate approaches such as principal component analysis (PCA) or non-negative matrix factorization (NMF) can be useful either in denoising,¹⁵ or with providing cleaner constituent signals, respectively.¹⁶ If multiple measurements are spatially distributed and vary slowly across space, one can average within spatial regions of similar response to reduce noise. Parameters extracted from locally averaged signals can be used as priors for least-squares fitting of the individual measurements within the region. However, all of these approaches introduce compromises between spatial resolution and signal-to-noise. Recently, interest in using machine learning (ML) tools has resulted in application of ML functionality for the purposes of processing scanning probe microscopy data for measurement artifacts correction,¹⁷ identification of probe degradation¹⁸, and classification of force spectroscopy data.¹⁹ Generally, neural networks find multiple applications as tools for rational design of complex systems due to their ability to generate

¹Center for Nanophase Materials Science, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
Correspondence: Stephen Jesse (sjesse@ornl.gov)

Received: 22 August 2018 Accepted: 10 December 2018

Published online: 22 February 2019

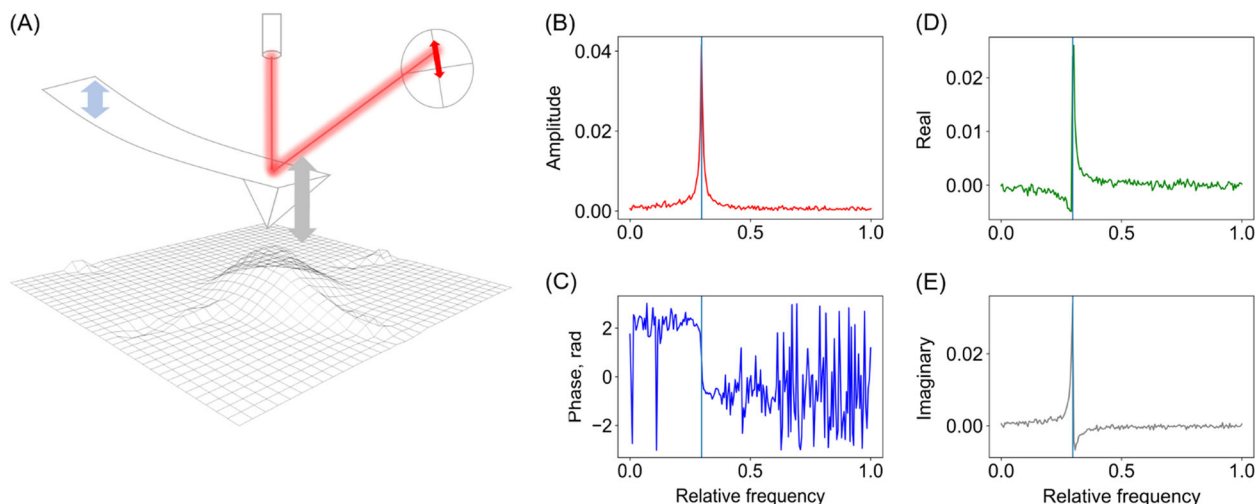


Fig. 1 Principal scheme of the atomic force microscope operating in tapping mode (a). This resonance of the oscillation is characterized by a complex function, which can be represented using (b, c) amplitude and phase or (d, e) real and imaginary components

multiparametric empirical models^{20–23} and provide parameter estimation,^{24,25} which is useful for analysis of complex scientific data such as potential energy surfaces.^{26–29} In addition, neural networks have been recently applied to condensed matter physics for the extraction of relevant degrees of freedom and identification of order parameters without a priori information.^{30–32}

Here, we introduce a fundamentally new method based on deep neural networks, to fit functional forms to noisy data based on a known physical model. We apply this method to the extraction of simple harmonic oscillator parameters³³ from piezoresponse force microscopy data, and show that by using a combination of both deep neural networks and least-squares fitting, we can probe signal responses in regimes an order of magnitude lower than with the traditional means, approaching the thermal limit for the excitation signal. As a model system, we demonstrate the extraction of damped simple harmonic oscillator parameters from band-excitation (BE) piezoresponse force microscopy (PFM) imaging of a layered ferroelectric compound. This approach of using deep neural network (DNN) is general and shows their utility as function approximators in both forward and reverse cases and that they work well in noisy environments.

RESULTS AND DISCUSSION

A typical AFM set-up is presented in Fig. 1a. Movement of the microscale beam with a sharp tip interacting with the sample is registered by monitoring its deflection via the laser beam. It is reflected from the cantilever surface into the photodiode. The resulting data are captured in the time domain and is converted into the frequency domain using a Fast Fourier Transform (FFT). The output of the FFT is a complex function that is commonly represented as amplitude (Fig. 1b), which is the absolute value of the complex number at a specific frequency and phase (Fig. 1c), and which is the angle between real and imaginary parts. However, the cantilever dynamics can be represented directly as real (Fig. 1d) and imaginary (Fig. 1e) parts as well.³⁴ Most commonly, SPM operates at a single frequency; however, capturing the broad-band cantilever response allows to extract the four parameters of the SHO equation and drastically improve the quality of the analysis. In addition to mechanical driving at the base (which is often the case in SPM), the SPM lever can also be excited at the tip by the electrically driven oscillation of the surface, as is the situation for piezoresponse force microscopy,³⁵ the technique of choice for electromechanically active materials including ferroelectrics and ionic conductors.^{36–38} In this case, the

tip stays in contact with the surface and the sample deformations induced by the converse piezoelectric effect (for piezoelectrics) or electrochemical strains (for ionics) drive cantilever motion. Such methods ensure that changes in contact mechanics are correctly interpreted with respect to the model of the oscillation, which would not be possible with a single-frequency measurement near a resonance.

