**Machine Learning in Scanning Transmission Electron Microscopy**

Sergei V. Kalinin,[1,*] Colin Ophus,[2] Paul Voyles,[3] Rolf Erni,[4] Demie Kepaptsoglou,[5,6] Vincenzo Grillo,[7] Andrew R. Lupini,[1] Mark P. Oxley,[1] Eric Schwenker,[8,9] Maria K. Y. Chan,[8] Joanne Etheridge,[10] Xiang Li,[11] Grace G. D. Han,[11] Maxim Ziatdinov,[1,12] Naoya Shibata,[13] Stephen J. Pennycook[14,15]

[1] Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN 37831

[2] NCEM, Molecular Foundry, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA, USA 94720

[3] Department of Materials Science and Engineering, University of Wisconsin-Madison, 1509 University Ave, Madison, WI 53706 USA.

[4] Electron Microscopy Center, EMPA – Swiss Federal Laboratories for Materials Science and Technology, Ueberlandstrasse 129, CH-8600 Dubendorf, Switzerland

[5] SuperSTEM laboratory, Sci-Tech Daresbury Campus, Daresbury, WA4 4AD, UK

[6] Department of Physics, University of York, Heslington York, YO10 5DD, UK

[7] CNR-NANO, via G Campi 213/a, I-41125 Modena, Italy

[8] Center for Nanoscale Materials, Argonne National Laboratory, Lemont IL 60439

[9] Materials Science and Engineering Department, Northwestern University, Evanston IL 60208

[10] Monash Centre for Electron Microscopy and Department of Materials Science and Engineering, 10 Innovation Walk, Monash University, Victoria, 3800 Australia

[11] Department of Chemistry, Brandeis University, 415 South Street, Waltham, MA 02453, USA

[12] Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

[13] Institute of Engineering Innovation, School of Engineering, The University of Tokyo, Yayoi 2-11-16, Bunkyo-ku, Tokyo 113-8656, Japan

[14] Department of Materials Science and Engineering, University of Tennessee, Knoxville, TN, 37996

[15] School of Physical Sciences and CAS Key Laboratory of Vacuum Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

* Corresponding Author sergei2@ornl.gov

## Abstract

Scanning transmission electron microscopy (STEM) has emerged as a uniquely powerful tool for structural and functional imaging of materials on the atomic level. Driven by advances in aberration correction, STEM now allows routine imaging of structures with single digit picometer-level precision for localization of atomic units. This Primer focuses on the opportunities emerging at the interface between STEM and machine learning (ML) methods. We review the primary STEM imaging modalities including structural imaging, electron energy loss spectroscopy and its momentum resolved modalities and 4D-STEM. We discuss the quantification of STEM structural data as a necessary step towards meaningful ML applications and its analysis in terms of relevant physics and chemistry. We show examples emphasizing the new opportunities offered by structural STEM imaging in elucidating the chemistry and physics of complex materials and how the latter connect to first-principle and phase-field models to yield consistent interpretation of generative physics. We present the critical infrastructural needs for the broad adoption of ML methods in the STEM community including the storage of data and meta data allowing the reproduction of experiment. Finally, we discuss the application of machine learning to automate experiment and enable novel scanning modes.

## [H1] Introduction

One of the most common methods of "seeing" atoms is scanning transmission electron microscopy (STEM). In STEM, images are generated by scanning a small probe, formed by focusing the electron beam, across a thin sample.[1] The probe is focused by the objective lens, commonly combined with other lenses designed to reduce the aberrations of the probe-forming optics. Scan coils, which generate a rapidly changing field, are used to scan the beam over the sample. Any signal that varies with probe position can be detected and used to form an image. A simple schematic of a STEM and various possible detector configurations is shown in Figure 1 (left). Some detector configurations can be used in parallel, which is advantageous because it allows multiple, complementary signals to be directly compared, pixel by pixel, both for imaging of simultaneous coherent and incoherent signals and for spectroscopic analysis such as X-ray and light emission and electron energy loss spectroscopy (EELS). The detected intensity is plotted on a monitor synchronized with beam scanning to form images. Because the distribution of transmitted and/or scattered electrons in the detector plane depends on the specimen thickness, composition and structure, the variation of detected intensity across the image tells us how the projected thickness, composition, and/or structure of the sample changes from point to point. The spatial resolution is effectively determined by the size of the electron probe. Today, the development of aberration correction technology for electron optics allows an electron probe to be routinely be narrowed to 50 pm or less full width half maximum (FWHM), meaning that the spatial resolution of aberration-corrected STEM has become smaller than the size of a hydrogen atom whose Bohr radius is approximately 53 pm. In a STEM image, the image contrast and the information it contains about the specimen depends on which region(s) of the detector plane we record from and how we detect the electrons by the post-specimen detectors.

The most common STEM measurement is to detect the electrons scattered to high angles (significantly beyond the probe-forming aperture radius) by a donut-shaped detector. Called annular dark field (ADF) imaging, the images obtained by this method are incoherent – essentially, all probed atoms contribute independently and additively to the total signal and interference effects between signals arising from different atoms are suppressed. This feature enables us to directly and robustly observe atomic columns and single atoms in materials and devices. Moreover, since the integrated intensity of high-angle scattered electrons varies approximately with the square of the atomic number $Z$ of the atoms under the probe, $Z$-contrast imaging can visualize the heavier element atoms in a crystalline lattice with high sensitivity. In contrast, the signal from the lighter atoms in a crystalline lattice is generally much lower, often making detection difficult. To increase the atomic signal and image light monolayer materials such as graphene, the inner angle of the detector can be decreased. This commonly used detection method is called annular bright-field (ABF) imaging. ABF mode uses a donut-shaped detector positioned to collect electrons scattered inside the angular-range defined by the incident probe (often referred to as the bright-field disk) and can directly visualize light atoms even in the presence of heavy atoms. Since ADF, ABF, and other type of images can be obtained simultaneously from exactly the same sample positions by using either many separate segmented detectors or a pixel-type detector (see Fig 1 (left)), both structural and chemical information at atomic dimensions can be obtained by STEM. When combining imaging with spectroscopic techniques, STEM allows us to extract multiple and complementary forms of information about a specimen from a very localized volume including its atomic dimensions. For example, imaging can be combined with electron energy loss spectroscopy (EELS), where each probe position records the energy spectrum of scattered electrons to measure inelastic scattering inside the specimen, and energy-dispersive X-ray spectroscopy (EDS or EDX), which measures X-rays emitted as a result of the electron probe interacting with the specimen, to simultaneously obtain a wealth of material information from the same sample positions.

In recent years, additional imaging modes, such as differential phase contrast and 4D-STEM as discussed below, have become available with advances in electron detectors, especially pixelated and multi-segmented detectors, which allow spatially-resolved probing of electric fields at atomic resolution and magnetic fields at nanoscale resolution.[2-4] Pixelated detectors allow a snapshot of electrons scattered through a range of angles for each probe position. This enables the detector configuration, for example ABF, ADF or even various multi-segmented configurations, to be selected after the data have been collected, with obvious advantage of selecting the imaging mode optimal for the sample without the need to do additional experimentation. They also permit special configurations to be selected such as for ptychography.[5] Increasing detector speeds offer significant advantages for low-dose imaging compared to conventional phase contrast imaging. Electric and magnetic field imaging also benefits from fast pixelated detectors. Improvements in aberration-correction and spatial coherence have led to increasingly large probe-forming apertures that enhance depth resolution to the nanometer scale, allowing optical sectioning and yielding 2D views of the sample at atomic resolution and at different depths.[6-10] In the future, optical sectioning by using even larger probe-forming apertures could yield atomic resolution as a function of depth, albeit with some interpretative complexity.[11,12] The more commonly-used tilt series tomography has already achieved atomic resolution in all three dimensions[13], while combined approaches have also been proposed.[14-16] Spectroscopy has also undergone a revolution in recent years with improved monochromator and electron emitter designs increasing energy resolution to the ~meV level[17] and allowing both phonon spectroscopy and direct measurement of sample temperature.[18]

The rapid advancement in STEM imaging and spectroscopy raises issues such as how to explore multidimensional data sets using a human operator and what type of quantitative information can be obtained from the measured data. Some questions include what materials-specific information can be obtained from microscopy data and at what level of confidence, for example when determining atomic coordinates from STEM or scattering potentials from 4D-STEM; how the imaging system affects materials-specific measurements; and whether materials-specific measurements can be improved with better knowledge of the system (such as knowing the beam parameters, the resolution function, or fully modelling the imaging system) and knowledge of a material's phase, structure, and composition. Questions remain about whether we can use materials-specific information with uncertainties determined by incomplete knowledge of the imaging system or intrinsic limitations to infer physics and chemistry. A common approach is the use of correlative models, where the variation of observables is compared. Without an understanding of the physical mechanisms involved, observed correlations do not necessarily indicate a causative relationship between the observed parameters. An alternative approach is generative statistical models which model the joint probability distribution between the observable variables and the target variables. The aim of generative statistical models is the recovery of generative physics such as force fields, exchange integrals, and other parameters. An example of a deep generative model is the variational auto encoder (VAE) which will be discussed in detail below. The question also remains as to whether correlative or causative materials-specific information can be used to reconstruct materials behavior in a broad parameter space such as, for example, phase diagrams via different temperatures and concentrations, and finally determine how the reliability of such predictions depends on the position in the parameter space. Similarly, we can harness the data stream from the microscope to engender real-time feedback, for example for automated experimentation and nanometer- and atomic scale matter manipulation.

Machine learning (ML) is becoming integral to answering these questions. In Fig 1 we illustrate a possible workflow for ML analysis of experimental data as it is acquired. The first

component is a local computer connected directly to the microscope acquisition computer, often referred to as an edge computer. This computer downloads the data from the microscope computer and carries out initial processing of the data to allow feature extraction and quantification of the datasets, usually using pretrained models. This direct connection is required because of the size of the datasets commonly acquired, especially 4D-STEM datasets. The edge computer is also connected to the local network and via this the internet. This allows access to web based resources such as online storage and Google Colab. However, the slower bandwidth provided by even commercial internet connections will be a bottle bottleneck preventing rapid retraining of models based on new datasets. The edge computer can also be connected via the local network to high performance computing (HPC) facilities. Here the data transfer will be rapid, but the calculations may be time consuming. These simulations may include structure inversion to determine the scattering potential, structural relaxations using density functional theory (DFT) or molecular dynamics (MD) or the determination of structural changes associated with quantum phenomena. Finally, the edge computer can use the analysis it made of the initial results to adjust the microscope parameters. This needs to be done rapidly to avoid specimen drift and other microscope instabilities. This will require the development of rapid algorithms and the ability to interact directly with the microscopes operating software. Multivariate statistical methods have become a technique of choice for visualizing multidimensional hyperspectral data sets.[19-21] Deep neural networks allow for robust image analysis with significant benefits for automatic location and classification of defects.[22,23] We may soon see a self-driving microscope,[24] potentially integrated with synthesis and computation to develop new materials. The microscope may even be able to place dopants into predetermined locations for an ultimate atom-by-atom assembly of materials.[25-27] For some common machine learning approaches, see Box 1.

In this primer, we first discuss the fundamental principles of imaging and spectroscopy in STEM for structural imaging, EELS and its momentum-resolved modalities and 4D-STEM, and highlight the need for ML methods as a way to analyze emerging multidimensional data sets. A more detailed description of modern STEM theory and experimentation can be found elsewhere.[28] We then discuss the quantification of structural STEM data in terms of atomic positions and the insights it yields into the physics and chemistry of materials. We emphasize the novel opportunities enabled by the synergy of correlative ML methods and physics-based ML and STEM as a way to link observations in STEM experiments to generative physical models of materials behavior. Finally, we discuss the associated challenges and scientific infrastructural needs of implementing ML with STEM before looking forward to the requirements for open-source software and data sharing within the community to facilitate the application of ML methods to STEM data.

## [H1] Experimentation

In this section we briefly outline the major components of a STEM, with an emphasis on the parts we believe are likely to be most relevant for machine learning.

### [H2] STEM fundamentals

A STEM is closely related to both a conventional transmission microscope and a scanning electron microscope. A beam of electrons is extracted from a source, also called a gun or a tip, and accelerated by a high-voltage typically between 30 and 300 kV. To avoid the beam colliding with gas molecules, the interior of the microscope has to be under high vacuum. Much of the support equipment surrounding a modern STEM will be associated with maintaining the

vacuum, high-voltage and other power supplies. In a modern STEM, the ability to image single atoms places extreme requirements on the cleanliness of the vacuum system and sample in a similar way to conventional surface science methods[29-31]. A dedicated sample holder is needed to position and tilt the sample, which is inserted into the microscope through an airlock.

Electron lenses, typically round electromagnets that generate an intense field on the beam axis, are used to shape the electron beam because electrons are charged particles affected by electric and magnetic fields. As seen in Fig 1, a series of condenser lenses first shape and demagnify the beam of electrons. The beam is then converged onto the sample by an objective lens to form a probe. The final size of the probe is limited by diffraction and by the aberrations of the electron lenses. The diffraction limit[32] depends on the convergence angle of the probe and the De Broglie wavelength of the electrons, which is determined by the accelerating voltage. As in optics, a larger convergence angle (which is equivalent to a larger probe forming aperture) corresponds to a finer resolution. However, unlike in conventional light optics where, in principle, arbitrary lens shapes may be aberration-free, conventional round electron lenses always have some aberrations[33]. These aberrations become worse at higher convergence angles, which limits the largest aperture size that can be used. Reducing or eliminating these aberrations via aberration correction has been an enduring challenge in electron microscopy (see [34] for a comprehensive history of aberration-correction). The first successful modern aberration-correctors in STEM[35,36] and in TEM[37-40] saw broad uptake during the early 2000s.[41-43]. The most relevant aspect of these correctors for the present Primer is that they depart from rotational symmetry by using non-round lenses and consist of multiple elements, which makes them complicated to use. The development of advanced computer controls to measure the aberrations and optimize their conditions is therefore an essential part of the operation of these systems[35]. With careful optimization of the lens settings, the probe size at the sample can now be about the size of an atom.

