

PAPER • OPEN ACCESS

## Towards automating structural discovery in scanning transmission electron microscopy\*

To cite this article: Nicole Creange *et al* 2022 *Mach. Learn.: Sci. Technol.* **3** 015024

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

### You may also like

- [An On-Demand Processing Framework for Faster Remote Sensing Big Data Analysis](#)  
Zhenchun Huang
- [A Green Bank Telescope Survey of Large Galactic H ii Regions](#)  
L. D. Anderson, W. P. Armentrout, Matteo Luisi et al.
- [From biologically-inspired physics to physics-inspired biology](#)  
Alexei A Kornyshev



## PAPER

## OPEN ACCESS

RECEIVED  
28 July 2021REVISED  
25 October 2021ACCEPTED FOR PUBLICATION  
10 November 2021PUBLISHED  
11 February 2022

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

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



# Towards automating structural discovery in scanning transmission electron microscopy\*

Nicole Creange<sup>1,\*\*</sup> , Ondrej Dyck<sup>1</sup>, Rama K Vasudevan<sup>1</sup> , Maxim Ziatdinov<sup>1,2</sup> and Sergei V Kalinin<sup>1</sup><sup>1</sup> Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States of America<sup>2</sup> Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States of America

\*\* Author to whom any correspondence should be addressed.

E-mail: [creangenc@ornl.gov](mailto:creangenc@ornl.gov)**Keywords:** STEM, Bayesian optimization, automated experiments, data acquisition workflows

## Abstract

Scanning transmission electron microscopy is now the primary tool for exploring functional materials on the atomic level. Often, features of interest are highly localized in specific regions in the material, such as ferroelectric domain walls, extended defects, or second phase inclusions. Selecting regions to image for structural and chemical discovery via atomically resolved imaging has traditionally proceeded via human operators making semi-informed judgements on sampling locations and parameters. Recent efforts at automation for structural and physical discovery have pointed towards the use of ‘active learning’ methods that utilize Bayesian optimization with surrogate models to quickly find relevant regions of interest. Yet despite the potential importance of this direction, there is a general lack of certainty in selecting relevant control algorithms and how to balance *a priori* knowledge of the material system with knowledge derived during experimentation. Here we address this gap by developing the automated experiment workflows with several combinations to both illustrate the effects of these choices and demonstrate the tradeoffs associated with each in terms of accuracy, robustness, and susceptibility to hyperparameters for structural discovery. We discuss possible methods to build descriptors using the raw image data and deep learning based semantic segmentation, as well as the implementation of variational autoencoder based representation. Furthermore, each workflow is applied to a range of feature sizes including NiO pillars within a La:SrMnO<sub>3</sub> matrix, ferroelectric domains in BiFeO<sub>3</sub>, and topological defects in graphene. The code developed in this manuscript is open sourced and will be released at [github.com/nccreang/AE\\_Workflows](https://github.com/nccreang/AE_Workflows).

## 1. Introduction

Over the last several decades, electron microscopy has become the mainstay of multiple fields ranging from materials science to condensed matter physics and biology [1–3]. In particular, the introduction of aberration correction in the late 1990s and its broad commercialization in subsequent decades has enabled advancements ranging from high veracity structural imaging [4–6], electron energy loss spectroscopy (EELS) with single atom sensitivity [7], and depth-resolved imaging [8, 9]. The last decade has significantly expanded this range of capabilities due to the emergence of sub-10 mV resolution EELS of plasmonic and phonon excitations [10], 4D scanning transmission electron microscopy (STEM) imaging [3, 11–13], and beams with orbital momentum [14, 15]. Similarly, momentous progress has been made in dynamic STEM

\* This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC0500OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

imaging, allowing the visualization of thermally, chemically, and beam induced transformations [16, 17]. Recently, these beam induced reactions were harnessed to enable patterning [18], manipulation [19, 20], and assembly [21, 22] at the atomic level. However, despite the gamut of advances in the development of source and detectors in STEM, the fundamental principle of scanning microscopies has remained largely the same, namely the sequential acquisition of data and the data collection workflow.

In most materials systems the distribution of data of interest is non-uniform in the image plane. For example, in ferroelectric thin films, the phenomena of interest occurs at interfaces and topological defects which appear in non-uniform patterns. In nanoparticle systems, the phenomena of interest occurs at the individual nanoparticles and especially at their junctions, whereas the empty space between them is often featureless. Finally, in atomically resolved data of systems such as graphene and 2D chalcogenides, phenomena of interest occurs at atomic and extended defects, which can occupy a relatively small fraction of the image plane. Correspondingly, an optimal imaging strategy will preferentially discover these physically-relevant features.

Human-operated STEM workflows are the mainstay of the microscopy field and consist of rastering over a square grid in multiple regions before finding an area of actual interest; however, this process is extremely limiting. First and foremost, the latency time of a human operator is well below that of the image data acquisition in STEM. Hence, the identification of regions of interest is currently human limited, making studies of dynamic phenomena such as catalysis, growth, or electron-beam manipulation slow. Secondly, the human eye is generally well suited for the identification of certain localized features, but the selection is highly biased by the operator experience and interest, introducing a large degree of observation bias in the exploration. Similarly, the discriminative capability of the human eye is limited and can be strongly affected by the chosen visualization settings such as image contrast and color scale. Finally, subtle changes in the image, such as a slight periodicity change, are usually undetected by the human eye and can be revealed only by post-experiment image analysis.

These considerations, as well as recent advances in autonomous robotics and vehicles, have sparked a tremendous interest towards the development of automated experimentation (AE) in general microscopy [23–25] and STEM [26, 27] in particular. Some of the specific targets for AE include the identification and subsequent imaging of specific objects of interest in static systems, aiming at: minimizing the total measurement time and consequent beam damage, imaging dynamic systems, and ultimately controlled/completely automated probe manipulation. In the most general case, implementation of the AE in microscopy requires a set of synergistic instrumental and control advances, including controlled beam/probe motion, rapid analysis of the data to extract descriptors of interest and feedback based on these descriptors.

For scanning probe microscopies, many of the instrumental aspects of these developments are far from trivial, however, control of the probe scan path was demonstrated fairly early by Schwartzentruber [28] for scanning tunneling microscopy (STM) and Ovchinnikov *et al* [29] and Requicha *et al* [30–32] for atomic force microscopy (AFM). Furthermore, broad implementation of non-rectangular scans has started over the last several years [33–35], paving the way into the development of automated processes in AFM and STM. These developments in turn necessitate the development of mathematical tools that can convert the data acquired over complex scan patterns into the standard 2D images, including compressed sensing [33, 34, 36, 37] and Gaussian process (GP) [38] methods. Due to the computational complexity, this has become viable only over the last 5 years. Comparatively in STEM, the progress has been slower due to the closed hardware architecture of commercial microscopes.

The key remaining element is the identification of the structures of interest and generation of the feedback signal. Structural discovery is not a universal concept, as it can be driven by different considerations and often reward functions are defined *a priori*. For example, in atomic manipulation experiments [19–21, 39–46], we seek to find regions of clean graphene, away from contaminations and free of defects. Reversely, in experiments focusing on exploring the chemistry of graphene, we want to find regions with high defect density. For specific applications, we may be interested in grain boundaries rather than point defects, or more generally, we can be interested in more- or less-crystalline regions, exploring the variability of possible defect structures, or finding all defects of a particular type. Here we demonstrate the general framework of how these workflows can be constructed in the absence of prior information.