The simple harmonic oscillator (SHO) serves as a good model for quantitative analysis of data generated by various scanning probe microscope set-ups. It describes the vibrational motion of the cantilever beam as a function of the frequency of induced oscillation. SHO equation (1) relates the response of the oscillator f to the frequency ω and has four major parameters: drive amplitude A_0 , resonant frequency ω_r , quality factor Q , and phase φ :

$$f(\omega) = \frac{A_0 \omega_r^2 e^{i\varphi}}{\omega^2 - \frac{i\omega\omega_r}{Q} - \omega_r^2} \quad (1)$$

The physical meaning of these parameters is well-defined, which allows using them to describe the behavior of the system being analyzed. Drive amplitude corresponds to the strength of the periodic external force acting on a system, quality factor reflects the dampening, phase or phase angle describes the shift between the excitation and response of the sample, and finally, resonant frequency reflects the overall stiffness of the system. Thus, a complete physical characterization of the sample via SPM would strongly benefit from the ability to extract these parameters from the experimental data.

Hence, the meaningful characterization of a sample using PFM requires the correlation of the observed broad-band signal with the SHO model equation. In practice, however, the noise presented in the experimental set-up, complicates extraction of the parameters. High driving amplitude may not be desirable due to the change or degradation of the sample, alternatively, the system may have intrinsically low response. In both scenarios, a low signal-to-noise ratio becomes a concern. Here, we demonstrate that using deep neural networks allows one to perform fitting and extract useful information from noisy datasets. Previously we have demonstrated application of machine-learning tools for the analysis of scanning probe³⁹ and electron microscopy data.⁴⁰

One of the major advantages of deep learning over traditional approaches in fields such as computer vision has been its remarkable ability to deal with noisy environments, including such examples as images, speech, and natural language processing.⁴¹ Indeed, neural networks, which have a single hidden layer,

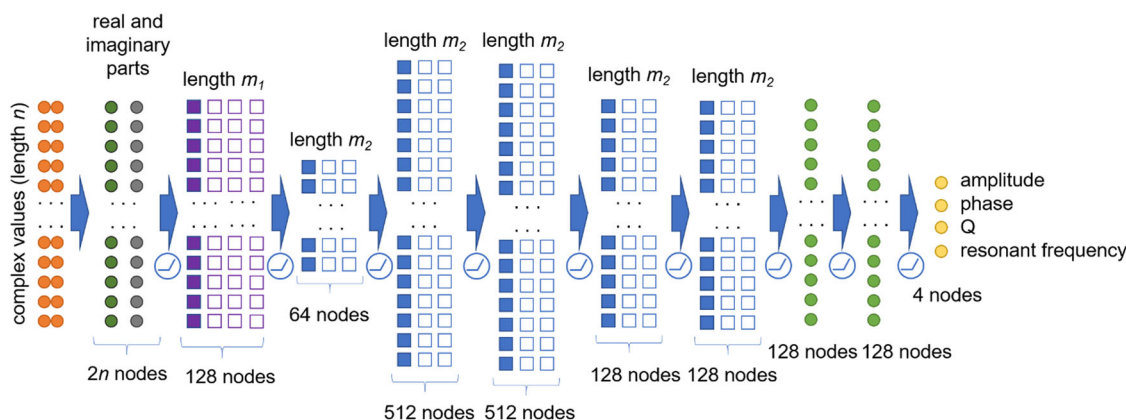


Fig. 2 The scheme of the neural network used for the functional fitting. The complex function is represented using real and imaginary parts and then passed to the first layer as a $2 \times N$ matrix. It is then followed by a series of densely connected layers consisting of vectors with length m_2 . Two dense flattened layers are connected to four outputs. The neural network is trained using parameters of the fit as outputs

but are infinitely wide are effective function approximators for any non-linear function,⁴² as follows from the universal functional approximation theorem. For known functions, it may seem counterintuitive to take a model with a handful of known parameters and then use a deep neural network with millions of parameters to approximate the same result. However, unlike least-squares method for parameter extraction, deep learning does not rely on the use of priors and can be trained on millions of examples in all manner of noisy conditions.

The deep neural network used for the fitting of the SHO data is presented in Fig. 2. The simulated SHO response with varying drive amplitudes, Q factors, resonant frequencies, phase shifts, and noise levels are split into two arrays corresponding to the real and imaginary parts of the complex-valued response. In our implementation, we process a complex output of the SHO equations by splitting it into two vectors corresponding to real and imaginary parts stacked along the new axis. For example, a complex array of size 100 would turn into a real (100,2) array. Thus, the input is a sequence of 100 vectors with 2 channels each. The first layer is the convolutional layer with 128 nodes with the kernel size 15 which determines the length of the convolutional representation. It is followed by the second convolution with 64 nodes with the kernel size of 5. This is followed by a series of dense layers with 512, 512, 128, and 128 nodes, respectively. The node is a vector with length m_2 as up until the Flatten layer, nodes of dense layers that follow Conv1D are not single numbers but rather vectors with the same length as Conv1D kernel. Proceeding this, there are two more densely connected layers with 128 nodes each. The output of the network is the four parameters of the SHO model. The rectified linear response (ReLU)⁴³ activation has been used for all layers of the network. Here, convolutional layers operate as feature-detection tools with high generalizability, which is required to return SHO response close to the ground truth for a wide range of fitting parameters. A series of dense layers with vast number of nodes composed of kernels with the length 5 was added to ensure that the structure of the network is complex enough to support multivariate optimization. Finally, flattening of the NN layers was used to yield SHO parameters. In this paper we focused on the proof-of-principle study and its applications for a relevant use case of SPM, and the optimization of the NN architecture as well as outlining minimal requirements for the number of layers, and their nature will be considered in future publications.