A series of projector lenses transfers the electrons that are transmitted through the sample to various detectors or an EEL spectrometer. The intensity measured by the detectors, which is equivalent to the number of electrons hitting the detectors, forms images as a function of the probe position. This means that the image points are acquired in series over time and a variety of detectors can be used to record different signals for each probe position. Note that the STEM mode of operation is different from a conventional TEM where a larger sample area is continuously illuminated by the beam and the projector lenses magnify the image of the sample.

The electrons scattered out to high angles strongly depend on the atomic number $Z$ of the elements and the signal recorded on a high-angle annular dark field detector (HAADF) is known as a $Z$-contrast image. The benefit of the HAADF mode is that it allows relatively simple image interpretation to a good approximation. In many cases, the bright spots on a $Z$-contrast image relate to the positions of the atoms and spot intensity relates to atomic number in an intuitive way. Finding atomic positions therefore depends on locating local maxima and image intensity provides clues as to the number of atoms or their atomic numbers. This simplicity is particularly convenient for ML because initial testing can often be done with very simplified models and training data can typically be classified by human operators.

A critical aspect of STEM experiments is sample preparation. The electron beam has a mean free path for scattering measured in tens of nanometers for most materials, which requires very thin specimens to produce high quality measurements in transmission modes. Samples can be thinned using a variety of methods ranging from mechanical polishing to focused ion beam (FIB) milling. The thinness of the sample causes most of the beam electrons to remain near the optical axis and the electron interaction with the sample primarily changes the phase of the beam. A small detector close to the axis gives a bright field (BF) image. One of the key difficulties in BF imaging is that detectors record the intensity whereas most of the details of

the beam-sample interaction are contained in the phase. To obtain a phase contrast BF image, a small amount of defocus is typically used to give an extra phase shift and cause the total intensity at the detector to depend on the phase.[28] An annular detector, typically with a bigger diameter, yields an ABF image. The type and size of detector used will affect the information contained in the image.

It is also possible to used segmented detectors to record finer angular ranges. A very common example is the use of four or more segments arranged as pie slices on the unscattered electron probe in order to measure beam shifts in differential phase contrast.[44,45] A more recent addition is the use of high-speed direct electron detectors to record a full image of the diffracted electron beam and generate a four-dimensional dataset via 2D images of the probe scanned over a 2D grid of probe positions, a technique often referred to as 4D-STEM.[46]

An additional reason for requiring thin specimens is to aid interpretation of even the most basic imaging modes. This is because the probe electron undergoes multiple scattering, commonly referred to a dynamical scattering in electron microscopy, as it propagates through the specimen. This causes the probe to change shape depending on the local sample structure. For example, a probe placed above an atomic column tends to travel down along the column, a phenomenon often referred to as channeling. For modern high-resolution STEM, the distance the probe channels down the column is crucially defendant on the atomic mass of the atoms in the column, meaning different columns are illuminated to different extents. This can make direct quantitative interpretation of images complicated for specimens of more than 10nm or so depending on the accelerating voltage used.

## [H2] Electron Energy Loss Spectroscopy

One of the key aspects of using an annular detector to form a STEM image is that only a small fraction of the electrons are scattered to high-angles. The disadvantage is that ADF-STEM can be an inefficient way to form an image; however, the benefit is that on-axis signals can be collected simultaneously from the same position to provide complementary specimen information. As the fast electrons in the beam go through the sample, they interact with and exchange energy and momentum with the sample. These effects are separated into elastic scattering, where the total kinetic energy is conserved, and inelastic scattering, where energy is transferred to the sample in some other form. For example, if the fast beam electron excites a sample electron to a higher-energy state, it loses a corresponding amount of energy, which can be measured with EELS. An excellent introduction to EELS can be found in [47].

The first EEL spectrometers[48] consisted of a magnetic prism followed by several quadrupoles and were subsequently optimized[49]. In an electron spectrometer, a magnetic field is used to disperse the beam dependent on its kinetic energy. A position sensitive detector is used to provide an EEL spectrum. This spectrum will have several peaks at energies that depend on the elements in the sample, with shapes that depend on sample thickness and electronic structure. Various complications in the interpretation arise because the beam electrons can interact with the sample in a variety of ways, and perhaps multiple times, especially for thick samples. Untangling these competing effects could be a promising area for ML techniques.

EELS in the STEM has emerged as a powerful analytical technique; the small electron probe dimensions means that excitations associated with core-loss,[50] plasmon[51] and optical transitions in materials can now be routinely interrogated, in some cases at single atom resolution.[52,53] We typically divide the EEL spectrum into different regions depending on how the fast beam electrons interact with the sample. The zero-loss peak contains electrons that have not lost a measurable amount of energy. Interactions with the atomic vibrations of the material lattice can excite phonons and such measurements are commonly referred to as vibrational spectroscopy. Collective excitation of the sample electrons are referred to as

plasmons , which are closely related to the electrical and photonic properties of a material and depend on material shape and size.[54-56] Energy losses that correspond to excitations of core electrons in the sample are referred to as core-loss and characterize the elements present at the probe location. The detailed shape of the core-loss edges will depend on the characteristics of the initial and final state of the sample electrons as well as many of the same thickness and orientation effects that affect the imaging modes.

Recording a spectrum at each probe position gives a spectrum image, also referred to as a hyperspectral data set.[57,58] The strength of EELS in a STEM is that it provides information on the local chemistry and electronic structure while the *Z*-contrast image simultaneously provides a map of the atomic configuration. This concept can be extended to other imaging modalities and combined with multiple signals. For example, when an excited sample electron relaxes back to its ground state it will emit the excess energy as a photon, which can simultaneously be analyzed with a cathodoluminescence or X-ray detector. A recent review article contains a more comprehensive discussion of these and associated methods.[59]

One of the factors that limits the energy resolution in EELS experiments is that the beam of electrons has a small but finite energy spread. Depending on the type of electron source, the beam width will typically be between 0.3 and 1 eV. Peaks that are separated from each other by less than this intrinsic energy spread cannot be easily resolved, which will mask many physically interesting phenomena. In addition, there are extreme demands on the performance of the microscope power supplies: we might want to measure energy losses of a few meV for a beam energy of several tens (or hundreds) of keV, which corresponds to a stability above 1 part per million. Perhaps even more importantly, the tails of the distribution extend out over a far larger energy range, meaning that very weak signals are easily swamped by the background.[60] Solving the energy spread problem requires an electron monochromator. In a monochromator, energy-dispersive elements disperse the beam before the sample and a range of energies is selected using a slit or an aperture.[61] This selection necessarily reduces the amount of current available to form the probe. Since the slit used to select the energy has a finite size and the monochromator may introduce aberrations, using an electron monochromator typically degrades the spatial resolution. However, a new generation of monochromated instruments has been carefully designed to allow high energy and high spatial resolution while minimizing the loss of current. Just as for aberration-correctors, computer control is usually essential to the operation of electron monochromators.[62]

One of the most exciting aspect of these new monochromated STEMs to date is the exploitation of modes such as aloof spectroscopy, where the beam passes near the sample instead of through it[63-65]. Such modes might open new non-damaging analysis modes or be used to provide data on sensitive biological or complex quantum states. Another surprising example is electron energy gain spectroscopy,[18] where the fast beam electron can gain energy from the sample in a reversal of the normal energy-loss process. Such novel modes are likely to allow access to new information streams about the local nanoscale properties of materials. For example, comparing energy loss and energy gain probabilities can now be used to measure the local temperature of materials in a parameter-free model.[18,66].

## [H2] Momentum Resolved EELS

STEM-EELS experiments are typically performed at small scattering angles with the spectrometer entrance aperture positioned along the optical axis to exclude the collection of electrons with a scattering angle $\vartheta$ larger than the collection semi-angle $\beta$.[67] In the STEM optical geometry where a converged probe is used, the probed signal is dominated by inelastic scattering wave vectors parallel to the forward scattering direction $q_{\parallel}$ and averaged with

contributions from other non-parallel wave vectors determined by the probe convergence semi-angle $\alpha$.[67]

In momentum or angular-resolved EELS measurements (q-EELS), the spectra are acquired as a function of scattering angle or momentum vector $q$. The measurements are performed by displacing the collection aperture relative to the optical axis. This procedure is shown schematically in Figure 2a. In practice, this is achieved by tilting the incident electron beam by an angle, whose magnitude determines the effective displacement wave vector $\mathbf{q}'$ of the spectrometer entrance aperture with respect to the forward scattered direction. Orientation dependence of core-loss edges from anisotropic materials can be determined by angle-resolved inelastic scattering of fast electrons.[68,69] Alternatively the effective displacement along $q'$ can be achieved by shifting the relative position of the spectrometer entrance aperture using post-specimen electronics.[70]

Momentum-dependent experiments have been successfully implemented in the entire range of the energy loss spectrum. High-loss q-EELS has been employed to probe the anisotropy of chemical bonds[68,69] and the detection of a magnetic dichroic signal.[71] In the low-loss energy range, the different q-space dispersion relations of the spectral peaks can be used to probe the dielectric characteristics,[72] plasmonic[73] and excitonic dispersions[74] and to identify optically forbidden transitions.[75] More recently, the introduction of improved monochromator designs with energy resolutions of a few ~meV has allowed for momentum-resolved vibrational EELS.[63,76]

While extremely powerful, STEM q-EELS is not without challenges, the first being the fundamental physical limit of momentum versus spatial resolution, with the momentum resolution, $\Delta q$, being inversely proportional to the electron probe size as determined by the beam convergence semi-angle $\alpha$ [77]. The large beam convergence angles used in STEM result in broad, overlapping diffraction discs, limiting the q-space resolution of the spectral data. Nevertheless, by careful choosing the experimental conditions, the momentum resolution[63,76] can be balanced with the achievable spatial resolution as recently demonstrated by atomically resolved STEM-EELS phonon spectroscopy experiments.[64,65,74]. The second challenge lies in the weak signals associated with vibrational EELS specifically and the rapidly decaying EELS signal off the optical axis,[78] which require lengthy experiments and the sacrifice of different types of resolution (such as spatial, momentum, or energy resolution) in favor of signal detection. The introduction of new rapid, high-efficiency direct electron detectors[79] promises to expand the range of q-EELs experiments in the future.

## [H2] 4D-STEM

New faster and more sensitive detectors are becoming available in electron microscopy. This is partly due to general advances in microelectronics devices and partly due to novel developments such as direct electron detectors[79]. Pixelated detectors can be used to acquire the electron intensity depending on the scattering angle. Different areas can then be selected during post-processing to effectively allow a variety of images to be constructed. In addition, pixelated detectors can acquire multidimensional sets of data (see Fig 2b and Box 2), which can potentially be used to reveal additional information about the structure and properties of materials at higher resolutions.

In electron microscopy, the reciprocity theorem of Helmholtz suggests that swapping the source and the detector should give the same image in the case of elastic scattering.[80]. Applying this principle to the STEM thus indicates that a small on-axis detector in STEM should give a BF image equivalent to a conventional TEM image. However, in order to obtain a highly coherent image, the collection angle must be very small and most of the electrons that have

interacted with and possibly damaged the sample will not be used, which is why BF-STEM imaging has traditionally been regarded as inefficient due to the low signal to noise ratio. Another way to look at coherence issue is to think of each part of the detector as forming an image at a different tilt angle.[80] Adding these slightly different images together, for instance by using a larger detector, will tend to blur some of the fine details and limit spatial coherence[81]. However, all the transmitted electrons can be collected as a function of angle at a particular probe position using a pixelated detector to obtain a nanodiffraction[82-85] pattern, also known as an electron Ronchigram.[86] In STEM, the electron Ronchigram provides several routes to measure aberrations, which is important for aberration correction[35,87-93] and for measuring spatial coherence, which is a resolution-limiting factor in aberration-corrected instruments and important for quantitative interpretation of images.[94,95] Nanodiffraction patterns can be collected at every probe position to generate a 4-dimensional (4D)-STEM dataset, also called a scanning nanodiffraction dataset. Such operating modes potentially use all the transmitted electrons to reconstruct an image and can be more efficient.[96-99] 4D-STEM has become a commonly adopted method with the commercial availability of high-quality detectors, with a recent comprehensive review published in [46].

The strength of 4D imaging is the extra information available within the collected data. First, as the electron beam transmits through the sample, it might be deflected by fields inside the sample. This deflection can be differentially detected[2,44,100] for differential phase contrast with multiple detectors, potentially at atomic resolution,[44] and can be used to reconstruct the measured fields.[101] Differential phase contrast contains details of the sample's electrical and magnetic fields, although how to untangle all of these subtle effects is not always clear, particularly in the presence of dynamical scattering. Other information that can be obtained includes polarization domains,[102] strain fields,[103-107] octahedral tilts,[108] and local symmetry.[99] An ongoing question is how to maximize the amount of information that can be extracted.

One way to maximize information extraction uses ptychography, where the redundancy in combining real-space and diffraction-space data can be used to solve for both the phase and amplitude of the electron beam[109] and permit more details of the sample to be reconstructed. Modern ptychography solutions are typically iterative methods derived from the Gerchberg-Saxton algorithm[110], with seminal implementation found in[5,111-113]. The main issue with these iterative methods is that they are still computationally intensive and have yet to incorporate the full effect of dynamical scattering, which limits their applicability to strongly scattering specimens with high-$Z$ elements and/or thick specimens that are > a few nm. While one solution is to simply apply more computer power to the problem, using artificial intelligence/machine learning (AI/ML) techniques could present a better route.[114] Similarly, a significant amount of 3D information encoded in the electron Ronchigram and 4D datasets,[115,116] could also be extracted using AI/ML techniques. There are several other applications for ML in the context of 4D imaging. Because only a small fraction of the beam electrons strongly interacts with the sample, almost all signals tend to be noisy or blurred by instrumental effects, meaning that, for example, even simple deconvolution or denoising procedures can be extremely useful.