Despite the relatively small number of demonstrations, currently the enabling instrumentation and software controls for automated experiments in STEM, if not necessarily widespread, are at least accessible. The proliferation of the software environments such as NION SWIFT [47] or PyJem [48] further suggest that automated experiments will soon become commonplace. Missing now are the algorithms and workflows that allow the implementation of automated experiments in STEM, i.e. convert the video stream from the detector into the control signals to the beam scanning, optimizing (in a certain sense) information acquisition.

Here, we discuss several general paradigms for the automated experiment in STEM utilizing classical rectangular scan workflows and develop the experimental workflow for structural discovery. In this

discussion, we consider the factors such as the existence of the partial or full prior knowledge on the system, knowledge of the exploration targets, the nature of the collected signal used to guide exploration, and the presence and relevance of the dynamic changes in the system. We further discuss how these approaches can generalize to the non-rectangular predefined scans and freeform scans.

## 2. Methods

The simplest form of AE in imaging is when the objects of interest are known *a priori*, from prior experiments or theoretical considerations. These can be formulated based both on the positive criteria (known atomic defect configurations, ferroelectric domain walls) or negative criteria (i.e. breaking of ideal lattice). In this case, the identification of the object of interest can be based on a simple function that converts the local signal vector into the measurement of interest. The signal vector can be either the sub-image, including classical rectangular scan, the scan over an arbitrary preset trajectory, or the scan over a dynamically adjusted trajectory, or a spectrum. For scans over a rectangular or static preset trajectory, the conversion from image data to the measurement of interest can be performed using the suitably chosen function, e.g. based on a simple correlation function or a more complex deep convolutional neural network (DCNN).

In the cases where the features of interest are unknown, the exploration algorithm can be based on the dynamic libraries of objects of interest. In this approach, the microscope exploits the unknown system and identifies the anomalies using the initial batch of experimental data. These are used to create the features of interest in the unsupervised manner. With these, the physical or statistical features of interest can be used to guide the probe motion over the image plane, e.g. using Bayesian Optimization. Here, we develop the approach for the automated experiment workflow in STEM based on a two-tiered approach including supervised atom identification via DCNNs and unsupervised feature of interest discovery via rotationally invariant variational autoencoders (rVAEs). The acquisition function is a function of the predicted value of the property of interest and the associated uncertainty (i.e. the output of a surrogate model). The maximum/minimum of the acquisition function is used to select the next measurement point. This approach is compared with the simpler method based on the pretrained linear transform-based AE workflows.

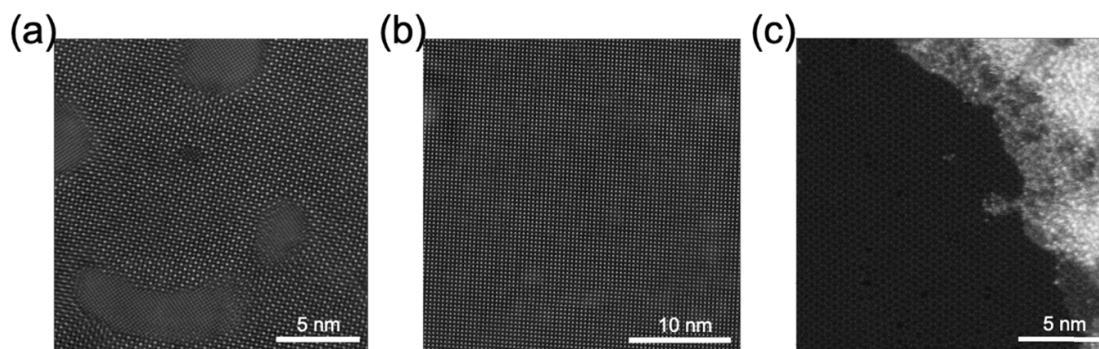
As a static data set, presented in figure 1, we chose high-angle annular dark-field images of various materials which each contain different feature lengths. The dataset includes: NiO pillars in a La:SrMnO<sub>3</sub> (NiO-LSMO) matrix [49–51], sample of BiFeO<sub>3</sub> (BFO) [52–54] and Si-containing graphene [55, 56]. Each of these images contain different information of interest, such as domain walls, defects, and dopants. In the NiO-LSMO system, the phases are obvious, however it is unknown if all the interfaces are the same or if any structural anomalies lie in the bulk. For Sm-BFO and graphene, the aim is to automatically identify topological defects but at different length scales.

The proposed analysis workflows are generally separated into three parts: image segmentation, linear or non-linear dimensionality reduction, potentially allowing for physics-based invariances, and data collection. Within each part, a variety of methods exist which may produce similar results. Here we will focus on two methods of image segmentation and two methods of image decomposition while the data collection is simulated via Bayesian optimization.

For image segmentation, the first method makes use of pre-trained DCNNs. This choice is dictated by the fact that in most cases of interest in the atomically resolved STEM and STM studies are the structures and transformations on the atomic level. Furthermore, DCNN based atom finding is robust and hence can be performed given the pretrained networks for a broad range of image sizes, sampling, signal to noise ratios, etc. The DCNN predictions can further be improved via ensemble learning and iterative training approaches [51, 55]. It is important to also note that DCNN can be trained for the semantic segmentation of mesoscopic images for determination of specific microstructural elements, e.g. particle and grain boundary finding, determination of the ferroelectric domain walls, etc, allowing for generalization of this approach towards other (prior known) objects of interest. Furthermore, the training of the DCNNs is done for a specific target (type of lattice) and more details on the training of neural networks and their adaptation to new data can be found in works by Ghosh and Ziatdinov [52, 57, 58].

The output of the DCNN is the softmax probability density that a given image pixel belongs to atom (single layer) or specific atomic class. This softmax probability density, different from a Bayesian posterior density or ensemble uncertainty, can be further used to reconstruct the centroids of the atoms, and hence determine the atomic coordinates. The latter can be further refined, either from original or semantically segmented images, using the functional fits or more complex Bayesian inference methods that include full or partial knowledge of the beam shape. Based on the atomic coordinates, the local neighborhood of the selected atom [59] can be used as a local structural descriptor vector.

The alternative image segmentation approach can be based on a suitable transform of the sub-image at an arbitrary position (i.e. not necessarily centered on an atom). This can be performed via fast Fourier



**Figure 1.** High-Angle annular dark-field (HAADF) STEM images of (a) NiO pillars in La:SrMnO<sub>3</sub> [49, 50], (b) BiFeO<sub>3</sub> [52–54] and (c) graphene [55, 56].