In order to ensure that the DNN is trained to recognize any combination of four fit parameters, we have used a batch generator, which created 100,000 SHO curves with randomized parameters, 80,000 of which were used to train the network and

20,000 of which were used for validation. To combat overfitting, each batch of 100,000 curves has been passed through the network only once (one epoch). This procedure was repeated 50 times, each time with a new set of randomly-generated SHO data. The training of the network was done on a desktop PC and takes several hours depending on the number of frequency bins of the data and computer specifications. The length of the input vector, however, does influence the overall quality of the NN prediction. Figure S1 (Supplemental information Section 1) displays the validation accuracies and losses after training on 10 batches of 80,000 curves for a series of neural networks with varying input length. It is evident that increasing input size progressively improves fit quality. We suggest that this effect can be explained by fact that if the SHO peak contains fewer points, less useful information can be extracted to determine SHO parameters. This consideration has direct implications for the practical use of NN-based fitting: the data resolution can impact the fit quality and experimental design should be selected accordingly.

The example of the SHO fitting of the actual experimental data done using DNN is presented in Fig. 3. This dataset was acquired using band-excitation PFM on a CuInP_2S_6 (CIPS)/ $\text{In}_{4/3}\text{P}_2\text{S}_6$ (IPS) composite flake of several μm thickness. While CIPS exhibits high piezoresponse and domains of positive and negative polarization orientation, the IPS phase is centrosymmetric and therefore not piezoelectrically active.⁴⁴ It is evident that although the DNN has no explicit information about the functional relationship between input and output, it can extract the parameters of the simulated data (Fig. 3a, b). However, there are cases of poor fitting like the one displayed in Fig. 3c, d where the resonant frequency is visibly off the optimal value. This is due to the fact that unlike traditional fitting algorithms, neural network recognizes the potential parameters of the fit without the optimization of their values. This behavior is a keystone feature of the DNN fitting, which also outlines its limitations. In order to further investigate the applicability of the DNN fitting method, it is necessary to compare its output with the least-square fitting method used for band excitation.

In the case of PFM, amplitude and phase of response are the two most important parameters of the system. The amplitude, which is proportional to the piezocoefficient, reveals the inactive IPS phases (blue areas on Fig. 4a), as well as the boundaries between piezoelectric CIPS domains (orange on Fig. 4a), while the phase allows to differentiate domains of different polarizations (green and blue correspond to CIPS while noisy areas correspond to IPS). Figure 4a, b displays the amplitude and phase derived using the least-square (LS) fitting algorithm, while Fig. 4c, d shows the optimal amplitude and phase determined by the DNN. The LS

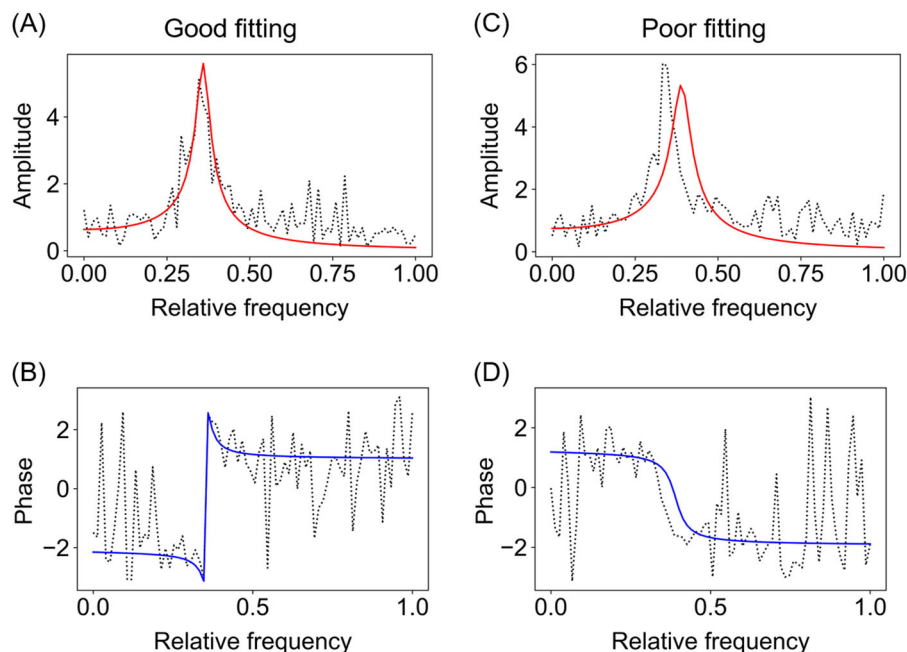


Fig. 3 Examples of fitting using a deep neural network. Overall, using the parameters generated by DNN to reconstruct the signal results in a good correspondence with the experimental data (**a**, **b**), however, it is not always perfect. The position of the resonant frequency of the fit may be slightly off (**c**, **d**)