Finally, there is a lot of scope for future development of 4D-STEM techniques. In principle it might be possible to record an energy loss at every scattering angle at every probe position combining q-EELS and 4D STEM information, or extend dimensionality further by combining 4D-STEM with tomography. One problem is that current detectors are 2D, which means that acquiring a higher dimensional dataset requires multiple scans of the same area. If the sample is damaged by the beam, changes over time, drifts away, or contamination builds up, repeating the same scan multiples times over the same area of the sample may not be possible. Thus, using ML methods to infer a model that informs an automated experiment to obtain a sub-set of this multidimensional data is one way to approach the multidimensional information challenge.

## [H2] Beam Engineering

While most efforts have concentrated on reducing probe size and increasing lateral and depth resolution, a new frontier of STEM is in beam engineering or shaping as a way to address emergent physical phenomena associated not usually observable using conventional imaging techniques.[117-120] Owing to new electron-optical elements, the amplitude and phase profile of the probe can be determined beyond the typical diffraction-, aberration- and coherence-limited shape. The functional beams created by beam shaping can widen the range of quantities that can be measured, determine the dynamical diffraction in the material,[121,122] and enhance the part of the electron scattering to be characterized.

Engineering of electron beam shape have evolved from the initial use of electron holographybased on a nanofabricated thin slab of material, to magnetic phase plates[123] to new applications based on electrostatic controllable phase modulators[124-126] that can be directly positioned in the condenser diaphragm. Below and in Fig 2 we give some examples of functional beams. A vortex beam, which is a doughnut-shaped beam with an azimuthal phase gradient, is shown in Fig 2c with a rotation of $2\pi$, although multiple integer rotations in the phase are also possible. The presence of the vortex in the phase results in zero intensity at the center of the probe. Vortex beams initially aimed to probe the spin state of atoms in magnetic materials by elastic[127,128] and inelastic scattering[129,130], and have also been used to form holograms to measure the vertical magnetic field.[127] More recently, nearly linear contrast images have been generated by matched illumination and detector interferometry ((MIDI)-STEM), which combines structured illumination based on modified zone plates phase with a high-speed direct electron detector[131] (see Figure 2d). The image of the beam recorded on the direct electron detector is post-processed with a mask matching the phase plate. Phase-shaped electron beams have also been used to probe the symmetry of localized surface plasmon resonances[132]. Appropriately tuned Gaussian-shaped beams[121] couple strongly with column and propagate almost unperturbed in the crystal. As shown in Figure 2e, this is predicted to produce contrast that does not oscillate as a function of specimen thickness and produce images that are simpler to quantitatively interpret. Bessel beams or equivalently hollow cone illumination can be used to increase the depth of focus of the beam[7,119], to facilitate strain analysis through nanodiffraction[133,134] or to improve the interpretability of low-loss EELS[135] by mimicking the idea of precession electron diffraction and combining beams of different directions to minimize the effects of channeling (see Fig 2f). This produces a more complex diffraction pattern that requires specialized post-processing and removes the need for expensive microscope modifications.

Finally, one key challenge in STEM is to control both the initial and the final state of the beam after scattering. For example, the HAADF-STEM detector has been used as a form of filter[136] but more sophisticated filters allow for the direct measurements of new final states without acquiring the full 4D-STEM, such as the measurement of orbital angular momentum through an orbital angular momentum sorter.[137]. It is clear that the increasing complexity of these new optical systems will require a more sophisticated control. For this reason, the techniques such as convolutional neural networks[138] as well as more advanced approaches such as Bayesian optimization[139] and reinforcement learning[140] will have an increasing role in the control of the complex optical system and in the alignment of the microscope in general.

## [H1] Results

Because modern electron cameras record data digitally, analysis of STEM experiments is typically carried out using software programs or computational methods. The initial analysis steps are to evaluate the quality of the experimental data and select one or more datasets for inclusion in a study. For STEM images, the minimum data processing required is to scale the image contrast and brightness and to crop around the features of interest. Quantitative measurements may also be performed, for example measuring the position of atomic columns, the length of atomic bonds, or comparing calibrated measurements signals to image simulations to estimate local structure or composition.

Higher-dimension STEM datasets often require substantially more analysis and data processing. Examples of 3D datasets include: time-series movies that may require drift correction to remove sample motion; tomographic tilt series requiring software to reconstruct 3D sample volumes; or spectroscopic datasets such as EDS or EELS that are usually interpreted from dictionaries of known signals or first principles calculations. 4D datasets such as qEELS or 4D-STEM also require specialized software codes for very large datasets.

## [H2] STEM Data Format and Analysis Programs

Modern STEM instruments are capable of producing large data streams from a variety of measurement channels to deliver important specimen information. This data is most valuable when analyzed in a quantitative, transparent, and reproducible manner. Most STEM datasets are currently recorded in proprietary formats defined by the hardware vendors. Some progress has been made towards saving data in open formats and defining standards for interoperability, but there is a long way to go before all STEM experiments follow findable, accessible, interoperable, and reusable (FAIR) data principles.[141]

Since STEM data is usually recorded on vendor software platforms, data analysis typically starts there. These platforms include the Gatan Microscopy Suite (GMS) with Digital Micrograph from Gatan, Velox from Thermo Fisher Scientific, Swift from NION, ESPRIT from Bruker, etc. These acquisition and analysis programs all offer the ability to visualize data as it is being recorded and are invaluable to provide feedback during experiments. They often include complex analysis methods such as quantification of EDS experiments from reference spectra. However, these vendor analysis platforms and methods are often closed source "black box" methods where the underlying code cannot be read or modified with, to date, the Nion Swift as the exception. Some of these platforms allow execution of external analysis scripts, which in the case of Digital Micrograph has led to a vast number of user-developed analysis methods. There is, however, no central listing of these methods and few have been vetted by the community. Nevertheless, many papers are published each year containing STEM results analyzed entirely on vendor software platforms.

To perform more complex analyses of STEM data, many users either write their own software or use community-developed codes. One of the most popular programs for analyzing imaging data is ImageJ, an open-source image processing program written in Java. Many scientists have written ImageJ plugins to read proprietary STEM data formats and perform various medium and atomic resolution corrections or measurements.[142,143] Another widely used programming language for analysis of STEM data is MATLAB, which despite not being open source does have a large library of user-created analysis tools either available on the central File Exchange or on researcher websites. There are also free languages that are mostly compatible with MATLAB such as Octave.

Recently, the Python programming language has started to overtake other choices for the analysis of STEM data, in no small part because Python has become the *de facto* standard for ML data analysis. For conventional analysis methods, there are various community developed Python codes devoted to various aspects of STEM analysis. The python package

most widely used in the electron microscopy community is HyperSpy, which was originally developed to analyze spectroscopic data such as STEM-EELS experiments. HyperSpy has grown into a general-purpose analysis toolkit for STEM data with many plugins targeting specific data types. These include atomic resolution image analysis with Atomap, luminescence spectroscopy data analysis with LumiSpy, 4D-STEM data analysis with pyxem,[144] and others. Another family of STEM analysis methods can be found in the Python-based Pycroscopy ecosystem,[144] which includes STEMTools toolkit (https://github.com/pycroscopy/stemtool), AtomAI library(https://github.com/pycroscopy/atomai) for applications of deep learning to microscopy data including deep kernel learning, and invariant representation learning, PyTEMLib library(https://github.com/pycroscopy/pyTEMlib) for model-based quantification analysis, and others. In addition, there are 4D-STEM-specific analysis codes currently under development including LiberTEM[145] and py4DSTEM.[146] All these packages are fully open source and can be freely modified to match the specific needs of a given research project.

Finally, simulation also plays a large role in STEM research. The most common simulations are imaging simulations incorporating the electron scattering process within the specimen together with the image formation system for the various imaging modalities, which can be performed using open source simulation codes including QSTEM,[147] $\mu$STEM,[148] Dr. Probe,[149] MULTEM,[150] STEMsalabim,[151] abTEM,[152] Prismatic,[153] and others. These codes often have specific focuses, for example abTEM includes the ability to use the electrostatic potential of a structure directly from density functional theory (DFT) calculations performed with the open source code GPAW,[154] while Prismatic focuses on fast calculations. Recently, it has become possible to simulate STEM spectroscopic signals such as plasmon scattering,[155,156] double-channeling core-loss STEM-EELS scattering,[157] very large crystals containing crystallographic defects,[158] and other modalities. Many new computational methods are under active development and large STEM simulations can be tackled with the widespread availability of GPU resources.

## [H2] Analysis of hyperspectral data

The discussion of ML in STEM necessitates a brief overview of exploratory data analysis that originally emerged in the context of EELS hyperspectral image analysis. Core-loss EELS provides a wealth of information such as chemical composition and local bonding but correct interpretation requires careful analysis and processing. For example, truly quantitative analysis of electron energy loss near edge structure (ELNES) requires the removal of plural scattering using methods such as Fourier-ratio deconvolution, where the simultaneously acquired zero-loss-peak region is used to deconvolve the plural scattering from the core-loss signal.[159] This process makes the edge structure more interpretable and provides increased signal to noise ratios (SNR). Quantitative interpretation of features such as the $L_{23}$ ratio also requires background subtraction and removal of the continuum components.[160,161] Background subtraction is often done using a simple power law fit to the spectrum prior to the edge of interest, but alternative methods are also available.[162] Following this preprocessing, quantification can be carried out by estimating peak heights and separation by either Gaussian fitting[163] or more sophisticated model fitting approaches.[164]

### [H3] Principal component analysis

Because the crosssections for ionization are quite small, especially for higher energy edges, the SNR of ELNES is generally quite low. This is also the case for other spectroscopies such as energy dispersive X-ray spectroscopy (EDX). To improve the SNR, multivariate statistical analysis of EELS was first demonstrated in the late eighties.[20,21] While many multivariate methods exist,[19] the most common method for denoising EELS data is principal component

analysis (PCA). PCA has been widely applied to both chemical mapping and near edge structure analysis[165-167] and has been used as a first step in more advanced analysis methods such as vertex component analysis[168] or the study of precipitates in manganese steels.[169,170] An alternative linear method is non-negative matrix factorization (NMF), which has been applied to core-loss EELS and EDX[171] and plasmonic data.[172] PCA and NMF are available in freely-available software packages and straightforward implementations are available in popular languages such as Python, making them an accessible approach to denoising EELS data.

PCA represents a dataset as a linear sum of weighted orthogonal components with the weighting determined by the commonality of each component. Features that are common throughout a dataset will receive a higher weighting while features such as noise, which is essentially random, get a much smaller weighting. Removing the lower weighted components before reconstruction results in the removal of most of the random noise. However, because features such as interfaces and defects cover only a small volume of the dataset, they receive low weightings as well. This means components containing valuable information regarding local structures can be removed along with those related to noise. While PCA has therefore been successfully applied to perfect crystals, the method can introduce unexpected artifacts in the processed data when the raw data contains interfaces and defects [57,162,173]. This can result in changes to the features to be measured, such as shifting apparent peak positions in near edge structures and changes in intensity. Since much of the most interesting physics happens at interfaces, defects and other localized structural features, this is highly problematic and approaches to successfully examine small changes in near edge structure due to local structural changes are required.

## [H2] Quantitative STEM

While conventional STEM analysis methods study contrast or image features, quantitative STEM makes use of the absolute scattered intensity in each pixel of a STEM image to enable more information about the specimen to be extracted. Quantitative STEM is essential to quantifying noise and information content and for quantitative comparison between STEM images and quantum mechanical image simulations.[174] The additional information that quantitative STEM can deliver about a specimen has been used in applications such as atom counting,[175] local composition measurements[63] and nanoparticle shape[176] and surface stability measurements.[177] Unless the sample allows for internal calibration of the image intensity,[178] calibration of the STEM detector and the microscope optics are essential.[179,180]

### [H3] Intensity measurement

Modern STEM detectors respond linearly to the incident electron current, with the dark count rate $D$ and the gain $G$ both adjustable by the user. The intensity in electrons in a pixel $I$ is connected to the measured digital counts $C$ by $I = (C - D)/G$. $G$ is measured by placing a signal of known intensity on the detector and there are several ways to measure it. For example, one convenient signal is the intensity of a single electron[181,182] $C_1$. $C_1$ can be measured by acquiring an image while allowing almost all of the beam to pass through the hole in the detector to yield $G = (C_1 - D)$. Another convenient signal is the full current of the beam $B$, which can be measured separately by using a Faraday cup for example, to yield $G = (C - D)/B$. Comparing STEM images to simulations requires normalizing the image intensity to the incident beam,[174] and is straightforward if $D$ and $G$ are known and $B$ is known in amperes. if $B$ is known only in digital counts as $B_C$, a quantitative image $I_n$, expressed as a fraction of in the incident beam current, can also be obtained. In this case, $I_n = (C - D) / (B_C - D)$.[174]