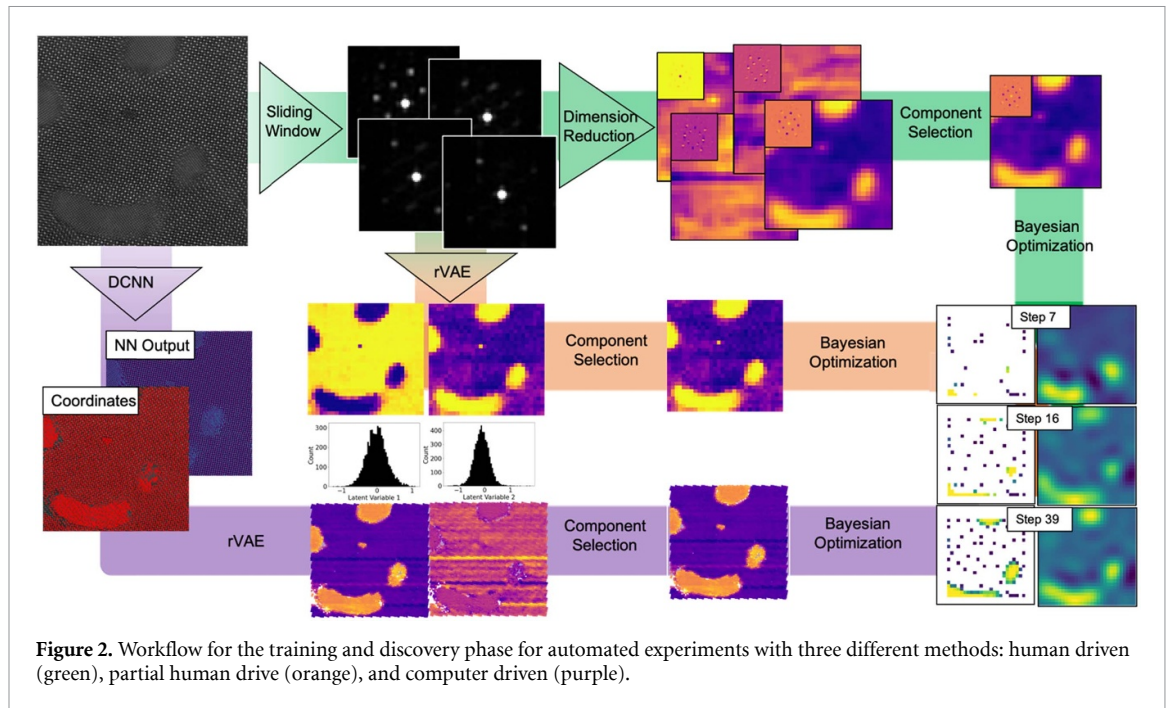
Transform (FFT) of the sub-image [60], or Radon or Hough transform of the mesoscopic data. However, in this case selection of the surrogate model requires prior knowledge of the exploration target. For example, this can be achieved via projection of the FFT on a specific linear unmixing component known *a priori*, or selection of the intensity at a given coordinate. In the proposed workflows, we will explore the use of a sliding window FFT to produce a stack of sub-images.

The second stage of the general workflow involves image decomposition in which two approaches are discussed: manual decomposition via non-negative matrix factorization (NMF)/single value decomposition (SVD) and unsupervised rVAEs. Briefly, variational autoencoders refer to the general class of the encoder-decoder networks in which the encoder projects the large dimensional vector space to a small number of latent variables. The decoder then draws variables from the latent space and up-samples them to the image. The encoder and decoder are trained jointly, providing for the efficient encoding of input object as a continuous latent variable. The rotationally invariant VAEs are configured so that three of the latent variables are rotational angle and x- and y- offsets, allowing to localize similar objects at different orientations within the image plane.

Both approaches produce low-dimensional representations which are then mapped back onto the image and selection of the appropriate dimension which represents the property of interest can be selected. On transition to the AE stage, the selection of the latent variable(s) is utilized to define the surrogate model used in BO. In the simplest case, the latent variable with the largest change across the latent space can be used; however, here the human intervention can be used to select the latent variable, or a combination of auto-selection and human selection can be used. Note that while VAEs belong to the unsupervised machine learning methods, they rely on the existence of a full data set for training. This is not the case for AE, where the information becomes accessible through the course of the experiment. Here, we explore the approach where the rVAE is trained during the discovery stage of the experiment.

We further implement the Bayesian optimization methods for AE with GP as the surrogate models. Generally, GP refers to an approach for reconstruction of relationships between two sets of parameters and yields the predicted values of a function and the associated uncertainty. This surrogate model is then used in conjunction with an acquisition function during the Bayesian optimization process. In this whole process, we have several factors which control the relationship between parameters including the GP covariance function (commonly referred to as kernel) and the type of acquisition function. The choice of kernel sets the base ‘shape’ of the relationship and common kernels include Gaussian and Matern kernels [61]. The search strategy is based on the balance between exploration and exploitation. In the exploration regime, we are able to search a large area of the sample for regions of interest, whereas in the exploitation regime, we probe near an existing area of interest to gain more information around the area. Here we use three different acquisition functions: expected improvement, upper confidence bound, and probability of improvement. The feature length scale and particular data set are critical to the specific kernel and acquisition function used. For example, well-defined features and large feature sizes could make use of a number of different kernels and acquisition functions and yield very similar results whereas a small feature which is highly local may see vastly different results with a change in the kernel or acquisition function.

Our proposed workflows, outlined in figure 2, are comprised of three different pathways for the automation of data, with varying degrees of human supervision: (a) sliding window FFT, dimensionality reduction via SVD, non-NMF or N-FINDR [62], component selection, BO (green line), (b) sliding window, rVAE, latent variable selection, BO, (green to orange line) and (c) DCNN, rVAE, latent variable selection, BO (purple line). Note that rVAE analysis can be performed with and without atom finding. These workflows



**Figure 2.** Workflow for the training and discovery phase for automated experiments with three different methods: human driven (green), partial human drive (orange), and computer driven (purple).

vary in effectiveness depending on the feature length and are performed on three datasets with various feature lengths.