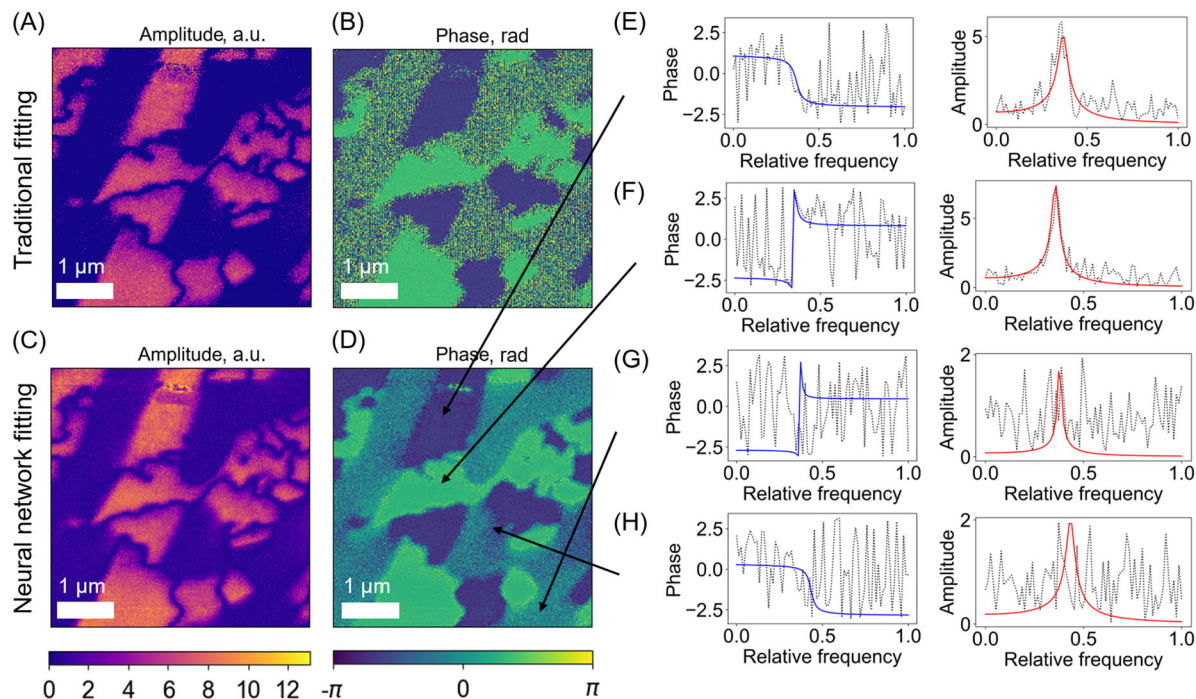


Fig. 4 The comparison of the experimental dataset fitting done by least-squares (LS) method (**a**, **b**) and by DNN (**c**, **d**). The amplitude values for LS **a** and DNN **c** correspond well to each other, however, the phases estimated by DNN **d** are less noisy compared to LS fitting **b**. The regions with strong piezoelectric response are fitted well for both phases (**e**, **f**) by DNN. When there is no piezoelectric response, DNN fitting results in phase values close to 0 (**g**). Remarkably, this method is capable of identifying phase even when the signal is very weak (**h**). IPS shows no piezoresponse and appears as blue areas on A and noisy areas on B while CIPS shows strong response and appears as orange on A and blue/green on B

fitting algorithm used here is included in *pycroscopy* package and uses contextual-driven initial guesses of the SHO fit. It is evident, that both methods can function to a satisfactory degree when the signal is strong. The signal collected from the center of ferroelectric domains is processed correctly by the neural network

(Fig. 4e, f). The signal from the inactive IPS phase (low amplitude regions) has little physical meaning in the absence of the PFM signal and renders any estimation of the phase shift values (Fig. 4g) meaningless. However, in the regions where the PFM signal is weak, DNN is capable of identifying the amplitude and

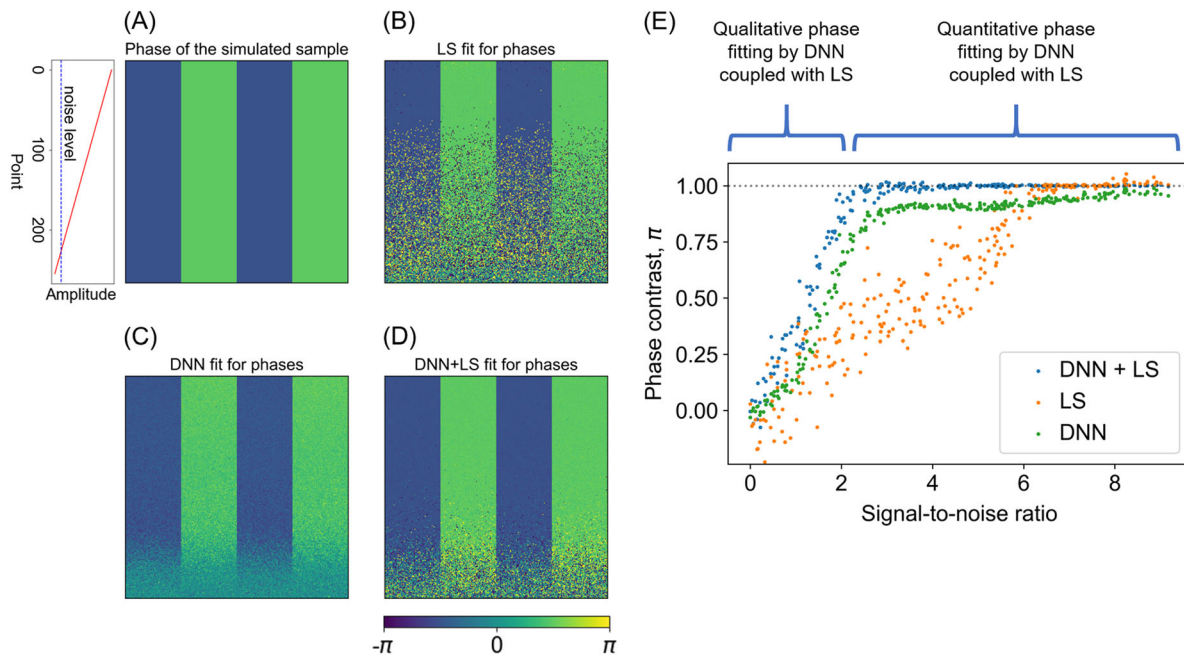


Fig. 5 Fitting of the simulated piezoelectric domains (a) with decreasing amplitude of the signal with the constant noise level highlights the difference between LS and DNN fitting: while LS results better fit for low-noise cases, it quickly loses accuracy as the signal-to-noise ratio drops (b). DNN is not as precise, however, it is much more robust and works even when the signal-to-noise is very low (c). When the results of the DNN fit are used as a guess for LS, the advantages of both methods are combined (d, e)

phase signal, while LS fitting fails to do so due to high noise (Fig. 4h). The frequency-dependent change in the phase occurs slightly to the left from the NN-fit (blue curve).