### [H3] Calibration

Comparison to simulations requires careful calibration of a variety of other microscope factors that must then be included in the simulations. These factors include detector inner angles[183] and outer angles, cutoffs or shadowing from complex pre-specimen optics such as aberration correctors,[179] and the non-uniform response of common scintillator detectors.[179,183] Finally, the simulated image intensity must be convolved using a calibrated incoherence function[174]. For aberration-limited systems, the incoherence function is typically a Gaussian with FWHM slightly smaller than the STEM image resolution and can be calibrated by comparing experiments to simulations for a known crystal sample, with thickness independently measured by position-averaged convergent beam diffraction.[184].
The incoherence function also accounts for fast instrument instabilities on the pixel acquisition time scale, such as high-frequency mechanical vibrations or probe jitter. These effects can be measured separately and may not be Gaussian.[94,185] When made in an appropriate optical plane, these measurements incorporate the effect of all mechanical and electronic instabilities in the STEM system, as well as the partial spatial coherence of the electron emitter.[186]
For aberration-corrected systems, the mathematical form of the incoherence function at high spatial frequencies can become important for quantitative analysis and needs to be measured in the absence of any other adjustable parameters.[95,180] In systems corrected to higher angles, i.e. with larger probe forming apertures, temporal incoherence becomes significant and may also need to be measured and incorporated in quantitative analysis.[187-189]

## [H2] From Images to Atomic Positions

The materials information contained in the data of an atomic-resolution STEM image can often be captured in a list of the positions, intensities, and shapes of the atomic columns[190,191] and uncertainties in those quantities. The combination of positions and intensities encodes crystallographic phase and orientation, the positions and crystallography of interfaces and defects, and strain fields. Intensities can be used to count atoms in each column along the beam direction[175,192] and the intensities and column shapes can be used to determine the presence and depth of impurities in a column.[6,193,194]

### [H3] Classical approach

The classical method for determining atomic column parameters from atomically resolved STEM images is to use a particle-finding approach, potentially informed by the translation symmetry of the underlying lattice,[195] to find approximate positions for all the columns in the image, then refine the parameters for each column by the least-square fitting of its intensity $I(x, y)$ to an assumed functional form for the microscope point spread function.[192,196,197] These approaches assume that the atomic column position corresponds to a specific form of intensity distribution, typically a local intensity maxima. This assumption has greatest validity in the case of *Z*-contrast images but care must be taken to understand the influence of dynamical scattering, which can shift intensity maxima in an image relative to the actual position of the atomic column.[198] An example of a *Z*-contrast image of GaN $[11\bar{2}0]$ [110] is shown in Figure 3. Figure 3(a) shows the image with red dots marking the initial, approximate positions of the Ga columns (the light N columns are not detected), determined by fitting to a two-dimensional Gaussian:

$$I(x,y) = I_0 + A\exp\left[\left(\frac{-1}{2(1-c^2)}\right)\left(\left(\frac{x-x_0}{x_w}\right)^2 + \left(\frac{y-y_0}{y_w}\right)^2 - \left(\frac{2c(x-x_0)(y-y_0)}{x_w y_w}\right)\right)\right]. \quad (1)$$

Where $I_0$ is the local background intensity, $A$ is peak intensity of the column, $(x_0, y_0)$ is the column position, $x_w$ and $y_w$ are the widths, and $c$ controls the rotation of the major and minor axes of the fitted Gaussians, accounting for the slight variations in the atomic shape. Figure

3(b) shows the residual between the data and the fit, which is both small and random, indicating a good fit in this case. Obtaining meaningful uncertainties in the fitting parameters requires weighting the data by its uncertainty. The uncertainty in $N$ detected electrons in a pixel is never less than $\sqrt{N}$ and may be greater due to detector noise for example.

Equation (1) describes local fitting around each atomic column, so the procedure must be repeated for every column in the image. If several atomic columns are close enough together that their images overlap, simultaneous fitting to the sum of several 2D Gaussians may be required for an "all at once" fitting of every column in the image,[197] which necessitates a more complicated function for $I_0$ with additional fitting parameters. For a high SNR and low-distortion data, which can be achieved by distortion correction and averaging, this fitting procedure can result in sub-pm precision in locating atomic columns even if the columns are ~100 pm wide.[192] High-quality images also lead to high success rates for the initial approximate column finding and reliable convergence for fitting. For SNRs and distortions more typical of single STEM images, the achievable precision is closer to 15 pm[199] and initial column finding and fitting convergence may require more manual operator tuning. Smoothing or image denoising using methods such as Fourier filtering, total variational denoising, or non-local algorithms such as block matching and 3D filtering[200] can improve atom finding, but as these methods do not typically preserve image intensities, the fitting step should still be performed on the original intensities. Classical fitting methods are therefore less suitable for single-shot images and prohibitively time-consuming to apply to very large images or to a time series of images.

**[H3] Computer vision approach**

Computer vision methods such as convolutional neural networks (CNNs) have recently been used to identify atom column positions in STEM images.[201-203] CNNs mimic human vision by identifying image features, which are patterns in intensity, rather than by numerically comparing intensities to a model, for example by using least squares. CNNs must be trained on pre-analyzed example images or on simulated data similar to the images that will be analyzed. Training an entirely new deep CNN from scratch is a major undertaking and requires an enormous volume of labeled examples. Fortunately, STEM simulations can be used to generate computer-labeled example images without the human effort of locating the atom positions and data augmentation can further increase the scope of the training data set.[201,202,204]

Figure 3(c) and 3(d) show an example STEM image of graphene and the deep convoluted neural network (DCNN)-derived atomic positions, respectively. Despite the low quality of the single-shot image, the DCNN correctly identifies all the atom positions. In general, DCNN atom identification is robust against both noise and distortion if those were part of the training set, making it a powerful approach for single-shot images. The precision in atomic column positions can exceed traditional methods.[203,205] In addition, most of the computational cost in using a DCNN lies in the training. Once trained, execution of the DCNN is very fast, making analysis of large images[206] and time series[207] straightforward. Current atom finding DCNNs do not generate the entire set of atom column parameters in Eq. (1) but the DCNN-derived positions can serve as initial positions for fitting. It is not yet clear how to quantify the domain of applicability for a particular DCNN given its training data, and therefore successful application to images of, for example, unusual crystal structures or heavily distorted atomic columns may require some retraining. Applications of DCNNs can go beyond simply identifying atom positions to identifying clusters of atom positions characteristic of defects as shown in Figgure 4(e) and 4(f)[206].

**[H3] Autoencoders**

One way of denoising data while retaining small changes in the dataset is to use exploratory data analyses based on autoencoders. An autoencoder consists of two neural networks and learns a low-dimensional embedding of the data, called latent representation, in an unsupervised manner[208]. In an autoencoder, the first encoder neural network compresses the data into a small number of latent variables and the second decoder neural network tries to reconstruct the original data from the latent code. In the process, the autoencoder learns the optimal representation of high-dimensional data, such as images, while rejecting noise, which makes it a great tool for both non-linear dimensionality reduction and image or spectrum denoising. Latent spaces can often highlight regions where differences in spectra occur and lead to additional scientific insights. The autoencoder concept can be extended towards learning correlative relationships between structure in an image and property in spectral data as has been demonstrated with the *im2spec* encoder-decoder models.[209]
Finally, transformation-invariant variational autoencoders (VAE) build upon classical autoencoders by making the reconstruction process probabilistic and incorporating prior knowledge into the latent space structure[210]. Figure 4 shows the application of rotationally invariant VAE to the analysis of graphene data.

Typically, a VAE is a directed latent-variable probabilistic graphical model that learns a stochastic mapping between observations $\boldsymbol{x}$ with a complicated empirical distribution and latent variables $z$ whose distribution can be relatively simple.[211] A VAE consists of a generative model as a decoder that reconstructs $\boldsymbol{x}_i$ from a latent code $z_i$, and an inference model as an encoder whose role is to approximate a posterior of the generative model via amortized variational inference[212]. Implementation-wise, both encoder and decoder models are approximated by deep neural networks whose parameters are jointly learned by maximizing the evidence lower boundary via a stochastic gradient descent with randomly drawn mini-batches of data. VAEs can therefore build relationships between high-dimensional data sets and a small number of latent variables in a way reminiscent of manifold learning.

One important aspect of the VAEs similar to many manifold learning methods is that the variability of the behaviors in the latent space allows one to reveal relevant features of the system behavior, equivalent to primary non-linear degrees of freedom. Another important aspect of VAEs here is their parsimony–the training process generates the best short descriptors representing the data. The primary limitation of the classical VAE approach is VAE sensitivity to the details of experimental parameters; for example the presence of autoencoder networks to parameterize atomic STEM data imposing a known structural relationship between latent variables as an input for statistical analysis, including Bayesian inference and Gaussian process regression.

## [H2] Distortion Corrections

High-resolution STEM experiments are distinguished from conventional plane-wave TEM experiments by the small dimensions of the converged electron probe, which is able to reach sizes below atomic bond lengths with modern aberration-correction technology. The small probe dimensions requires scanning the electron beam over the sample surface in order to collect spatially resolved information in the detector plane, which leads to relatively long experimental acquisition times on the scale of seconds per image for most experiments. Any relative motion of the probe with respect to the sample during the acquisition time will introduce artifacts that can be due to thermal sample motion, mechanical vibration, local fields induced by charging or electronic instabilities of the electron beam.[213] In order to perform

precise measurements of atomic positions at the maximum instrument resolution, these residual drift artifacts must be measured and removed from STEM experiments.[214]

The most straightforward method for reducing errors due to sample motion is to simply record an image series, align all the images, and take the mean value of each pixel.[215,216] However, this method does not efficiently make use of all available information since it does not attempt to recover any information lost by distortions in the imaging system. A more advanced correction method uses the STEM's ability to scan in any direction and rotate the orientation of the scanning direction relative to the sample[217] to measure linear sample drift due to thermal motion and correct acquired images by applying an affine transformation.[218] This approach can be extended to measure and correct local nonlinear or nonrigid distortions in acquired images by using gradient descent.[219-221] These correction approaches have also been applied to STEM-EELS and 4D-STEM data.[222,223]

Another family of methods for distortion correction makes use of information measured from the sample itself. For example, if the crystalline lattice parameters of a material are known, the linear drift can be removed from a single image.[224] In atomic-resolution images, it is also possible to directly track the measured position of atoms in a time series to estimate the undistorted configuration.[225-228]

Finally, STEM is not restricted to recording square or rectangular scan patterns. A variety of complex scan patterns have been proposed to produce a more uniform acceleration of the electron beam. Some examples include spiral scans[229,230], randomized beam shifts[231,232], blanking[233] and fractal Hilbert space-filling curves.[234] More complex patterns can help decouple the direction of motion of the sample from the movement of the beam, allowing for more isotropic information transfer in all directions. We note however that all of the above-mentioned methods only correct relative errors between images. Length measurements on an absolute scale still require precise calibration.[235]

## [H1] Applications

Advances in STEM over the last decade have allowed routine visualization of atomic structure of solids and localization of atomic columns with sub-pm precision. Descriptors that are strongly correlated with the functional properties of materials, such as atomic bond lengths and angles, can now be measured at the single-atom level. For example, the bond length in carbon compounds is directly linked to reactivity and bond strength. Similarly, bond length and bond angle in perovskites are strongly correlated with metal-insulator and ferromagnetic-antiferromagnetic transitions.[236-238] Traditionally, these descriptors have only been accessible on a macroscopic level via X-ray and neutron scattering; advances in STEM naturally lead to questions as to how it can be used to explore the local physics and chemistry of crystalline and disordered solids. For example, one advantage of local imaging is the measurement of bond lengths and bond angle variations within a material, for example at surfaces, interfaces and defects, to map of strain fields via direct measurement of atomic column positions. Correspondingly, multiple examples of strain mapping in the vicinity of second phase inclusions, dislocations, and surfaces have been reported[239-241], with the experimental structure compared with the classical solid-state mechanics models[242] in many cases. In this section, we show examples of the application of HAADF STEM imaging to several topical materials science problems. While this work focusses on HAADF imaging, the STEM has many other imaging modes that can be applied to materials science,[243] but applications to biology are also common.[244]

## [H2] Mapping ferroelectric phenomena

The high spatial resolution of STEM makes it a perfect tool for exploring the physics of materials with strong coupling between the order parameter and structural distortions.[245,246] Seminal works using TEM[247,248,249] and STEM[250] demonstrated that quantitative measurements of atomic column positions can be used to map the polarization order parameter field. This approach was rapidly extended to other physical functionalities strongly coupled to structure, including octahedra tilting in perovskites in both the image plane[251-253] and the beam direction[254] and chemical and physical strain fields[255,256]. Common to this approach is an *a priori* postulated relationship between the observed contrast (for example atomic column positions) and the physical descriptor (for example polarization).

The observation of the order parameter field and its evolution near surfaces and interfaces opens a pathway to learn the mesoscopic physics of the systems, such that mesoscopic models with a small number of free parameters can be matched to STEM observations. For example, the correlation and interfacial terms in the Ginzburg-Landau free energies can be extracted from order parameter profiles across domain walls and interfaces[253,256,257], while analyses of ferroelectric vortex shapes can be used to derive the numerical values of the flexoelectric tensor.[258] Recently, these approaches were extended into the Bayesian domain to take into consideration any prior knowledge about the system and evaluate changes in our understanding of material behavior given new experimental data.[259] Comparing experimental data to models can systematically address issues such as the resolution and information limits required to observe specific physical phenomena and whether prior knowledge of the system enable additional insights from the experimental data.

Although determining the atomic positions in an image to assess polarization in a ferroelectric domain is a valid approach,[260] it is an indirect way of revealing a material's ferroelectric property. A more direct approach that maps polarization in ferroelectric materials is based on differential phase contrast STEM, where an annular detector split into at least four independent azimuthal segments is used to derive changes to the center of mass in the diffraction pattern due to the in-plane component of the electrostatic field in the sample.[44] A center of mass change is determined by calculating the intensities difference of two opposite detector segments and is proportional to the electrostatic field in the specimen, which affects the angular propagation of the beam while interacting with the ferroelectric specimen. This approach has recently successfully mapped large polarization gradients across naturally formed domains in a doped ferroelectric material.[261] Although clear features due to ferroelectric polarization can be observed at the atomic scale, the differential phase contrast image intensity requires careful interpretation as the nanoscale field component due to the ferroelectric field is superimposed on the electrostatic field of the atomic potentials.[262] As mentioned above, any in-plane electrostatic field in the sample leads to a change of the center of mass in the diffraction pattern formed behind the sample. The same is also true for magnetic fields, which similarly affect the propagation of the electron beam. By carefully disentangling the electrostatic contribution from its magnetic counterpart, differential phase contrast STEM can also be used to measure magnetic properties of materials on the nanoscale,[3] similar to off-axis electron holography[263] carried out in broad-beam TEM mode.