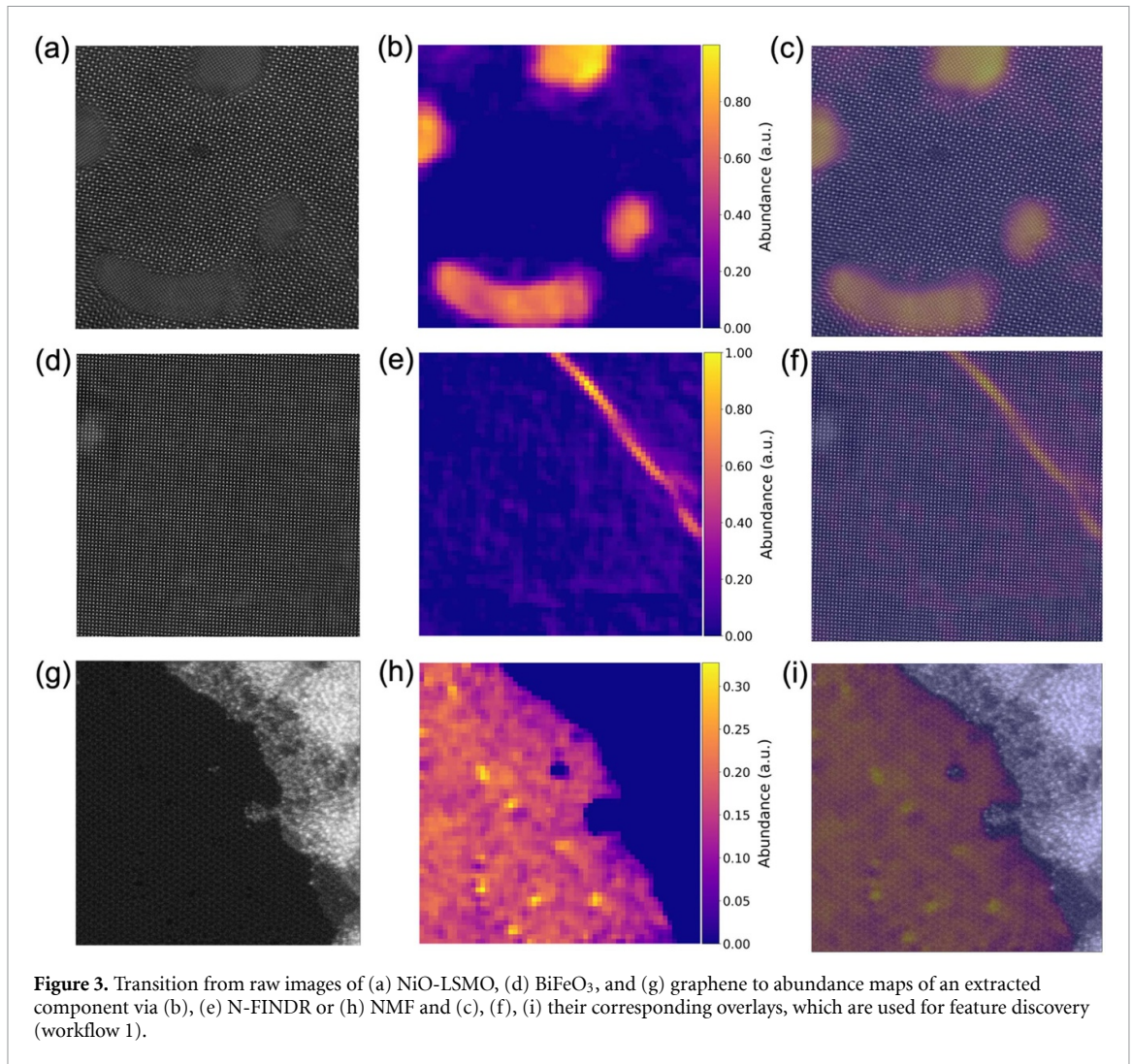
### 3. Discussion

The selection of a workflow will be highly dependent on the individual dataset; however, some trends of each workflow have been identified. Throughout all three datasets, workflow 1 and 3 consistently identified features of interest. The use of rVAE was found to be less reliable as the existence of multiple minima hindered convergence of the network.

Figure 3 highlights the component selection process for NiO-LSMO (a)–(c), Sm-BiFeO<sub>3</sub> (d)–(f), and graphene (g)–(i) using a primarily human driven workflow, workflow 1. In this figure, the second column, figures 3(b), (e), and (h), reveals the abundance of the chosen component and the third column, figures 3(c), (f), and (i), provide the overlay of the component abundance and the raw image. The selection of a component which is representative of the feature of interest is key as this component will inform the BO process. In workflow 1, the image decomposition into individual components is performed using N-FINDR [62] for NiO-LSMO and Sm-BiFeO<sub>3</sub>, whereas NMF is used for graphene. It was noted that N-FINDR becomes less useful for the selection of small features, in which case a simpler image decomposition such as SVD or NMF can be used. Although the use of Bayesian unmixing, SVD, and NMF are simple decomposition techniques, they appear to be robust for feature isolation; however, human decision is necessary for the selection of technique, window size, and eventual component selection.

A similar process is performed using a sliding window FFT and rVAE, workflow 2, but the component extraction is less robust than the simple image decomposition techniques used in workflow 1. The lack of robustness stems from the lack of network convergence, as the latent space contains numerous minima and prevents full convergence. In datasets where rVAE convergence was obtained, the latent variables were not always directly interpretable. Figure 4 outlines the variable outputs from trained rVAE networks where in each case, the input data was obtained via a sliding window FFT. In simple datasets, such as the NiO-LSMO seen in figure 4(a), the latent variable clearly separates the NiO pillars from the LSMO matrix and thus the use of the latent variable seen in figure 4(b) would provide easy identification of the different phases. On the other hand, the latent variable for graphene, seen in figure 4(e), is less interpretable. Overlays of the two datasets, figures 4(c) and (f), clearly show the distinction between the interpretability of the latent spaces. As such, the use of this particular workflow would best be suited for data with large feature sizes or well-separated features.

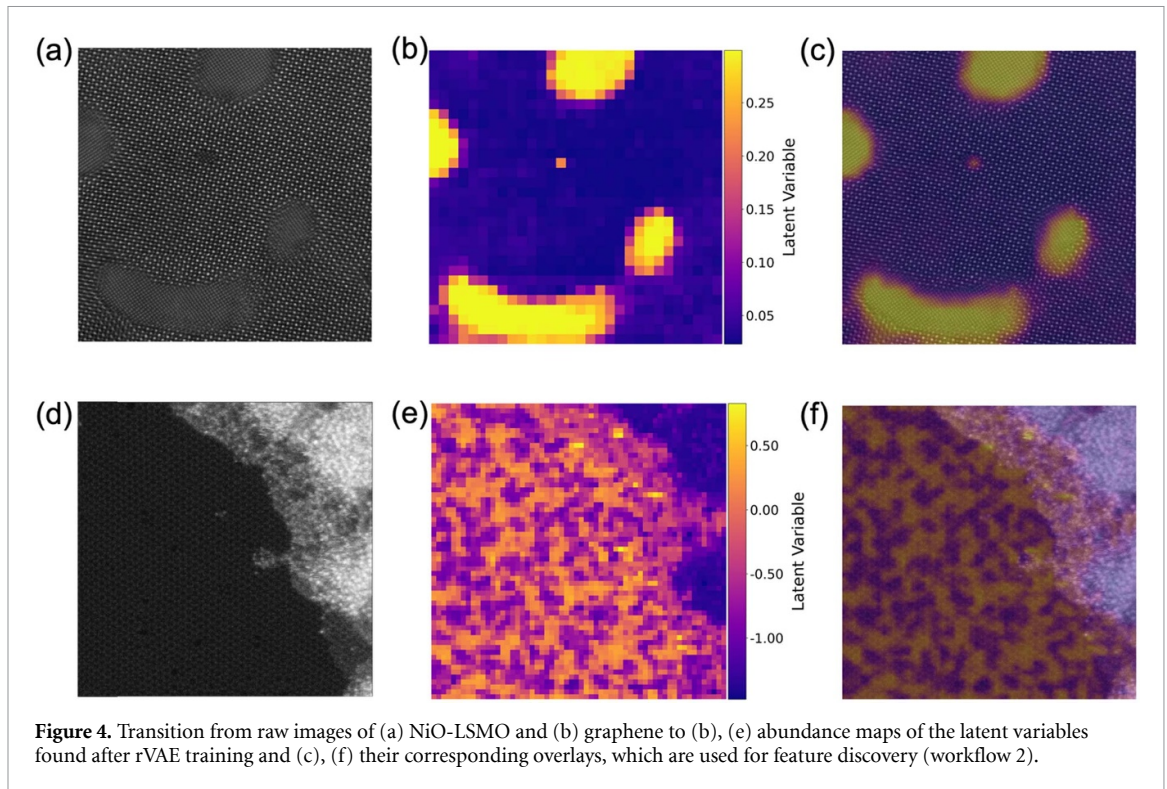
To obtain atomic level precision and realize semi-supervised automation, we utilized a DCNN for atom finding prior to use of the rVAE. As previously discussed, a pre-trained DCNN is implemented on a collected image in order to obtain each atomic column position. A sliding window is used to create a stack of images, each centered on an atomic column, which is then used for training of the rVAE network. Figure 5 reveals the



latent variables found for each of the datasets used throughout this work. Here, the latent variables highlight clearly the separated phases of NiO-LSMO and the ferroelectric domain wall of Sm-BiFeO<sub>3</sub>. The latent variable for graphene is less obvious though the encoded angle derived from rVAE is able to separate the two trimer orientations in graphene. The use of the DCNN and rVAE for feature identification is shown to be useful for both large and small features. Though the use of a much simpler workflow, such as that of the sliding FFT window and NMF decomposition, may be of a more practical use for datasets with large feature sizes.

After obtaining and selecting a component or latent variable, it is used as the target for BO. Data points are collected according to the parameters within the BO algorithm and the image reconstruction and uncertainty is calculated using the data collected thus far. Figure 6 details the BO process for the NiO-LSMO system using the unsupervised workflow of a pre-trained DCNN and latent variable selection via rVAE. As these workflows are performed outside of a microscope, the latent variable map is used as the data feedback and reduced to a sparse grid of one data point to start the BO algorithm, alternatively, in a live experiment each point would be the location of the electron beam and subsequent data collection point. Additionally, to speed up the data collection, batches of 10 points are used for each step of the BO process, such that at the 10th step, 101 data points have been collected. Upon implementing BO and the collection of 200–300 points, we are on average able to reconstruct the full original image with a low degree of uncertainty. Furthermore, most points collected are performed in the regions of interest. This decrease in collection time and collection of just around the feature of interest decreases the total beam time the sample is exposed to, ultimately decreasing the total beam damage. The precise number of points is highly dependent on multiple parameters including the feature size, the acquisition function and the kernel.