Thus, a signal buried in the noise can be picked up by the DNN more reliably, even though, just like in Fig. 3c, d the output of neural network-based fitting may not perfectly correspond to the ground truth. This result follows what one would expect from a DNN. Namely, that it can generalize well, but perhaps not yield the exact answer, and that the output could benefit from subsequent optimization to find a more exact answer.

Comparison between the LS and DNN approaches thus far has shown that there are regimes in which LS methods produce better results, and ones (in particular low SNR) where DNN is superior. However, the complimentary nature of the methodologies and behaviors of the two approaches suggests that further improvements can be achieved through their synergistic combination. The remarkable robustness of the DNN-based fitting with respect to noise can be used to estimate parameters that can undergo further optimization via iteration. In order to test this hypothesis, we created an artificial dataset. The Q factor and resonant frequency were unchanged throughout, phase of the response was varied as stripe domains, and the drive amplitude was linearly decreased from top to the bottom (Fig. 5a) of a simulated scan while the noise level was kept the same. This roughly models a sample with four domains. The LS fit that used uniform guess across the simulated dataset shows excellent match when the noise is low; however, as amplitude decreases, the fit becomes progressively worse (Fig. 5b). DNN phase fit, however, maintains its utility much better and shows much clearer phase contrast. At the same time, it is not very accurate, and the estimated phase contrast appears to be smaller than it should be (Fig. 5c). When the results of the DNN are used as inputs for the LS optimizer, however, the robustness to noise of the DNN is combined with the accuracy of LS (Fig. 5d). This is summarized in Fig. 5e. The bottom axis is the signal-to-noise ratio (SNR) calculated as the maximum amplitude of the noiseless simulated signal in Fourier domain divided by the standard deviation of the noise. The left axis is the phase contrast defined as the difference between average values

of the phase estimated by a given fitting method. For a perfect fit, this difference must be equal to π . It is clear that fitting algorithms have two regimes of approximation. At high SNR values, the phase contrast is close to π , which we label as quantitative fitting. At low SNRs, the phase difference is manifested, but it is progressively noticeably smaller than π as the SNR decreases. This regime, which we term qualitative fitting, can be used to process data and can yield contrast between domains of different polarization, but the results of the fit lack accuracy. LS fit switches to qualitative regime at SNR ~ 6 . The DNN fit is much more robust and experiences regime transition at SNR ~ 2 ; however, its prediction above that value deviates from π , which is consistent with the previous observation. The serial combination of DNN followed by LS shows superior performance over each of them used separately. In combination, the DNN component, which can be characterized as a broadly sweeping and holistic means to assess data, provides a reliable initial guess for LS even into low SNR regimes, and with a reliable guess, the iterative and pedantic nature of LS provides a more accurate determination of parameters that are not possible by DNN. This way, the LS algorithm is initialized within the correct minima of the multiparametric space and more easily converges to it. When the amplitude of the signal approaches zero, the phase contrast for all methods approaches zero as well. Practically this suggests that there is no systematic false-positive identification. Overall, this highlighted the functionality of the DNN for processing of the SPM broad-band data: its best use is to provide a guess for the LS optimizer.

Since hybrid fitting showed a unique combination of accuracy and stability to high noise thresholds, we have further explored its applicability as a tool for data analysis. Specifically, we have chosen a material with strong and known ferroelectric properties (Bismuth ferrite (BiFeO₃, or BFO) and investigated its response at decreasing values of piezoelectric drive amplitude (Fig. 6). In this set-up, we can directly compare information provided by the fitting of experimental data with varying SNR ratios. We have used four methods of fitting: least-square with uniform guesses (A) (the same as was used for Fig. 5), least-square with contextual-driven guesses (B) (the same as was used for Fig. 4, implemented in