## [H2] Grain boundaries and interfaces

STEM imaging is a valuable probe of the atomic structure at grain boundaries and interfaces in crystalline materials. However, despite improvements in spatial resolution and depth estimation[3], it remains difficult to determine the 3D atomic structure of interfaces. Computational methods such as DFT can be used to supplement STEM results to improve structure determination,[264-266] but extending such methods to simulating long-range behavior

across interfaces of practical size is not straightforward. While ML and materials informatics have played a significant role in the development of crystal structure and material property descriptors[267] to extend length scales and serve as a proxy for expensive calculations,[268,269] reliable descriptors and data-driven modeling paradigms are lacking for interfacial systems.

To transform existing trial-and-error approaches,[270] high-throughput computation is used together with energetic evaluation from atomistic modeling, STEM image simulations, and computer vision-based image comparison in order to determine 3D atomic structures at grain boundaries and interfaces.[271] For energetic modeling of practical interfaces, interatomic potentials provide a computationally efficient alternative to electronic structure methods such as DFT while often capturing the same essential physics. Interatomic potential structure search schemes such as basin hopping and genetic algorithms sample a space of lowest energy solutions and are often sufficient if the task is to find the most stable structures. This is problematic if the system of interest contains possible metastable configurations, as is the case with grain boundaries and interfaces, because these interatomic potentials will often work to minimize energy regardless of whether the resultant structure is consistent with observations. With suitable constraints at the boundaries, simulated annealing using empirical potentials has been used to determine structural motifs during grain boundary migration in aluminum oxide (Fig. 5a-c).[272] However, there is no guarantee that the computational structures obtained in the simulated annealing simulations match the experimental structures and the process of matching can be labor intensive. Therefore, it is desirable to have an additional constraint that involves a measure of similarity between simulated and experimental STEM images to ensure that the space of structures explored is somewhat consistent with observations.

For image comparisons, similarity measurements are functions or processes that quantify how alike two images are. Some similarity measurements are direct pixel measurements, where pixel values are explicitly considered when making a comparison. In the simplest case, a direct-pixel measurement is a sum of the errors between pixel values or the mean squared error. A more sophisticated method such as the structural similarity index measure (SSIM)[273] processes collections of pixels as patches and provides similarity measurements based on a multiplicative combination of intensity, contrast, and structure terms. By normalizing image patches for intensity and contrast differences, SSIM reveals the structure of the image signal. Finally, state-of-the art image comparison for image and video processing applications are often based on a visual information fidelity in the pixel domain (VIFP).[274,275] The principle behind VIFP is that image quality (in our case similarity with respect to a reference) corresponds to a measure of Shannon information fidelity loss between the reference and distorted image relative to the information of the reference, using a combination of sub-band-coding, distortion models, and models of the human visual system.

Figure 5d-g shows a HAADF-STEM image of a cadmium telluride grain boundary,[276] from which the corresponding 3D structures are obtained using high-throughput computation with an iterative basin hopping scheme. The optimization objective combines the interfacial energy obtained from the Stillinger-Weber interatomic potential with image similarity between the simulated and experimental STEM image using SSIM image matching. Rather than a single 3D structure, a family of structures is obtained that can be further investigated using first principles computation and statistical analysis. It can be seen that the 3D structures obtained using this method are not composed of simple dislocation core models and cannot otherwise be conjectured.

## [H2] Chemistry

STEM is widely used to investigate interactions between atoms and atomically thin supports owing to its sub-Ångström resolution. In particular, HAADF-STEM can provide a

higher image intensity for heavy elements compared to lighter elements based on *Z*-contrast imaging, where the signal intensity is proportional to the *v*-th power of the atomic number of an element for $I \sim Z^v$. The exponent, *v*, ranges typically from 1.6 to 1.9 depending on the radius of the atoms, their scattering factor, the surrounding atomic structure and the specimen thickness.[28,277-279]

### [H3] Heavy atoms on 2D substrates

HAADF-STEM enables the tracking of heavy metal atoms on monolayer 2D substrates, shedding light on their diffusion, aggregation, and the physics of various atomic interactions. For example, the atomic interactions between single platinum atoms and monolayer molybdenum disulfide, a 2D substrate, have been extensively investigated, with the migration of single platinum atoms on the surface of molybdenum disulfide well-tracked by sequential imaging[280]. The atomic model of the hopping of platinum atoms between sulfur vacancy sites is shown in Fig. 6a. All platinum atoms are on sulfur instead of molybdenum sites, confirming the affinity of platinum atoms to sulfur vacancies on the clean molybdenum disulfide surface. The trapping of platinum nanoclusters at dislocations in molybdenum disulfide has also been studied.[281] Fig 6b shows that upon annealing of a chloroplatinic acid precursor at 350 ℃ on an *in situ* heating holder, seeds and clusters of platinum appear at the grain boundaries of polycrystalline molybdenum sulfide. The platinum nanocrystals at the grain boundaries reveal a tilt angle of 60° between the two grains, which is confirmed by fast Fourier transform, and highlight the interactions between platinum dopants and grain boundaries of 2D materials. The presence of hydrocarbon contamination on 2D substrates, commonly introduced during the growth of 2D materials, has also been shown to play a significant role in the trapping and stabilization of metal atoms and nanoclusters. For example, gold, iron and chromium atoms primarily reside on the amorphous carbon contamination of graphene substrates[282].

### [H3] *In situ* heating

*In situ* heating in aberration-corrected STEM is a powerful tool for investigating the epitaxial growth of nanocrystals on a 2D substrates. Upon heating of a precursor to 800 ℃, platinum atoms diffuse on the surface of molybdenum disulfide and grow into nanocrystals (Fig. 6c).[283] The (111) plane of the platinum nanocrystal aligns along the zigzag direction of molybdenum disulfide (Fig. 5d) while the (002) plane aligns with the armchair direction of the substrate. The *d* spacing of the platinum nanocrystals ($d$ = 2.28 Å), obtained from a Fourier transform analysis of the image of the platinum nanocrystal and a reference material, is comparable to that of a bulk platinum crystal ($d$ = 2.265 Å), indicating little strain formed in the crystals (Fig. 6e). In addition, strategies have been developed to control the morphologies of metal nanoclusters forming on 2D materials. One example is a metal precursor incorporating a bulky organic ligand around a metal to suppress the formation of 3D atom clusters on the 2D support at elevated temperature and yield 2D clusters with flat surfaces and single metal atoms after annealing.[284] Various reports where the epitaxial interactions between the metal nanoclusters and the substrates are analyzed by STEM imaging include the growth of 2D palladium diselenide nanocrystals on monolayer molybdenum disulfide[285], monolayer lead iodide nanodisks on graphene[286], a 2D molybdenum diselenide film on hexagonal boron nitride[287] and pyrochlore $Nd_2Ir_2O_7$ grown on yttria-stabilized zirconia[288].

The van der Waals interactions between a flat aromatic molecule and the basal plane of an underlying substrate has also been elucidated by STEM. In [289], a planar aromatic molecule containing platinum metal as a marker for tracking was designed. During deposition, the molecules were found on the basal plane of the molybdenum disulfide near the edges of nanopores owing to the strong van der Waals interactions between the aromatic core of the molecule and the underlying molybdenum disulfide substrate. Molecule configurations were hypothesized by measuring the distance between the platinum markers and the edges of

molybdenum disulfide pores in the STEM images. This heavy-metal tagging method presents new opportunities for studying the interactions between tagged molecules and the underlying 2D substrates.

### [H3] *In situ* reactions

Recently, STEM has contributed to elucidating chemical reaction mechanisms during catalysis by directly imaging and tracking metal catalysts. For example, in the gold core of a core-shell nanoparticle with a nickel phosphide shell, gold atoms have been shown to diffuse into the nickel phosphide shell during annealing and cause an inward expansion of the structure.[290] The direct visualization of transition metal catalysts has provided chemical insight into the performance and mechanism of the hydrogen evolution reaction,[291] hydrogenation,[292] dehydrogenation,[293] and carbon monoxide oxidation[294,295]. STEM has also been used to monitor the structural change of stimuli-responsive organic molecules.[296] When platinum atom markers were incorporated in azobenzene photoswitches, the photo-irradiation of a flat *trans* isomer of azobenzene caused it to switch to a bent *cis* structure and reduced the platinum-platinum distance markers from ~2.1 nm to ~1.4 nm (Fig. 6f). Finally, the self-assembly of mesoscale systems has been investigated by STEM using heavy metal marked macromolecules. Platinum-marked porphyrin hexamers were imaged on graphene at a low electron dose to minimize beam-induced damage[297] and shown to align along the graphene zigzag direction owing to strong epitaxial interaction with the substrate even in the presence of hydrocarbon contamination. In addition, the self-assembly and packing of non-planar, bowlike macromolecules has been studied on graphene.[298] The W-marked calix[4]azoarene molecules stacked in a head-to-tail fashion, forming molecular chains and lamellar structures at high concentrations. In low concentration areas, various types of intermolecular interactions were detected, such as side-by-side and head-to-head interactions.

When studying processes such as chemical reactions or the transformation of small catalytic clusters whose properties strongly depend on the instantaneous atomic configuration, temporal resolution of the imaging process becomes critical. As mentioned above, common STEM acquisition speeds are in the range of seconds per frame. With a temporal resolution in the range of seconds, important dynamics might be missed. However, advanced scan engines allow for faster acquisition speeds. Recently, recording speeds in STEM mode exceeding 100 frames/s have been used to unravel structural transitions and the stability of small catalytic clusters.[299] One major drawback of such high frame rates in STEM mode is increased noise in the data because the electron dose is limited and needs to be balanced against the stability of the sample under investigation. With dwell times in the range of 100 ns or less, only a few tens of electrons transmit through the sample per scan position and fewer are eventually scattered to the detector to form the noisy signal, which suffers from shot noise. Noisy image series stemming from time-resolved data or from beam-sensitive samples, which do not tolerate a high enough electron dose for achieving a suitable signal to noise ratio, can be processed with classical, powerful denoising algorithms[300] that are typically slow and require long processing times. However, denoising of data is an area where ML can bring big benefits. Recently implemented denoising algorithms based on DCNNs[301] can process (both denoise and restore) time-resolved image series in near real-time. A well-trained algorithm therefore enables reliable restoration of atomic resolution data and paves the way for further increases of the temporal resolution in STEM and further minimization of the electron dose for radiation-sensitive materials.

## [H2] Structure of solids from atoms up

In modern condensed matter physics and materials science, current understanding of condensed matter and quantum systems hinges on macroscopic symmetry. Formalized via point and space group theory,[302] symmetry underpins areas such as structural analysis and serves as the basis for the descriptive formalism of quasiparticles and elementary excitations, phase transitions, and mesoscopic order-parameter based descriptions. The natural counterpart of symmetry-based descriptors is the concept of physical building blocks. Crystalline solids or magnetic or ferroelectric behavior can generally be described via a combination of unit cells with discrete translational symmetry of the lattice. Other systems such as Penrose structures possess well-defined building blocks but do not possess long-range translational symmetry. Finally, a broad range of materials fully lack translational symmetry, with examples ranging from structural glasses to ferroelectric and magnetic morphotropic systems.[303-311]. Typically, symmetry-based descriptors have led to much deeper insights into the structure and functionalities of materials with translational symmetries compared to partially and fully disordered systems.[312-314]

To date, the analysis of atomically-resolved imaging data has almost invariably been based on mathematics developed for macroscopic scattering data[245,313] despite the fundamentally different nature of microscopic measurements. As a simple example, consider an ideal crystal containing a macroscopic number of structural units. The symmetry of the diffraction pattern represents the symmetry of the lattice and the width of the peaks in the Fourier space is determined by intrinsic factors such as angular resolution of the measurement system rather than disorder in the material. The presence of symmetry breaking distortions, such as a cubic to tetragonal state transition, is instantly detectable via diffraction peak splitting. In comparison, only a small part of the object is visible in microscopic observations. The positions of the atoms are known only within an uncertainty interval and this uncertainty can be comparable to the magnitude of the symmetry-breaking feature of interest, such as tetragonality or polarization. Hence, two question arise: at which image size is it justified to define symmetry from atomically-resolved data and at which level of confidence can symmetry be defined? Ideally, such an approach should be applicable to structural data and more complex multidimensional data sets such as EELS[315,316] and ptychographic imaging.[5,103,317]

The alternative to conventional top-down descriptors is a bottom-up structural analysis, where unsupervised or semi-supervised ML methods are used to determine common structural blocks and the patterns they form in a solid. For materials with an ideal crystalline lattice, the conventional way to describe structures is in reference to the lattice. Fig. 7 shows a bottom-up structural analysis for a ferroelectric material[318]. The experimental image (Fig 7a) can be converted to a stack of sub-images describing either raw STEM contrast or a DCNN-based segmented image and subsequent application of linear or non-linear dimensionality methods yields the components and loading maps (Fig 7d). The component describes the salient elements of materials structure, whereas the loading map describes the structure of the solid.