We note that the acquisition function can have multiple close to degenerate maxima, particularly in the cases when large regions of the image are uniform. While almost equivalent from the BO perspective, these maxima can differ strongly from the microscope perspective, i.e. a certain measurement sequence can be



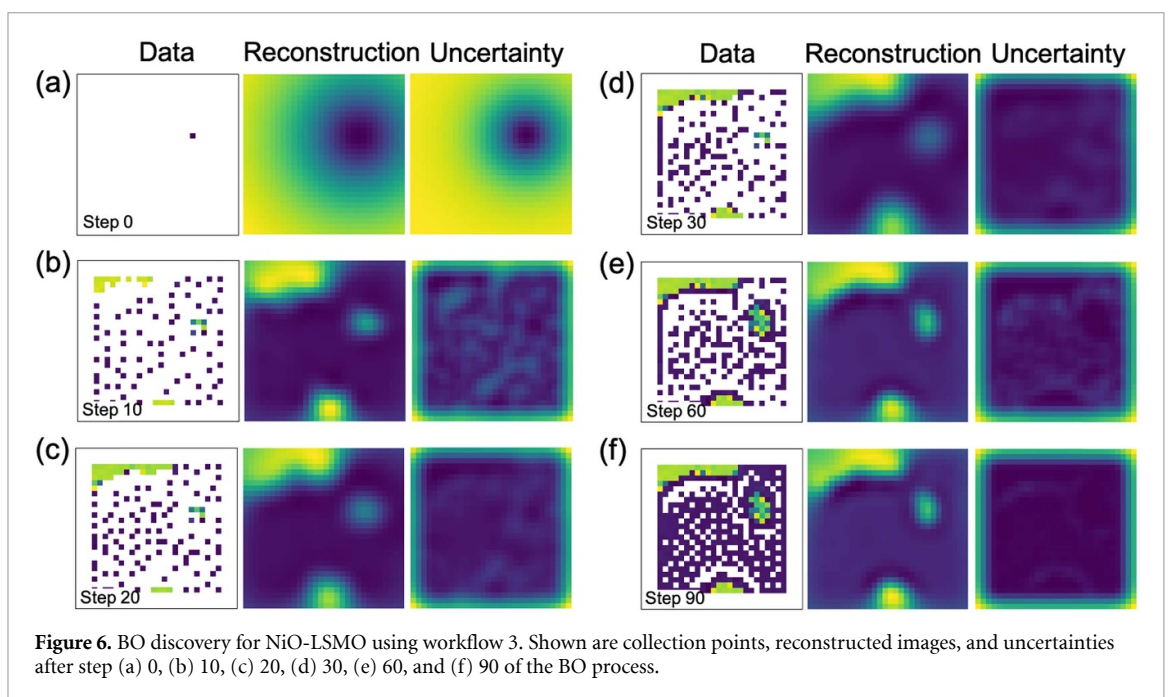
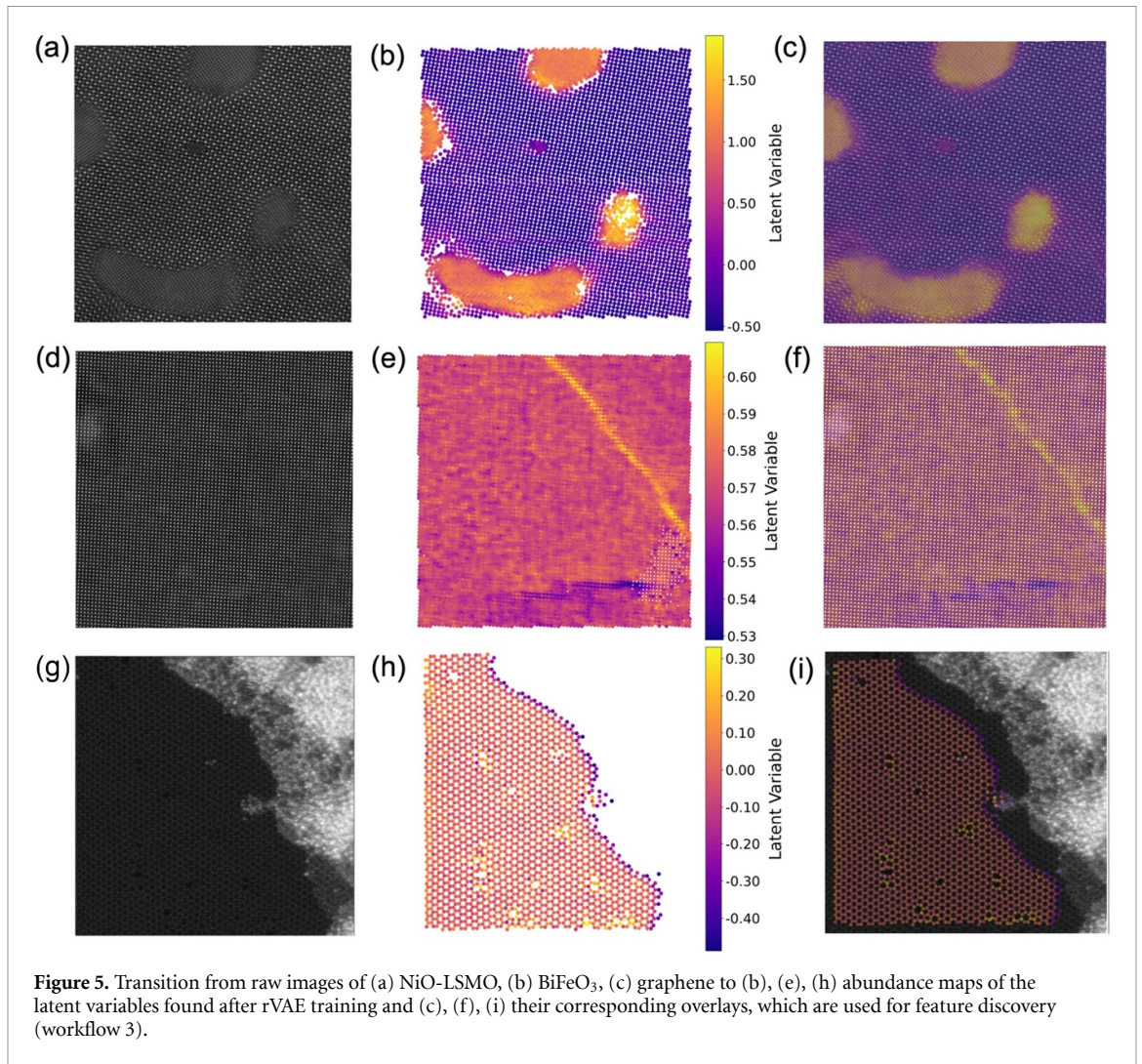
preferential. For example, measurements of the near regions are easiest since probe motion is minimized, whereas scanning between opposite regions of the image field can be sub-optimal. That said, it is obviously sub-optimal to scan very close to already-scanned regions. Practically, this means that we need microscope-specific pathfinder functions that will allow balancing BO-derived points and microscope optimal performance.