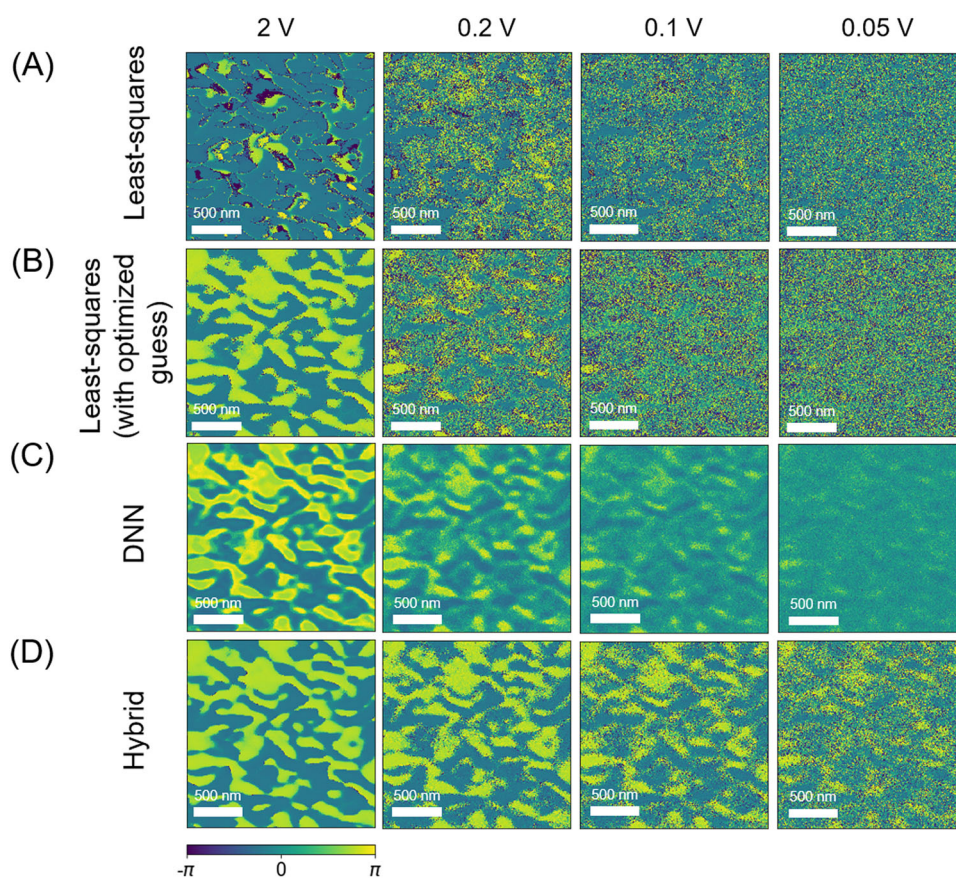


Fig. 6 Piezoresponse force microscopy phase maps obtained by fitting the lateral PFM signal with decreasing drive amplitude: a comparison of least-squares with uniform guesses (**a**), least-squares with guesses generated using traditional methods (**b**), deep neural network fitting (**c**) and a hybrid fit (**d**)

pycroscopy package)⁴⁵, deep neural network fitting (C), and hybrid fitting (D). It is evident that using uniform guesses result in poor fitting even for high driving voltages (Fig. 6a). However, supplying more meaningful initial guesses strongly improves the fit quality (Fig. 6b). This serves as another vivid demonstration that least-squares is a powerful method, however, its convergence is heavily dependent on the starting point. When the SNR is decreased by an order of magnitude from 2 to 0.2 V, traditional methods of finding this starting point are no longer effective. The results of DNN fitting are presented in Fig. 6c; however, hybrid fitting is found to be superior to all of the above-mentioned methods (Fig. 6d). In fact, a comparison between hybrid and state-of-art fitting reveals that former allows for phase contrast analysis with 10–20 times smaller SNR. The details of contrast extraction are discussed in Supplemental information Section 2.

We attribute this to the fact that a single spike in a spectrum might be interpreted as a resonance peak by the converging LS optimizer, while DNN considers correlations across both real and imaginary signals and across the entire band. At the same time, some peculiarities of the neural networks must be respected for the successful design of hybrid fitter. As it was previously mentioned, DNN does not directly utilize a concrete physical model in the process of the fitting. Consequently, it may generate physically unfeasible outputs (such as Q factors equal to 0). While this happens in less than one percent of cases, it is practically useful to bypass such values with some predetermined guess values. While the exact architecture of the network as well as function to be fitted can be customized of case-by-case basis, we believe that our approach in the current state can be readily adopted for other applications requiring fitting of a known function, which quantitatively describes certain physical

processes. In this case batch generation of synthetic datasets becomes a viable approach to train neural network and ultimately extract relevant multivariate parameters. We also suggest that the output of the NN fitter needs to undergo further optimization using any appropriate technique (such as least-square optimizer) to ensure the precision of parameter estimation.

We demonstrate a novel approach for the inverse problem solution and extraction of physical model parameters from spectral-imaging data-based least-squares fitting augmented by deep learning for determination of priors. Pattern recognition allows for accessing functional properties of materials with a signal that is more than an order of magnitude weaker than it was possible without it approaching thermal limit. Specifically, for the case of piezoresponse microscopy, we demonstrate imaging at the order of magnitude lower excitation voltages.

The use of deep learning as a tool to generate priors for functional fitting algorithms can be extremely beneficial in a broad range of instrumentation and measurement applications, helping to increase the range of materials that can be studied (via reduction in the amplitude of required excitation), as well as possible advances in the temporal resolution, due to the reduction in need to signal average in time^{46,47}. The DNN method is also relatively fast, taking ~ms for 100 curves on a good GPU. We further argue that this approach can be broadly applied to more complex physical models of the response. This approach is expected to be immediately applicable for other resonance-based SPM techniques including atomic force acoustic microscopy (AFAM)^{48,49}, magnetic force microscopy (MFM)⁵⁰, and KPFM.⁵¹

In the future, the implementation of these networks into hardware will greatly accelerate processing, and thereby enhance effective instrument capabilities with existing experimental

hardware. In essence, these approaches allow one to push the fundamental limits of the instruments via increased information extraction from the measured signals. We envision that fitting algorithms involving neural networks can be successfully applied to more general task finding inverse problems solutions by providing optimal initial conditions and guiding searches for traditional computation parameter extraction approaches.

METHODS

Piezoresponse force microscopy measurements

Band-excitation PFM was conducted using Cypher atomic force microscopes (Asylum Research) combined with National Instruments electronics and custom LabView codes for signal generation and data acquisition.

The composite CuInP_2S_6 (CIPS)/ $\text{In}_{4/3}\text{P}_2\text{S}_6$ (IPS) sample was prepared as described elsewhere.⁴⁴ Band-excitation PFM measurements of the vertical response were performed on a CIPS/IPS flake of several μm thickness attached to a copper circuit board using conductive silver paint. The drive voltage was 1 V, within a frequency band of 120 kHz centered around the contact resonance of ~ 300 kHz using a conductive probe (Nanosensor PPP-EFM, nominal force constant = 2.8 N/m, nominal free resonance = 75 kHz).