The analysis is considerably more complicated in cases where the system does not possess discrete translational symmetry. When materials have atomic bond disorder, linear dimensionality reduction methods do not perform well owing to the large number of orientational variants necessitating alternative descriptors. If the atomic positions are established, analyses can be based on the analysis of the nearest neighbors.[319-321] Alternatively, the analysis can be performed using other rotationally invariant representations such as graph networks.[322]

Finally, bottom-up analysis can be performed via transformation-invariant variational autoencoders (VAE).[210] Shown in Figure 3 is the application of rotationally invariant (r)VAE to the analysis of graphene data. Here, the graphene lattice undergoes structural transformations due to electron beam irradiation, which results in the formation of topological defects. The rVAE is able to segment chemical space by separating the graphene lattice from topological

defects and grouping some of the topological defects together in the latent space. This was achieved by explicitly separating the variation in orientations of individual building blocks from the variation in structural content. The rVAE approach was also shown to enable effective exploration of the chemical evolution of the system based on local structural changes[323] and may be extended to more complex systems.[324] Importantly, the rVAE allows discovery of the molecular building blocks and chemical reactions pathways in an unsupervised manner.

## [H2] Physics of atomic interactions

Solids can be described based on generative models that give rise to the equivalent stochastic atomic, dipole, or spin microstructure.[325] Generative models for systems with defined atomic lattices and chemical site disorder are represented by lattice models such as Ising,[326] Kitaev,[327] Heisenberg,[326] etc. For example, binary solid solutions can be fully described via the corresponding Ising-like Hamiltonian that gives rise to statistically similar microstructures, and this description is compact and generalizable to non-observed concentrations and temperatures. S/TEM data yields the microstates of physical systems and enables comparison between generative models and experimental observations. Direct observation of the mesoscopic degrees of freedom can be directly compared to the lattice model via statistical distance minimization.[325,328,329] Beyond statistical analyses, observations of the multiple metastable configurations have been used to reconstruct the force fields acting between atoms.[330,331] Learning the generative model from atomically-resolved data, incorporating prior knowledge, and yielding corresponding uncertainties as posterior parameter distributions is therefore a clear opportunity for characterizing the intrinsic properties of material systems.

# [H1] Reproducibility and data deposition

In this section, we summarize the requirements for data storage in STEM. In any experiment, it is vitally important to include the important metadata allowing the experiment to be reproduced. In a STEM study, the most important experimental parameters are the accelerating voltage of the microscope, the convergence angle of the STEM probe (also called the numerical aperture), the step size between adjacent probe positions, the dwell time, and at least a rough estimate of the beam current in the initial STEM probe. Specific imaging modalities require additional metadata. 2D images recorded using monolithic detectors require a precise description of the detector response and detector coordinates, for example the angular range for annular detectors or the position and rotation of detector quadrants in differential phase contrast measurements, as well as the position and orientation of the diffraction pattern relative to these coordinates. 3D datasets such as time series require the time stamp of each frame, while tomography tilt series must specify the stage tilt angles for each time and spectroscopic measurements such as EDS or EELS must specify the energy range and/or bin width. Spectroscopic measurements should specify any other parameters needed to reproduce the experiment, for example monochromation conditions of the beam if applied and the collection angles of the spectrometer used.

In pixelated measurements of the probe such as in 4D-STEM, one must specify the calibrated output pixel size or provide absolute reference data and the rotation and/or direction of the probe scan steps relative to the detector coordinates. In 4D-STEM, it is also useful to record an image of the STEM probe passing only through vacuum, which can be used a probe reference image for ptychography, orientation, or strain mapping.

Other microscope parameters such as the voltage settings of the source, focusing optics, corrector optics, projector system, the vacuum levels of the microscope, hardware models, and software version numbers can also be recorded as a matter of course. Detailed notes regarding the sample including the preparation, geometry, and beam exposure conditions should also be recorded. In STEM studies it is particularly easy to record data at vastly different magnifications, making the task of generating survey images of the sample layout straightforward. Of particular importance in most STEM experiments is the sample tilt used for each measurement; one of the strengths of STEM imaging is combining observations at different orientations to construct a more complete picture of the specimen.

While new technologies have led to dramatic improvements in data quality, they have introduced a parallel development which we think is even more important in the long run; the digitization of TEM and STEM studies. Because STEM imaging experiments typically use large monolithic detectors that directly output electronic signals, they provided many early examples of the benefits of running experiments online, where a local computer is used to drive the experiment, record and analyze the data, or both (see Fig 1).[332-335] The benefits granted by recording scientific data digitally are numerous. First, performing quantitative analysis is much easier on far larger scales using computational tools and methods. Some recent experiments only made possible by combining computational methods with STEM experiments include pm-precision measurements of atomic column positions,[192] atomic resolution 3D tomography[336] and imaging past the conventional information limit with ptychography.[317] The next step along these lines is to further close the loop between experiment and online analysis and perform data-driven STEM experiments.[337,338] Secondly, digital data can be shared easily with the wider scientific community, both in raw formats and after processing for further analysis, provided data transfer pipelines for big data are available. It has also become standard practice for many scientific journals to require authors to provide open access to data.[339] In parallel, the “open science" movement aims to increase transparency, efficiency, and reproducibility in scientific research.[340,341] Reproducibility is a major concern for all scientific disciplines including STEM[342]; researchers should make their data available in open source file formats such as ASCII, CSV, or hdf5 with all of the information required to understand and reproduce the results at the time of publication.

Data deposition can be performed using various online services such as the Materials Data Facility.[343] However, these services do not prescribe what metadata is required or any specific data formats and are therefore typically less useful than online databanks that target specific classes of data. The largest example in materials science is the Materials Project, which stores various calculated parameters for a large number of materials.[344] In biological electron microscopy, there are several examples of large databases that aid in transparency and reproducibility, including the Electron Microscopy Public Image Archive[345] and the Protein Data Bank.[346] Recently, a similar repository for atomic-resolution 3D structures for materials science applications called the Materials Data Bank[347] has also been established. There are a few examples of repositories of experimental or simulated STEM datasets such as in Ref [13,348]. There is currently, however, no repository that is both open-ended to support all kinds of STEM experiments while being specific enough to requires submission of all metadata and data required to reproduce published results. There is a strong need to develop this infrastructure for STEM research.[338,349,350]

## [H1] Limitations and Optimizations

The remarkable progress in high-resolution STEM imaging and spectroscopies over the last decades is based upon essentially unchanged instrument operation. The typical operation

of the microscope starts with stabilization tuning and focusing of the microscope, a process that can take minutes to hours and hinges upon predefined tuning routines and operator intuition. The operator then chooses the regions of interest on the sample and performs imaging and spectroscopy. The scanning process almost invariably uses a rectangular scanning pattern with beam rastering in a zig-zag shape. The process is well-recognized and accepted in the microcopy community but is often opaque to the general scientific community.

The development of ML and artificial intelligence methods over the past ten years has naturally led to the concept of automated and autonomous experiments in STEM and other microscopies, which typically includes the concept of a microscope that automatically chooses optimal imaging parameters, identifies the regions of interest on the sample, and performs the required measurements. In addition, the ability of STEM to potentially control beam-induced changes means it can be used as an atomic fabrication platform, which has been demonstrated by the creation of individual vacancies,[351,352] the insertion of dopant atoms,[26] the directed motion of atomic units,[26,353] and the assembly of homo- and heteroatomic molecules[354,355]. Below we touch on the requirements to enable the synergy between STEM and ML methods.

## [H2] Beam control and image reconstruction

The central premise for automated experiments is the ability to alter the probe path via external control, together with scanning along non-rectangular beam paths. While fairly straightforward, altering the probe raster path requires access to the microscope scan functions, which is fairly uncommon for high-resolution STEMs that often possess black-box manufacturer-developed controls. Scanning the beam using predefined non-rectangular paths such as spirals[229,230] has, however, been demonstrated.

The reconstruction of images from the data acquired along non-rectangular beam paths is another issue, which has been addressed using several approaches based on compressed sensing and a Gaussian process[230,356-358] or reconstructive autoencoders and VAEs. The compressed sensing and Gaussian process methods are generally interpolated between the acquired data points during the experiment and do not have inferential biases or prior information. Gaussian process methods also allow quantification of uncertainty maps and enable automated experiment workflows based on exploration or exploitation. In contrast, the VAE approach utilizes prior knowledge in the form of a trained network, which strongly increases the efficiency of this approach but makes it sensitive to out of distribution data.

## [H2] Automation and image-based feedback

Arbitrary scan paths provide a necessary engineering element for the development of automated experiment workflows. Another key element is the identification of objects of interest in the image plane, which depends on sensitivity to small variations in imaging conditions. For example, some of the early demonstrations of image-based feedback in STEM include detection of material crystallinity from the magnitude of the peak in the line-by-line fast Fourier transform of the image.[359] The broad introduction of deep learning image recognition networks offers another approach to automate experiments. However, despite significant and justified enthusiasm about potential of DCNNs in image analysis, these technologies are associated with strong inference biases. This implies that DCNNs will be sensitive to past data and generally sensitive to small changes in microscope parameters. A common challenge for pre-trained DCNNs comes from the distribution shift[360,361] when a neural network trained on one set of parameters (such as acquisition parameters, sample condition, etc.) fails to generalize to parameters outside of the training range. For spectroscopic experiments that aim to identify regions of sample where a particular behavior/functionality is

maximized (or minimized), one solution is the deep kernel learning approach[362] that actively learns a correlative relationship between structural data and functionality of interest and uses it to navigate the spectroscopic measurements.[24] Finally, target autoencoders are set by human operators and rapidly-emerging techniques such as curiosity-based learning offer an automation solution. In general, the likely pathways for adoption of autoencoders in STEM are high-level decisions at the human time scale with ML making fast low-level decisions.

## [H2] Atomic fabrication

Electron beams can induce changes in the structure of materials from amorphization to complete evaporation. Correspondingly, minimization of beam damage has been a primary driver in electron microscopy development, with its breakthrough achieved with the invention and rapid adoption of aberration correction, which enables high-resolution low-voltage imaging below many material's knock-on damage thresholds. In this regime, beam damage can often be sufficiently slow and associated with reconstruction of the chemical bond network rather than its complete destruction. Phenomena such as beam-induced phase transformations,[363] vacancy ordering,[364,365] and crystallization and amorphization,[366-370] have been reported. In monolayers[371] of graphene and layered dichalcogenides, STEM studies have demonstrated the presence of a broad spectrum of chemical transformations including vacancy formation,[351] grain boundary motion,[372] fast beam-induced motion of dopant atoms, and the dynamic motion of molecular groups.[373] Observing such precise dynamic changes under the action of a 50-100 kV electron beam is highly surprising and the mechanisms responsible for the observed phenomena are still actively discussed.[374-376 377,378] However, irrespective of the exact mechanisms, these observations clearly suggest the potential of using the electron beam for direct atomic fabrication. In bulk materials, beam-induced crystallization of silicon and strontium titanate was harnessed using image-based feedback.[379] Recently, the direct formation of vacancies and site-specific dopants, directed motion of silicon adatoms on graphene, and assembly of homo- and heteroatomic molecules have been demonstrated.[26] However, the atomic fabrication process to date has been driven by a human operator, typically at the time scale of minutes and tens of minutes for a single operation.

The simple examination of intrinsic latencies of STEM suggests that electron beam assembly can be accelerated by many orders of magnitude but necessitates the seamless combination of image recognition on potentially noisy and out of distribution data to identify the objects of interest together with reinforcement learning or similar methods.[380] Compared to well-popularized examples using reinforcement learning such as Atari or Go games, the rules that control electron beam transformations are unknown. Hence, we need either determination of these rules or algorithms that can base reinforcement learning on observations only such as muZero[381].

## [H1] Outlook

In this Primer, we hope to have illustrated the tremendous potential of deep learning for post-acquisition analysis, physics extraction from data, and especially automated and autonomous experiment. Fully realizing this potential necessitates significant developments on multiple levels, from instrumental platforms to common workflows, shared data and codes. These developments also necessitate progress in ML methods, which will be common to many other areas of physics.

On the instrument level, realizing the full potential of ML methods requires enabling open software architecture to control microscope operation and allow for custom experiments.

General software platforms for such development are well-established and exemplified by industry standards such as LabView. In recent years, some of the microscope manufacturers have provided open software for microscope operation such as Nion Swift and JEOL PyJem. Complementary to open software architectures will be the development of local computing capabilities to provide the computational power necessary to run complex calculations at the latencies of microscope operation. It should be noted that rapid progress in computational infrastructure now offers multiple opportunities for these developments, from extremely light computational platforms such as RaspberryPi and NVIDIA Jetson series to DGX and cluster solutions.

On the facility level, progress necessitates the development of universal yet flexible analysis workflows on sample preparation, imaging, and data analytics. This allows reproducible and traceable measurements and also serves as a necessary condition for transition to automated and autonomous experiments. Complementary to this are data repositories for storage of data, metadata providing context for the measurements and data meaning, and codes used in the original analysis. Examples of such workflow developments are given by cryo-EM imaging of biological systems.[382] It should be noted that historically, developments of workflows do not exclude the human operator from the research process. Rather these workflows allow delegation of the low-level, low-latency operations to automated systems so that a human operator can focus on high-level decision making.

On the STEM community level, there is a clear imperative for community-wide development and sharing of data analysis and, when possible, instrument control codes. Platforms such as GitHub that enable effective distributed code development are now mainstream and we hope to see the development of a code-sharing culture and credit within the community, from home institutions and from sponsor agencies. Equally important is data sharing, both within the STEM community and as a bridge to the broader physics community.

Finally, the most serendipitous developments are possible at the interface between STEM and the general scientific community. STEM offers a treasure trove of precise data on atomic position and functionalities linked to deep electronic levels, the Fermi level, and collective excitations. Extracting accurate physical information from this data and linking it to materials physics will revolutionize our understanding of condensed matter physics and chemistry on the atomic level. This will necessitate matching developments in ML, including physics-based ML, deep kernel learning, and active learning methods. Special sets of opportunities and requirements emerge in the context of autonomous experimentation, necessitating engineering controls, development and deployment of Bayesian optimization and reinforcement learning algorithms and their seamless integration into STEM workflows. These can enable automated tuning of the microscope, search and exploration of regions of specific interest, and ultimately atomic fabrication. To close this primer, we bring forth a quote from Feynman – “What I cannot make, I cannot understand.” With imaging, quantification, and fabrication capabilities enabled by the synergy of the experimental method with ML, STEM holds the promise for true understanding of the atomic world.