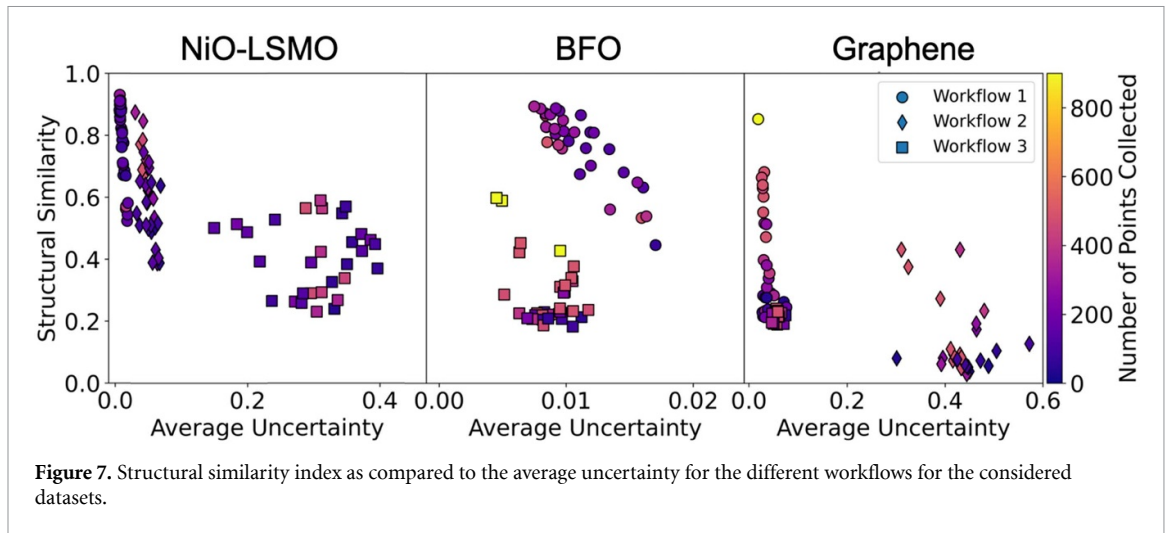
These considerations suggest that a strategy is needed to configure the space. While multiple strategies can be explored, here we suggest a change in the correlation length (d), the kernel type and the acquisition function parameters to strike a balance between exploration and exploitation. For each dataset considered here, combinations of correlation length, kernel type and acquisition function were explored. We note that the correlation length will be highly dependent on the size of the data and the feature of interest size. In this work, the correlation length was considered between 1 and 20, the kernels consisted of RBF, Matern52, and Rational Quadratic, and the acquisition functions were varied between confidence bound, expected improvement, and probability of improvement. To evaluate the ability of full image reconstruction using the sparse dataset of features of interest, a structural similarity index measure (SSIM) is calculated at each parameter combination between the target data and the reconstructed image. Figure 7 outlines the SSIM and average GP uncertainty at the point of diminishing returns, where the collection of more data points does not improve the SSIM or the image reconstruction. While the full range of combinations have not been explored, those explored here provide insight into combinations which might work well for similar data sets.

From the BO parameter combinations explored thus far, we can observe that there is a wide spread of results for each dataset, though workflow 1 appears to be robust no matter the parameters used or the type of data. This robustness is lacking in the other two workflows, seemingly due to the use of rVAE and the subsequent lack of network convergence. As workflow 1 does not require any network training, the performance in finding features of interest is steady throughout different types of data. However, the limitation appears in the level of quantification, as the sliding window FFT lacks the ability to obtain atomic level precision. When a dataset contains a feature of interest which is on the scale of single atoms, such as graphene defects or ferroelectric domain walls, workflow 3 performs adequately. Although the SSIM for workflow 3 is primarily below 0.5, the collection of data points is largely within the features of interest. The decrease in SSIM appears within the image reconstruction of small features, leading to a reconstruction which does not match the structural pattern of the target data.

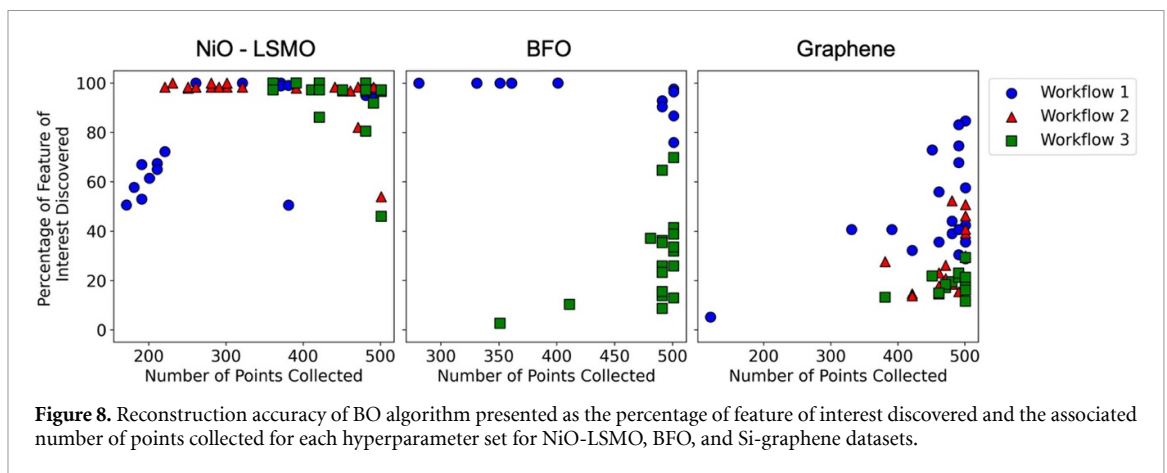
The efficiency of each workflow is evaluated by finding the percentage of the feature of interest which is discovered at each BO step. The step at which the percentage is maximized is outlined in figure 8, with each data point in a workflow and dataset representing a different set of hyperparameters. It is shown that workflow 1 has a robust nature in which, in most cases, over half of the feature of interest was discovered







**Figure 7.** Structural similarity index as compared to the average uncertainty for the different workflows for the considered datasets.



**Figure 8.** Reconstruction accuracy of BO algorithm presented as the percentage of feature of interest discovered and the associated number of points collected for each hyperparameter set for NiO-LSMO, BFO, and Si-graphene datasets.

regardless of the hyperparameters used or the feature of interest size. Workflows 2 and 3 are highly variable in their efficiency, as the hyperparameters and the convergence of rVAE training can highly effect the outcome. However, given a large feature size or a well-defined feature, such as the NiO pillars in LSMO, workflow 2 can be highly efficient, with 99% of the feature of interest being discovered in fewer than 200 points. Further exploration and calibration of the hyperparameters would produce similar results to those seen in NiO-LSMO data, with highly efficient data collection and minimal user oversight.

Besides calibration of the workflows, the implementation on the microscope may present some complications. First, the use of a pre-trained DCNN may not work if the collected image is not within the distribution of data used to train the network. If this occurs, the network will need to be re-trained, or the data would need to be transformed to lie within the training distribution. Secondly, the collection of data within the BO process must be converted in a similar manner as the latent variable used. This could be a simple FFT transform or a more complicated process to obtain the latent variable obtained via rVAE. However, the manipulation of data can be trivial if the software environment which controls the microscope is open source such as NION SWIFT [47] and PyJem [48]. Furthermore, given that the software environment is accessible, the manipulation of the electron beam to specific locations determined via BO is trivial however there is some limitation in the speed of data collection when solely using the software environment. Nonetheless, the use of a field-programmable gate array connected directly to the microscope can bypass the software to enable the collection of data at a much faster rate and opens the avenue for real-time *in situ* processing.

#### 4. Conclusion

To summarize, broad incorporation of automated experiments in STEM and SPM necessitates the development of machine learning algorithms that will guide the selection of the objects of interest. Given the specificity of STEM imaging, these approaches should take into account the generic features of the system, such as the presence of atomic columns in atomically resolved images or specific features in mesoscale

images, and account for the presence of characteristic structural elements and patterns, such as phase separation and ferroelectric domains.