Band-excitation PFM on a bismuth ferrite (BFO) thin film (thickness = 100 nm) grown on a SRO/STO substrate and mounted on a grounded support was conducted using Multi75-G Budget sensor probes (nominal force constant = 3 N/m, nominal 75 kHz). For BFO, the lateral band-excitation PFM response was acquired. Maps of the ferroelectric domains were imaged using seven values of the driving voltage: 2 V, 0.2 V, 0.1 V, 0.05 V, 0.03 V, 0.02 V, and 0.01 V, within a frequency band of 30 kHz centered around the contact resonance of ~ 620 kHz.

Neural network implementation

Data processing was done using Python 3.6. Keras with TensorFlow backend was used to build up and train a deep neural network. Intel Xeon CPU E-5-1650 v3 3.50 GHz processor and 40 GB of RAM were used to perform the computations.

DATA AVAILABILITY

Scanning probe microscopy data as well as Python scripts used for the analysis are available from the authors upon request.

ACKNOWLEDGEMENTS

This research was conducted at the Center for Nanophase Materials Sciences, which is a DOE Office of Science User facility. Data analysis effort (NB, RKV, SJ) was sponsored by the Laboratory Directed Research and Development Program (as a part of the AI Initiative) of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy (DOE). This research used resources of the Compute and Data Environment for Science (CADES) at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725

AUTHOR CONTRIBUTIONS

Concept development was done by N.B., S.V.K., O.S.O., and S.J. Data acquisition was done by N.B. and S.N. Manuscript writing was done by N.B. and R. K. V.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Computational Materials* website (<https://doi.org/10.1038/s41524-019-0148-5>).

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Chen, S. W., Chen, H. C. & Chan, H. L. A real-time QRS detection method based on moving-averaging incorporating with wavelet denoising. *Comput. Methods Prog. Biomed.* **82**, 187–195 (2006).
- Masciotti, J. M., Lasker, J. M. & Hiescher, A. H. Digital lock-in detection for discriminating multiple modulation frequencies with high accuracy and computational efficiency. *IEEE T. Instrum. Meas.* **57**, 182–189 (2008).
- Sonnaillon, M. O. & Bonetto, F. J. A low-cost, high-performance, digital signal processor-based lock-in amplifier capable of measuring multiple frequency sweeps simultaneously. *Rev. Sci. Instrum.* **76**, 024703 (2005).
- Boonstra, A. J. & van der Veen, A. J. Gain calibration methods for radio telescope arrays. *IEEE T. Signal Proces.* **51**, 25–38 (2003).
- Stark, M. & Guckenberger, R. Fast low-cost phase detection setup for tapping-mode atomic force microscopy. *Rev. Sci. Instrum.* **70**, 3614–3619 (1999).
- Fan, Y. et al. Laser photothermoacoustic heterodyned lock-in depth profilometry in turbid tissue phantoms. *Phys. Rev. E* **72**(5 Pt 1), 051908 (2005).
- Dazzi, A., Saunier, J., Kjoller, K. & Yagoubi, N. Resonance enhanced AFM-IR: a new powerful way to characterize blooming on polymers used in medical devices. *Int. J. Pharm.* **484**, 109–114 (2015).
- Rodriguez, B. J., Callahan, C., Kalinin, S. V. & Proksch, R. Dual-frequency resonance-tracking atomic force microscopy. *Nanotechnology* **18**, 6 (2007).
- Sommerhalter, C., Matthes, T. W., Glatzel, T., Jager-Waldau, A. & Lux-Steiner, M. C. High-sensitivity quantitative Kelvin probe microscopy by noncontact ultra-high-vacuum atomic force microscopy. *Appl. Phys. Lett.* **75**, 286–288 (1999).
- Jesse, S. et al. in *Annual Review of Physical Chemistry* Vol. 65 (eds Johnson, M. A. & Martinez, T. J.) 519–536 (Palo Alto, 2014).
- Jesse, S. & Kalinin, S. V. Band excitation in scanning probe microscopy: sines of change. *J. Phys. D Appl. Phys.* **44**, 1–16 (2011).
- Jesse, S., Kalinin, S. V., Proksch, R., Baddorf, A. P. & Rodriguez, B. J. The band excitation method in scanning probe microscopy for rapid mapping of energy dissipation on the nanoscale. *Nanotechnology* **18**, 1–8 (2007).
- Budil, D. E., Lee, S., Saxena, S. & Freed, J. H. Nonlinear-least-squares analysis of slow-motion EPR spectra in one and two dimensions using a modified Levenberg-Marquardt algorithm. *J. Magn. Reson. Ser. A* **120**, 155–189 (1996).
- Nowak, W. & Cirpka, O. A. A modified Levenberg-Marquardt algorithm for quasi-linear geostatistical inversing. *Adv. Water Resour.* **27**, 737–750 (2004).
- Jesse, S. & Kalinin, S. V. Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy. *Nanotechnology* **20**, 085714 (2009).
- Kannan, R. et al. Deep data analysis via physically constrained linear unmixing: universal framework, domain examples, and a community-wide platform. *Adv. Struct. Chem. Imaging* **4**, 6 (2018).
- Martinek, J., Klapetek, P. & Campbell, A. C. Methods for topography artifacts compensation in scanning thermal microscopy. *Ultramicroscopy* **155**, 55–61 (2015).
- Rashidi, M. & Wolkow, R. A. Autonomous scanning probe microscopy in situ tip conditioning through machine learning. *ACS Nano* **12**, 5185–5189 (2018).
- Karatay, D. U., Zhang, J., Harrison, J. S. & Ginger, D. S. Classifying force spectroscopy of DNA pulling measurements using supervised and unsupervised machine learning methods. *J. Chem. Inf. Model.* **56**, 621–629 (2016).
- Yin, F. et al. Back propagation neural network modeling for warpage prediction and optimization of plastic products during injection molding. *Mater. Des.* **32**, 1844–1850 (2011).
- Yun, S. Y., Namkoong, S., Rho, J. H., Shin, S. W. & Choi, J. U. A performance evaluation of neural network models in traffic volume forecasting. *Math. Comput. Model.* **27**, 293–310 (1998).
- Gupta, V. K. et al. Prediction of capillary gas chromatographic retention times of fatty acid methyl esters in human blood using MLR, PLS and back-propagation artificial neural networks. *Talanta* **83**, 1014–1022 (2011).
- Lee, W. Y. et al. Empirical modeling of polymer electrolyte membrane fuel cell performance using artificial neural networks. *Int. J. Hydrog. Energ.* **29**, 961–966 (2004).
- Fairbairn, E. M. R., Paz, C. N. M., Ebecken, N. F. F. & Ulm, F. J. Use of neural networks for fitting of FE probabilistic scaling model parameters. *Int. J. Fract.* **95**, 315–324 (1999).
- Peyada, N. K. & Ghosh, A. K. Aircraft parameter estimation using a new filtering technique based upon a neural network and Gauss-Newton method. *Aeronaut. J.* **113**, 243–252 (2009).
- Dawes, R., Thompson, D. L., Guo, Y., Wagner, A. F. & Minkoff, M. Interpolating moving least-squares methods for fitting potential energy surfaces: computing high-density potential energy surface data from low-density ab initio data points. *J. Chem. Phys.* **126**, 184108 (2007).
- Handley, C. M. & Popelier, P. L. Potential energy surfaces fitted by artificial neural networks. *J. Phys. Chem. A* **114**, 3371–3383 (2010).