## Figures captions

**Figure 1**. **STEM as a quantitative tool.** a) schematic of a STEM microscope with different detector configurations. In STEM, the probe is an image of the source focused onto the sample, which is scanned across the specimen. Here we show the source at the bottom of the column which is typical of VG and Nion microscopes, but most manufacturers place the source at the top of the column. The detector plane is in diffraction space, and multiple detectors or a pixelated detector can be used as in 4D STEM. b) Adoption of ML methods will allow a transition from qualitative imaging of materials where only features of interest are extracted toward a more quantitative tool yielding information on the structural and electronic phenomena, collective excitations and their dispersions, and magnetic and spin effects. In-line ML methods further open pathways toward automated tuning and image optimization, automated exploration and discovery, and electron beam modification and atom by atom assembly.

**Figure 2. Angle-dependent STEM-EELS, 4D-STEM and beam engineering.** a) Schematic representation of one of the possible scattering geometries used in angle-dependent STEM-EELS experiments[70]. In this setup, the angular resolution is achieved by displacing the EELS spectrometer entrance (collection) aperture (schematically represented by the small overlapping red disks in the figure) along a vector $\boldsymbol{q}'$ relative to the optical axis. This relative shift is experimentally achieved by projecting the center of the spectrometer entrance aperture of the bright field disc (BF disk) using the microscope post specimen electronics. The aperture is shifted at increasingly higher values of momentum **q**, represented by the the resultant scattered wave vector $\boldsymbol{q} = \boldsymbol{k}_1 - \boldsymbol{k}_0$, where $\boldsymbol{k}_0$ and $\boldsymbol{k}_1$, are the incident and scattered vectors, respectively. In this representation the optical axis lies along the incident vector $k_0$ in the figure. The momentum resolution $\boldsymbol{\Delta q}$ is dependent on the beam convergence semi angle **α** and spectrometer collection angle **β**; larger beam convergence angles result in broad, overlapping diffraction discs, liming the q-space resolution of the spectral data. b) In 4D-STEM, simultaneous recordings of a STEM image and a 2D diffraction pattern for each probe position yields a highly redundant 4D dataset containing rich sample information. c-f. Examples of uses of structured illumination and detection in STEM. (c) Vortex beams can be used to probe magnetism, here a large quanta of Orbital angular momentum can probe vertical magnetic field (Reprinted from [122]), (d) Matched illumination and detector interferometry (MIDI)-STEM can be used to probe light material through linear contrast *(Reprinted from [120] ),*. (e) Gaussian probe can be used to simplify the electron probe propagation in crystals *Reprinted figure with permission from Rotunno et al. Phys. Rev. Appl. Electron-Beam Shaping in the Transmission Electron Microscope: Control of Electron-Beam Propagation Along Atomic Columns 11, 044072 (2019) Copyright (2019) by the American Physical Society* [117](f) Scanning a Bessel probe can be used for nano-diffraction and strain analysis *Reprinted from G. Guzzinati et al. Appl. Phys. Lett. 114, 243501 (2019) with the permission of AIP Publishing* [121].

**Figure 3: Identifying positions of atoms and atomic columns from STEM images.** (A) High quality HAADF-STEM image after distortion correction and averaging with refined atom positions marked by red dots. (B) Residual between image intensity in red box in (A) and best-fit Gaussian. (C) Lower-quality, single-shot ADF-STEM image of graphene. (D) Atomic positions determined using CNN. (E, F) Localization of point defects from noisy STEM data in graphene by applying a simple graph analysis to the output of a DCNN (green: Si, red: C) (A) and (B) adapted from Ref. [192]. (C) and (D) adapted from Ref. [204]. (E) and (F) adapted from [205]

**Figure Application of VAE to graphene. 4.** (a) Simplified schematic of VAE. The encoder (inference) network compresses input image data into a small number of latent variables. By default, the first three latent variables are designed to absorb rotations and translations of structures in the input images. Remaining latent variables aim at disentangling variations in the structure itself. (b) Encoded angle and (c) one of the latent variables for each atom in a snapshot (single STEM frame) of graphene undergoing structural transformations under e-beam irradiation. (d) Latent space manifold learned by VAE from data in unsupervised fashion.

**Figure 5. HAADF imaging of grain boundaries.** (a-c) HAADF STEM images of grain boundary migration in $Al_2O_3$, overlaid with structural models from simulated annealing, adapted from Ref [272]. (d-g) Reconstruction of three-dimensional atomistic structure of a $CdSe_{0.25}Te_{0.75}$ grain boundary.[276] (d) the experimental image, (e) the DFT optimized structure based on the experimental image, (f) a convolution image based on (e) and (g) the overlay of (d) and (f).

**Figure 6. STEM examples in chemistry**. (a) Schematic of single atom platinum migration on molybdenum disulfide 2D surface.[280] (b) High-magnification ADF-STEM image showing deposition of platinum nanocrystals on a grain boundary. Inset is fast Fourier transform of (b), in which one set of reflections is observed[281]. (c) Image of a region containing many platinum nanocrystals epitaxially grown a 2D surface[283]. (d) Magnified ADF-STEM image of yellow box in (c) showing platinum nanocrystal attached to the edge of molybdenum disulfide. (e) Fast Fourier transform analysis of (d) where reflections from platinum and molybdenum disulfide are labeled as white and yellow/green, respectively. (f) Structural change of photo-switches.[296] Schematic of isomerization of azobenzene derivative (left) and ADF-STEM image showing change of platinum marker distances (right).

**Figure 7. Bottom-up structural analysis for a ferroelectric material** (a) Experimental STEM image of La-doped $BiFeO_3$. (b) Local descriptors (sub-images) of different sizes centered on atoms from one of the sub-lattices extracted from CNN output. (c) Scree plot for principle component analysis (PCA) performed on the full stack of extracted descriptors (subimages) showing explained variance as a function of number of PCA components for different sized sub-images. (d) PCA decomposition of local descriptors into four components. PCA eigenmodes (associated with distortions) and corresponding loading maps (showing where distortion occurred in image) are shown in top and bottom rows of (d), respectively

## Boxes

**Box 1: Common machine learning approaches.**

Convolutional neural networks: A deep neural network with an input layer, hidden layers, and an output layer. Usually used for 2D image, they allow a computer to classify and discover patterns in an input image.

Bayesian optimisation: A method for finding the global minimum of a function that expensive to evaluate and for which the gradient is unknown.

Reinforcement learning: A machine learning method where an agent uses trial and error to learn from it previous actions. The reward function is supplied by the programmer.

Manifold learning: A method that transforms high dimensional data to a lower dimensional space.

Curiosity-based learning: A reinforcement learning method where the agent builds its own reward function.

Deep kernel learning: A method combining neural networks with uncertainty estimates using Gaussian processes.


## Acknowledgements:

This work 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 (S.V.K., M.P.O., and A.R.L.) and was performed and partially supported (M.Z.) at Oak Ridge National Laboratory's Center for Nanophase Materials Sciences (CNMS), a U.S. Department of Energy, Office of Science User Facility. V.G. acknowledge the support of the European Union Horizon 2020 Research and Innovation Programme under Grant Agreement No 766970 Q-SORT (H2020-FETOPEN-1-2016-2017). SuperSTEM (D.K.) is the UK National Facility for Advanced Electron Microscopy funded by the Engineering and Physical Sciences Research Council (EPSRC). P.M.V. acknowledges support form DOE BES (DE-FG02-08ER46547). G.G.D.H. and X.L. acknowledge the support from Brandeis NSF MRSEC, Bioinspired Soft Materials, DMR-2011486. Work at the Molecular Foundry (C.O.) was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. C.O. acknowledges support from a DOE Early Career Research Award. M.K.Y.C acknowledges support from the Center for Nanoscale Materials (DOE SUF) under Contract No. DE-AC02-06CH11357. C.O. and M.K.Y.C acknowledge support from DOE Early Career Research Awards. N.S. acknowledges support from the JSPS KAKENHI Grant numbers 20H05659 and 19H05788. J.E.acknowledges the support of the Australian Research Council Discovery Project grants DP150104483 and DP160104679. The authors are extremely grateful to Dr. Karren More (CNMS) for careful reading and editing the manuscript.

383

**Author contributions**

Introduction (SJP, MPO); Experimentation (CO, ARL, DK, RE, VG); Results (XL, GGDH,); Applications (ES, MKYC); Reproducibility and data deposition (CO, MZ); Limitations and optimizations (CO, SVK); Outlook (CO, MZ, SVK); Overview of the Primer (S.V.K.).

**Competing interests**

The authors declare no competing interests.

**Peer review information**

*Nature Reviews XXX* thanks [Referee#1 name], [Referee#2 name] and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Glossary

Aberration: imperfection usually referring to the electron optics of the microscope.
Coherent imaging: Measurements where the local contrast is dominated by the phase-alignment of the electron wavefronts; constructive interference (in phase) leads to higher signals and destructive inference (out of phase) leads to lower signals.
Incoherent imaging: When the coherence length of the electron waves are smaller than the resolution element of the measurement, the total signal is given incoherently by the sum of individual electron wavefunction intensities, and the relative phase of these wavefronts does not affect the measured intensity.
Contrast: spatial variation of intensity
Z-contrast imaging: STEM-HAADF imaging method, where the image contrast scales roughly monotonically with the atomic number Z of the atom(s) being imaged, approximately as $Z^{1.7}$.
Ptychography: A method of generating images from many coherent diffraction patterns formed at different probe positions in the STEM. It is also widely practiced in X-ray scattering experiments.
tilt series tomography: By tilting the specimen and recording projected images at different angles, computer algorithms can be used to reconstruct the 3D sample structure.
transmission modes: Imaging modes in electron microscopy where the electron beam passes through the specimen.
differential phase contrast: A method that measures the change in the convergent beam diffraction pattern as a function of probe position either using segmented or pixilated detector. These changes can be related to the local change in the sample's potential and corresponding fields.
Phonons: A quantized collective vibration of atoms in a crystalline sample, which can be excited by the electron beam and characterized by STEM-EELS or diffraction measurements.
plasmons: A quantized collective oscillation of electrons relative to the fixed ions in a sample, which can be excited by the electron beam and characterized by STEM-EELS.
core-loss edges: Excitation of inner shell electrons (ionization) by the electron beam, where the energy loss can be probed by STEM-EELS for features referred to as "edges."
dynamical scattering: A term commonly used in electron microscopy to describe the multiple scattering of the incident electron probe as it propagates through the specimen.
electron optical elements: Electromagnetic lenses used to focus or otherwise shape the electron beam.
electron holography: A technique for viewing the phase of the exit surface wave function using the interference of a scattered and unscattered electron beam.
azimuthal phase gradient: A wavefunction where the phase is a linearly proportional to the angle in polar coordinates, and the total phase shift is an integer multiple of 2 \pi for each revolution (see OAM).
orbital angular momentum: Orbital Angular Momentum are quanta given by the number of multiples of 2 \pi in the phase of an electron beam, per angular revolution in polar coordinates. (see Azimuthal Phase Gradient).
electron energy loss near edge structure: The intensity variation of the EELS signal as a function of energy loss near the onset of the core-loss signal.
$L_{23}$ ratio: The ratio of the L3 to L2 peaks formed by the transition of the $2p_{3/2}$ and $2p_{1/2}$ electrons to empty states.
Latent variables: A variable that is not directly observable, often obtained using VAEs.
Latent spaces: A vector space spanned by the latent variables.
Evidence lower boundary: The lower bound of the probability of observing a particular result for a given model.

Dark count rate *D*: *D* is the mean value of a STEM image acquired with the beam blanked preferably near the gun by, for example, closing the gun vacuum valve.
Gain *G*: Adjustment to ensure the measured signal covers the optimal range of the amplifier.
Faraday cup: A conductive cup which can capture charged free particles, which can estimate electron beam current by integrating the recorded signal.
Residual: The difference between the fitted image and the experimental image after atom location.
Penrose structures: **l**ocal structural units which, when displaced and rotated, can fully tile space, but do not have periodic translational symmetry. Such atomic structures can be found in quasicrystals.
Electron beam irradiation: when an electron beam induces changes in a specimen due to energy transfer, often called beam damage.
Dwell time: the time period of the data collection in each pixel
Inferential biases:
Exploration: uncertainty minimization
Exploitation: balancing exploration and pursuing target functionalities
Out of distribution data: when observational conditions change between experiments and preclude a direct comparison between experiments.
Distribution shift: in machine learning, when training and test sets do not come from the same distribution.
knock-on damage thresholds: The energy of the incident electron required to remove an atom from the crystal lattice.

# References


1 Keyse, R. J., Garratt-Reed, A. J., Goodhew, P. J. & Lorimer, G. W. *Introduction to scanning transmission electron microscopy*. (Routledge, 2018).

2 Shibata, N. *et al.* Electric field imaging of single atoms. *Nature Communications* **8**, 15631, doi:10.1038/ncomms15631 (2017).

3 Campanini, M., Nasi, L., Albertini, F. & Erni, R. Disentangling nanoscale electric and magnetic fields by time-reversal operation in differential phase-contrast STEM. *Appl. Phys. Lett.* **117**, 154102, doi:10.1063/5.0026121 (2020).

4 Müller, K. *et al.* Atomic electric fields revealed by a quantum mechanical approach to electron picodiffraction. *Nature Communications* **5**, 5653, doi:10.1038/ncomms6653 (2014).