Here, we explored the workflows for AE in STEM which combined DCNN semantic segmentation and rVAE encoding in addition to a lower-level segmentation using sliding window FFT. On the training stage, the network defines and selects a specific parameter to explore, defined as the latent variable or component, while on the automated experimental stage, Bayesian optimization rapidly explores the sample space for features of interest. We note that the workflows can be made much faster, or we can obviate the training stage entirely via transfer learning, when the initial scans are used both for guiding the exploration strategy and training rVAE.

Overall, the use of a specific workflow is dependent of the dataset and feature of interest. Workflow 1, consisting of a sliding window and image deconvolution, is robust and ideal for well-defined features of interest, while workflow 3, consisting of DCNN and rVAE, is best used for atomic scale features. For real deployment, there is a need to also introduce pathfinding functions, which balance the BO gain and the microscope scanning. These workflows can be generalized for different systems depending on the training of the DCNN. In addition, parameter tuning and main network training would need to be performed *a priori* to data collection, as they are time-consuming; however, fine-tuning of the DCNN network could be performed within the workflow. Nevertheless, the workflows developed provide a gateway for rapid feature discovery, limited beam damage, and the collection of data primarily isolated within features of interest.

## Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: [https://github.com/nccreang/AE\\_Workflows](https://github.com/nccreang/AE_Workflows).

## Acknowledgments

This effort (ML and STEM) is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), Materials Sciences and Engineering Division (OD, SVK) and was performed and partially supported (RVK, MZ) at the Oak Ridge National Laboratory's Center for Nanophase Materials Sciences (CNMS), a U.S. Department of Energy, Office of Science User Facility. We thank Jacob Swett for preparing and providing the graphene sample, Wenrui Zhang and Gyula Eres for the preparation and providing the NiO-LSMO sample, Matthew Chisholm for acquiring the NiO-LSMO STEM data, Ziaohang Zhang and Ichiro Takeuchi for preparing and providing the Sm-BFO sample and Chris Nelson for acquiring the Sm-BFO STEM data.

## ORCID iDs

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

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

Maxim Ziatdinov  <https://orcid.org/0000-0003-2570-4592>

## References

- [1] Pennycook S J 2008 *Advances in Imaging and Electron Physics* vol 153, ed P W Hawkes *et al* (Cambridge: Elsevier Academic Press Inc) p 327
- [2] Pennycook S J, Varela M, Lupini A R, Oxley M P and Chisholm M F 2009 Atomic-resolution spectroscopic imaging: past, present and future *J. Electron Microsc.* **58** 87–97
- [3] Jiang Y *et al* 2018 Electron ptychography of 2D materials to deep sub-angstrom resolution *Nature* **559** 343
- [4] Yankovich A B, Berkels B, Dahmen W, Binev P, Sanchez S I, Bradley S A, Li A, Szlufarska I and Voyles P M 2014 Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts *Nat. Commun.* **5** 4155
- [5] Krivanek O L *et al* 2010 Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy *Nature* **464** 571–4
- [6] Jia C L, Urban K W, Alexe M, Hesse D and Vrejoiu I 2011 Direct observation of continuous electric dipole rotation in flux-closure domains in ferroelectric Pb(Zr,Ti)O<sub>3</sub> *Science* **331** 1420–3
- [7] Varela M *et al* 2004 Spectroscopic imaging of single atoms within a bulk solid *Phys. Rev. Lett.* **92** 095502
- [8] Borisevich A Y, Lupini A R and Pennycook S J 2006 Depth sectioning with the aberration-corrected scanning transmission electron microscope *Proc. Natl Acad. Sci. USA* **103** 3044–8
- [9] Ishikawa R, Pennycook S J, Lupini A R, Findlay S D, Shibata N and Ikuhara Y 2016 Single atom visibility in STEM optical depth sectioning *Appl. Phys. Lett.* **109** 163102
- [10] Idrobo J C, Lupini A R, Feng T, Unocic R R, Walden F S, Gardiner D S, Lovejoy T C, Dellby N, Pantelides S T and Krivanek O L 2018 Temperature measurement by a nanoscale electron probe using energy gain and loss spectroscopy *Phys. Rev. Lett.* **120** 095901
- [11] Zeltmann S E, Müller A, Bustillo K C, Savitzky B, Hughes L, Minor A M and Ophus C 2020 Patterned probes for high precision 4D-STEM bragg measurements *Ultramicroscopy* **209** 9