28. Manzhos, S., Wang, X., Dawes, R. & Carrington, T. Jr. A nested molecule-independent neural network approach for high-quality potential fits. *J. Phys. Chem. A* **110**, 5295–5304 (2006).
29. Blank, T. B., Brown, S. D., Calhoun, A. W. & Doren, D. J. Neural-Network Models of Potential-Energy Surfaces. *J. Chem. Phys.* **103**, 4129–4137 (1995).
30. Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431 (2017).
31. van Nieuwenburg, Evert P. L., Liu, Y.-H., Huber & Sebastian, D. Learning phase transitions by confusion. *Nat. Phys.* **13**, 435–439 (2017).
32. Broecker, P., Carrasquilla, J., Melko, R. G. & Trebst, S. Machine learning quantum phases of matter beyond the fermion sign problem. *Sci. Rep.* **7**, 8823 (2017).
33. Gannepalli, A., Yablon, D. G., Tsou, A. H. & Proksch, R. Mapping nanoscale elasticity and dissipation using dual frequency contact resonance AFM. *Nanotechnology* **22**, 355705 (2011).
34. Lozano, J. R. & Garcia, R. Theory of multifrequency atomic force microscopy. *Phys. Rev. Lett.* **100**, 076102 (2008).
35. Kalinin, S. V., Karapetian, E. & Kachanov, M. Nanoelectromechanics of piezoresponse force microscopy. *Phys. Rev. B* **70**, 184101 (2004).
36. Chen, Q. N. et al. Delineating local electromigration for nanoscale probing of lithium ion intercalation and extraction by electrochemical strain microscopy. *Appl. Phys. Lett.* **101**, 063901 (2012).
37. Li, J. Y. et al. Strain-based scanning probe microscopies for functional materials, biological structures, and electrochemical systems. *J. Mater.* **1**, 3–21 (2015).
38. Proksch, R. Electrochemical strain microscopy of silica glasses. *J. Appl. Phys.* **116**, 066804 (2014).
39. Nikiforov, M. P. et al. Functional recognition imaging using artificial neural networks: applications to rapid cellular identification via broadband electro-mechanical response. *Nanotechnology* **20**, 405708 (2009).
40. Ziatdinov, M. et al. Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations. *ACS Nano* **11**, 12742–12752 (2017).
41. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
42. Hornik, K. Approximation Capabilities of Multilayer Feedforward Networks. *Neural Netw.* **4**, 251–257 (1991).
43. Glorot, X., Bordes, A. & Bengio, Y. Deep Sparse Rectifier Neural Networks. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, Vol. 15 (eds Geoffrey, G., David, D. & Miroslav, D.) PMLR: Proceedings of Machine Learning Research, 2011; pp 315–323.
44. Susner, M. A. et al. *ACS Nano* **9**, 12365–12373 (2015).
45. Somnath, S., Smith, C. R., Laanait, N. & Jesse, S. *Pycroscopy*. Computer software, 0.60.0; Oak Ridge National Laboratory: <https://pycroscopy.github.io/pycroscopy/about.html> (2018).
46. Somnath, S., Belianinov, A., Kalinin, S. V. & Jesse, S. Full information acquisition in piezoresponse force microscopy. *Appl. Phys. Lett.* **107**, 263102 (2015).
47. Collins, L. et al. Multifrequency spectrum analysis using fully digital G Mode-Kelvin probe force microscopy. *Nanotechnology* **27**, 105706 (2016).
48. Nikiforov, M. P. et al. Probing the temperature dependence of the mechanical properties of polymers at the nanoscale with band excitation thermal scanning probe microscopy. *Nanotechnology* **20**, 395709 (2009).
49. Jesse, S., Nikiforov, M. P., Germinario, L. T. & Kalinin, S. V. Local thermomechanical characterization of phase transitions using band excitation atomic force acoustic microscopy with heated probe. *Appl. Phys. Lett.* **93**, 073104 (2008).
50. Collins, L. et al. G-mode magnetic force microscopy: separating magnetic and electrostatic interactions using big data analytics. *Appl. Phys. Lett.* **108**, 193103 (2016).
51. Collins, L. et al. Band excitation Kelvin probe force microscopy utilizing photo-thermal excitation. *Appl. Phys. Lett.* **106**, 104102 (2015).



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2019