- [12] Campanini M, Nasi L, Albertini F and Erni R 2020 Disentangling nanoscale electric and magnetic fields by time-reversal operation in differential phase-contrast STEM *Appl. Phys. Lett.* **117** 4
- [13] Campanini M, Eimre K, Bon M, Pignedoli C A, Rossell M D and Erni R 2020 Atomic-resolution differential phase contrast STEM on ferroelectric materials: a mean-field approach *Phys. Rev. B* **101** 12
- [14] Ferullo V et al 2014 Generation of nondiffracting electron Bessel beams *Phys. Rev. X* **4** 7
- [15] Rotunno E, Tavabi A H, Yucelen E, Frabboni S, Dunin Borkowski R E, Karimi E, McMorran B J and Grillo V 2019 Electron-beam shaping in the transmission electron microscope: control of electron-beam propagation along atomic columns *Phys. Rev. Appl.* **11** 10
- [16] Ishikawa R, Mishra R, Lupini A R, Findlay S D, Taniguchi T, Pantelides S T and Pennycook S J 2014 Direct observation of dopant atom diffusion in a bulk semiconductor crystal enhanced by a large size mismatch *Phys. Rev. Lett.* **113** 155501
- [17] Yang Z Q, Yin L, Lee J, Ren W, Cheng H-M, Ye H, Pantelides S T, Pennycook S J and Chisholm M F 2014 Direct observation of atomic dynamics and silicon doping at a topological defect in graphene *Angew. Chem., Int. Ed.* **53** 8908–12
- [18] Lin J H et al 2014 Flexible metallic nanowires with self-adaptive contacts to semiconducting transition-metal dichalcogenide monolayers *Nat. Nanotechnol.* **9** 436–42
- [19] Susi T, Meyer J C and Kotakoski J 2017 Manipulating low-dimensional materials down to the level of single atoms with electron irradiation *Ultramicroscopy* **180** 163–72
- [20] Dyck O, Kim S, Kalinin S V and Jesse S 2017 Placing single atoms in graphene with a scanning transmission electron microscope *Appl. Phys. Lett.* **111** 113104
- [21] Dyck O, Kim S, Jimenez-Izal E, Alexandrova A N, Kalinin S V and Jesse S 2018 Building structures atom by atom via electron beam manipulation *Small* **14** 1801771
- [22] Jesse S et al 2018 Direct atomic fabrication and dopant positioning in Si using electron beams with active real-time image-based feedback *Nanotechnology* **29** 255303
- [23] Kalinin S V et al 2016 Big, deep, and smart data in scanning probe microscopy *ACS Nano* **10** 9068–86
- [24] Zahl P, Wagner T, Moller R and Klust A 2010 Open source scanning probe microscopy control software package GXSM *J. Vac. Sci. Technol. B* **28** C4E39
- [25] Krull A, Hirsch P, Rother C, Schiffrin A and Krull C 2020 Artificial-intelligence-driven scanning probe microscopy *Commun. Phys.* **3** 8
- [26] Kalinin S V, Borisevich A and Jesse S 2016 Fire up the atom forge *Nature* **539** 485–7
- [27] Dyck O, Jesse S and Kalinin S V 2019 A self-driving microscope and the atomic forge *MRS Bull.* **44** 669–70
- [28] Swartzentruber B S 1995 Direct measurement of surface diffusion using atom-tracking scanning tunneling microscopy *Phys. Rev. Lett.* **76** 459–62
- [29] Ovchinnikov O S, Jesse S and Kalinin S V 2009 Adaptive probe trajectory scanning probe microscopy for multiresolution measurements of interface geometry *Nanotechnology* **20** 255701
- [30] Requicha A A G et al *IEEE—Nano. (IEEE)* pp 81–86
- [31] Mokaberi B, Yun J, Wang M and Requicha A A G *IEEE Int. Conf. on Robotics and Automation (IEEE)* pp 1406–12
- [32] Requicha A A G, Arbuckle D J, Mokaberi B and Yun J 2009 Algorithms and software for nanomanipulation with atomic force microscopes *Int. J. Rob. Res.* **28** 512–22
- [33] Stevens A, Yang H, Carin L, Arslan I and Browning N D 2014 The potential for Bayesian compressive sensing to significantly reduce electron dose in high-resolution STEM images *Microscopy* **63** 41–51
- [34] Stevens A, Kovarik L, Abellan P, Yuan X, Carin L and Browning N D 2015 Applying compressive sensing to TEM video: a substantial frame rate increase on any camera *Adv. Struct. Chem. Imag.* **1** 1–20
- [35] Sang X, Lupini A R, Unocic R R, Chi M, Borisevich A Y, Kalinin S V, Endeve E, Archibald R K and Jesse S 2016 Dynamic scan control in STEM: spiral scans *Adv. Struct. Chem. Imag.* **2** 1–8
- [36] Li X, Dyck O, Kalinin S V and Jesse S 2018 Compressed sensing of scanning transmission electron microscopy (STEM) with nonrectangular scans *Microsc. Microanal.* **24** 623–33
- [37] Ede J M and Beanland R 2020 Partial scanning transmission electron microscopy with deep learning *Sci. Rep.* **10** 1–10
- [38] Kelley K P, Ziatdinov M, Collins L, Susner M A, Vasudevan R K, Balke N, Kalinin S V and Jesse S 2020 Fast scanning probe microscopy via machine learning: non-rectangular scans with compressed sensing and Gaussian process optimization *Small* **16** 6
- [39] Dyck O, Zhang C, Rack P D, Fowlkes J D, Sumpter B, Lupini A R, Kalinin S V and Jesse S 2020 Electron-beam introduction of heteroatomic Pt-Si structures in graphene *Carbon* **161** 750–7
- [40] Dyck O, Yoon M, Zhang L, Lupini A R, Swett J L and Jesse S 2020 Doping of Cr in graphene using electron beam manipulation for functional defect engineering *Appl. Nano Mater.* **3** 10855–63
- [41] Dyck O, Zhang L, Yoon M, Swett J L, Hensley D, Zhang C, Rack P D, Fowlkes J D, Lupini A R and Jesse S 2021 Doping transition-metal atoms in graphene for atomic-scale tailoring of electronic, magnetic, and quantum topological properties *Carbon* **173** 205–14
- [42] Susi T et al 2014 Silicon-carbon bond inversions driven by 60-keV electrons in graphene *Phys. Rev. Lett.* **113** 115501
- [43] Tripathi M, Mittelberger A, Pike N A, Mangler C, Meyer J C, Verstraete M J, Kotakoski J and Susi T 2018 Electron-beam manipulation of silicon dopants in graphene *Nano Lett.* **18** 5319–23
- [44] Su C et al 2019 Engineering single-atom dynamics with electron irradiation *Sci. Adv.* **5** 2252
- [45] Inani H, Mustonen K, Markevich A, Ding E-X, Tripathi M, Hussain A, Mangler C, Kauppinen E I, Susi T and Kotakoski J 2019 Silicon substitution in nanotubes and graphene via intermittent vacancies *J. Phys. Chem. C* **123** 13136–40
- [46] Park H, Wen Y, Li S X, Choi W, Lee G-D, Strano M and Warner J H 2021 Atomically precise control of carbon insertion into hBN monolayer point vacancies using a focused electron beam guide *Small* **17** 2100693
- [47] Meyer C, Dellby N, Hachtel J A, Lovejoy T, Mittelberger A and Krivanek O 2019 Nion swift: open source image processing software for instrument control, data acquisition, organization, visualization, and analysis using python *Microsc. Microanal.* **25** 122–3
- [48] JOEL PyJEM 2021 GitHub repository (available at: <https://github.com/PyJEM/PyJEM>)
- [49] Zhang W, Cheng S, Rouleau C M, Kelley K P, Keum J, Stavitski E, Zhu Y, Chisholm M F, Gai Z and Eres G 2020 Unusual electrical conductivity driven by localized stoichiometry modification at vertical epitaxial interfaces *Mater. Horizons* **7** 3217
- [50] Kalinin S V et al 2020 Deep bayesian local crystallography (arXiv:2012.07134v07131)
- [51] Ghosh A, Sumpter B G, Dyck O, Kalinin S V and Ziatdinov M 2021 Ensemble learning-iterative training machine learning for uncertainty quantification and automated experiment in atom-resolved microscopy *npj Comput. Mater.* **7** 100

- [52] Ziatdinov M, Creange N, Zhang X, Morozovska A, Eliseev E, Vasudevan R K, Takeuchi I, Nelson C and Kalinin S V 2021 Predictability as a probe of manifest and latent physics: the case of atomic scale structural, chemical, and polarization behaviors in multiferroic Sm-doped BiFeO<sub>3</sub> *Appl. Phys. Rev.* **8** 011403
- [53] Nelson C T, Vasudevan R K, Zhang X, Ziatdinov M, Eliseev E A, Takeuchi I, Morozovska A N and Kalinin S V 2020 Exploring physics of ferroelectric domain walls via Bayesian analysis of atomically resolved STEM data *Nat. Commun.* **11** 6361
- [54] Nelson C T, Zhang W, Takeuchi I and Kalinin S V 2020 Zenodo, Nature Communications
- [55] Ghosh A, Sumpster B G, Dyck O, Kalinin S V and Ziatdinov M 2021 Ensemble learning and iterative training (ELIT) machine learning: applications towards uncertainty quantification and automated experiment in atom-resolved microscopy (arXiv:2101.08449v08442)
- [56] Kalinin S V *et al* 2021 Unsupervised machine learning discovery of chemical and physical transformation pathways from imaging data (arXiv:2010.09196v09192)
- [57] Ziatdinov M, Dyck O, Maksov A, Li X, Sang X, Xiao K, Unocic R R, Vasudevan R, Jesse S and Kalinin S V 2017 Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations *ACS Nano* **11** 12742–52
- [58] Ziatdinov M and Kalinin S V 2021 AtomAI: open-source software for applications of deep learning to microscopy data *Microsc. Microanal.* **27** 3000
- [59] Vasudevan R K *et al* 2021 Investigating phase transitions from local crystallographic analysis based on statistical learning of atomic environments in 2D MoS<sub>2</sub>-ReS<sub>2</sub> *Appl. Phys. Rev.* **8** 011409
- [60] Vasudevan R K, Belianinov A, Gianfrancesco A G, Baddorf A P, Tselev A, Kalinin S V and Jesse S 2015 Big data in reciprocal space: sliding fast Fourier transforms for determining periodicity *Appl. Phys. Lett.* **106** 091601
- [61] Rasmussen C E and Williams C K I 2005 *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)* (Cambridge: MIT Press)
- [62] Winter M E 1999 N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data *Proc. SPIE 3753, Imaging Spectrometry V* pp 266–75