5 Nellist, P. D., McCallum, B. C. & Rodenburg, J. M. RESOLUTION BEYOND THE INFORMATION LIMIT IN TRANSMISSION ELECTRON-MICROSCOPY. *Nature* **374**, 630-632, doi:10.1038/374630a0 (1995).

6 Hwang, J., Zhang, J. Y., D'Alfonso, A. J., Allen, L. J. & Stemmer, S. Three-Dimensional Imaging of Individual Dopant Atoms in $SrTiO_3$. *Physical Review Letters* **111**, 266101, doi:10.1103/PhysRevLett.111.266101 (2013).

7 Xin, H. L. & Muller, D. A. Aberration-corrected ADF-STEM depth sectioning and prospects for reliable 3D imaging in S/TEM. *Journal of Electron Microscopy* **58**, 157-165, doi:10.1093/jmicro/dfn029 (2009).

8 Benthem, K. v. *et al.* Three-dimensional imaging of individual hafnium atoms inside a semiconductor device. *Appl. Phys. Lett.* **87**, 034104, doi:10.1063/1.1991989 (2005).

9 Borisevich, A. Y., Lupini, A. R. & Pennycook, S. J. Depth sectioning with the aberration-corrected scanning transmission electron microscope. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3044-3048, doi:10.1073/pnas.0507105103 (2006).

10 Nellist, P. D. & Wang, P. Optical Sectioning and Confocal Imaging and Analysis in the Transmission Electron Microscope. *Annual Review of Materials Research* **42**, 125-143, doi:10.1146/annurev-matsci-070511-155016 (2012).

11 Wang, P. *et al.* Bright-field scanning confocal electron microscopy using a double aberration-corrected transmission electron microscope. *Ultramicroscopy* **111**, 877-886, doi:https://doi.org/10.1016/j.ultramic.2010.10.012 (2011).

12 Ishikawa, R. *et al.* Single atom visibility in STEM optical depth sectioning. *Appl. Phys. Lett.* **109**, 163102, doi:10.1063/1.4965709 (2016).

13 Levin, B. D. A. *et al.* Nanomaterial datasets to advance tomography in scanning transmission electron microscopy. *Scientific Data* **3**, 160041, doi:10.1038/sdata.2016.41 (2016).

14 Alania, M. *et al.* Depth sectioning combined with atom-counting in HAADF STEM to retrieve the 3D atomic structure. *Ultramicroscopy* **177**, 36-42, doi:https://doi.org/10.1016/j.ultramic.2016.11.002 (2017).

15 Dahmen, T. *et al.* Combined Scanning Transmission Electron Microscopy Tilt- and Focal Series. *Microscopy and Microanalysis* **20**, 548-560, doi:10.1017/S1431927614000075 (2014).

16 Hovden, R. *et al.* Breaking the Crowther limit: Combining depth-sectioning and tilt tomography for high-resolution, wide-field 3D reconstructions. *Ultramicroscopy* **140**, 26-31, doi:https://doi.org/10.1016/j.ultramic.2014.01.013 (2014).

17 Krivanek, O. L. *et al.* Vibrational spectroscopy in the electron microscope. *Nature* **514**, 209-212, doi:10.1038/nature13870 (2014).

18 Idrobo, J. C. *et al.* Temperature Measurement by a Nanoscale Electron Probe Using Energy Gain and Loss Spectroscopy. *Physical Review Letters* **120**, 095901, doi:10.1103/PhysRevLett.120.095901 (2018).

19 Bonnet, N. Multivariate statistical methods for the analysis of microscope image series: applications in materials science. *Journal of Microscopy* **190**, 2-18, doi:https://doi.org/10.1046/j.1365-2818.1998.3250876.x (1998).

20 Trebbia, P. & Bonnet, N. EELS elemental mapping with unconventional methods I. Theoretical basis: Image analysis with multivariate statistics and entropy concepts. *Ultramicroscopy* **34**, 165-178, doi:https://doi.org/10.1016/0304-3991(90)90070-3 (1990).

21 Trebbia, P. & Mory, C. EELS elemental mapping with unconventional methods II. Applications to biological specimens. *Ultramicroscopy* **34**, 179-203, doi:https://doi.org/10.1016/0304-3991(90)90071-S (1990).

22 Dan, J., Zhao, X. & Pennycook, S. J. A machine perspective of atomic defects in scanning transmission electron microscopy. *InfoMat* **1**, 359-375, doi:https://doi.org/10.1002/inf2.12026 (2019).

23 Maksov, A. *et al.* Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2. *npj Computational Materials* **5**, 12, doi:10.1038/s41524-019-0152-9 (2019).

24 Roccapriore, K. M., Kalinin, S. V. & Ziatdinov, M. Physics discovery in nanoplasmonic systems via autonomous experiments in Scanning Transmission Electron Microscopy. *arXiv preprint arXiv:2108.03290* (2021).

25 Susi, T. *et al.* Towards atomically precise manipulation of 2D nanostructures in the electron microscope. *2D Materials* **4**, 042004, doi:10.1088/2053-1583/aa878f (2017).

26 Dyck, O., Kim, S., Kalinin, S. V. & Jesse, S. Placing single atoms in graphene with a scanning transmission electron microscope. *Appl. Phys. Lett.* **111**, 113104, doi:10.1063/1.4998599 (2017).

27 Dyck, O. *et al.* Electron-beam introduction of heteroatomic Pt–Si structures in graphene. *Carbon* **161**, 750-757, doi:https://doi.org/10.1016/j.carbon.2020.01.042 (2020).

28 Oxley, M. P., Lupini, A. R. & Pennycook, S. J. Ultra-high resolution electron microscopy. *Reports on Progress in Physics* **80**, 026101, doi:10.1088/1361-6633/80/2/026101 (2016).

29 Loth, S., Etzkorn, M., Lutz, C. P., Eigler, D. M. & Heinrich, A. J. Measurement of Fast Electron Spin Relaxation Times with Atomic Resolution. *Science* **329**, 1628-1630, doi:doi:10.1126/science.1191688 (2010).

30 Giessibl, F. J. AFM's path to atomic resolution. *Materials Today* **8**, 32-41, doi:https://doi.org/10.1016/S1369-7021(05)00844-8 (2005).

31 Park, J. Y., Maier, S., Hendriksen, B. & Salmeron, M. Sensing current and forces with SPM. *Materials Today* **13**, 38-45, doi:https://doi.org/10.1016/S1369-7021(10)70185-1 (2010).

32 Abbe, E. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für mikroskopische Anatomie* **9**, 413-468 (1873).

33 Scherzer, O. Uber einige Fehler von Elektonenlinsen. *Zeit. Phys.* **101**, 593-603 (1936).

34 Hawkes, P. W. in *Science of Microscopy* (ed Peter W. Hawkes & John C. Spence) 696-747 (Springer, 2007).

35 Dellby, N., Krivanek, O. L., Nellist, P. D., Batson, P. E. & Lupini, A. R. Progress in aberration-corrected scanning transmission electron microscopy. *Journal of Electron Microscopy* **50**, 177-185, doi:10.1093/jmicro/50.3.177 (2001).

36 Krivanek, O. L., Dellby, N. & Lupini, A. R. Towards sub-angstrom electron beams. *Ultramicroscopy* **78**, 1-11 (1999).

37 Haider, M. *et al.* A spherical-aberration-corrected 200 kV transmission electron microscope. *Ultramicroscopy* **75**, 53-60 (1998).

38 Haider, M. *et al.* Electron microscopy image enhanced. *Nature* **392**, 768-769, doi:10.1038/33823 (1998).

39 Rose, H. Elektronenoptische aplanate. *Optik* **34**, 285-311 (1971).

40 Zach, J. & Haider, M. Aberration correction in a low voltage SEM by a multipole corrector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **363**, 316-325 (1995).

41 Batson, P. E. Aberration correction results in the IBM STEM instrument. *Ultramicroscopy* **96**, 239-249 (2003).

42 Nellist, P. D. *et al.* Direct sub-angstrom imaging of a crystal lattice. *Science* **305**, 1741-1741, doi:10.1126/science.1100965 (2004).

43 Kisielowski, C. *et al.* Detection of single atoms and buried defects in three dimensions by aberration-corrected electron microscope with 0.5-angstrom information limit. *Microscopy and Microanalysis* **14**, 469-477, doi:10.1017/s1431927608080902 (2008).

44 Shibata, N. *et al.* Differential phase-contrast microscopy at atomic resolution. *Nature Physics* **8**, 611-615, doi:10.1038/nphys2337 (2012).

45 Shibata, N. *et al.* New area detector for atomic-resolution scanning transmission electron microscopy. *Journal of electron microscopy* **59**, 473-479, doi:10.1093/jmicro/dfq014 (2010).

46 Ophus, C. Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond. *Microscopy and Microanalysis* **25**, 563-582, doi:10.1017/s1431927619000497 (2019).

47 Egerton, R. F. *Electron Energy-Loss Spectroscopy in the Electron Microscope*. (Springer, 2011).

48 Baker, J. H. a. R. F. Microanalysis by Means of Electrons. *Journal of Applied Physics* **15**, 663, doi:10.1063/1.1707491 (1944).

49 Gubbens, A. J., Kraus, B., Krivanek, O. L. & Mooney, P. An imaging filter for high voltage electron microscopy. *Ultramicroscopy* **59**, 255-265, doi:10.1016/0304-3991(95)00033-W (1995).

50 Silcox, J. Core-loss EELS. *Current Opinion in Solid State & Materials Science* **3**, 336-342, doi:10.1016/s1359-0286(98)80042-9 (1998).

51 Nelayah, J. *et al.* Mapping surface plasmons on a single metallic nanoparticle. *Nature Physics* **3**, 348-353, doi:10.1038/nphys575 (2007).

52 Zhou, W. *et al.* Atomically localized plasmon enhancement in monolayer graphene. *Nat. Nanotechnol.* **7**, 161-165, doi:10.1038/nnano.2011.252 (2012).

53 Suenaga, K. & Koshino, M. Atom-by-atom spectroscopy at graphene edge. *Nature* **468**, 1088-1090, doi:10.1038/nature09664 (2010).

54 Colliex, C., Kociak, M. & Stéphan, O. Electron Energy Loss Spectroscopy imaging of surface plasmons at the nanometer scale. *Ultramicroscopy* **162**, A1-A24, doi:https://doi.org/10.1016/j.ultramic.2015.11.012 (2016).

55 El-Sherif, H., Jovanovic, S., Preston, J. & Basim, N. Electron Microscopy and Interface Plasmons Characterization of Cadmium Telluride Thin Film Grown Incommensurately with Weak Bonding on Sapphire. *Microscopy and Microanalysis* **26**, 3116-3118 (2020).

56 Yang, H., Garfunkel, E. L. & Batson, P. E. Probing free carrier plasmons in doped semiconductors using spatially resolved electron energy loss spectroscopy. *Physical Review B* **102**, 205427 (2020).

57 Bonnet, N., Brun, N. & Colliex, C. Extracting information from sequences of spatially resolved EELS spectra using multivariate statistical analysis. *Ultramicroscopy* **77**, 97-112 (1999).

58 Jeanguillaume, C. & Colliex, C. Spectrum-image: the next step in EELS digital acquisition and processing. *Ultramicroscopy* **28**, 252-257 (1989).

59 Polman, A., Kociak, M. & García de Abajo, F. J. Electron-beam spectroscopy for nanophotonics. *Nature Materials* **18**, 1158-1171, doi:10.1038/s41563-019-0409-1 (2019).

60 Hachtel, J., Lupini, A. & Idrobo, J. Exploring the capabilities of monochromated electron energy loss spectroscopy in the infrared regime. *Scientific Reports* **8**, doi:10.1038/s41598-018-23805-5 (2018).

61 Krivanek, O. *et al.* Towards sub-10 meV energy resolution STEM-EELS. *Journal of Physics: Conference Series* **522**, 012023, doi:10.1088/1742-6596/522/1/012023 (2014).

62 Krivanek, O. L. *et al.* Towards sub-10 meV energy resolution STEM-EELS. *Journal of Physics: Conference Series* **522**, 012023, doi:10.1088/1742-6596/522/1/012023 (2014).

63 Hage, F. S. *et al.* Nanoscale momentum-resolved vibrational spectroscopy. *Sci. Adv.* **4**, 6, doi:10.1126/sciadv.aar7495 (2018).

64 Hage, F. S., Radtke, G., Kepaptsoglou, D. M., Lazzeri, M. & Ramasse, Q. M. Single-atom vibrational spectroscopy in the scanning transmission electron microscope. *Science* **367**, 1124-+, doi:10.1126/science.aba1136 (2020).

65 Lagos, M. J., Trügler, A., Hohenester, U. & Batson, P. E. Mapping vibrational surface and bulk modes in a single nanocube. *Nature* **543**, 529-532, doi:10.1038/nature21699 (2017).

66 Boersch, H., Geiger, J. & Stickel, W. Interaction of 25-keV Electrons with Lattice Vibrations in LiF. Experimental Evidence for Surface Modes of Lattice Vibration. *Physical Review Letters* **17**, 379-381, doi:10.1103/PhysRevLett.17.379 (1966).

67 Egerton, R. F. Electron energy-loss spectroscopy in the TEM. *Reports on Progress in Physics* **72**, 25, doi:10.1088/0034-4885/72/1/016502 (2009).

68 Leapman, R. D., Fejes, P. L. & Silcox, J. ORIENTATION DEPENDENCE OF CORE EDGES FROM ANISOTROPIC MATERIALS DETERMINED BY INELASTIC-SCATTERING OF FAST ELECTRONS. *Physical Review B* **28**, 2361-2373, doi:10.1103/PhysRevB.28.2361 (1983).

69 Botton, G. A. A new approach to study bonding anisotropy with EELS. *J. Electron Spectrosc. Relat. Phenom.* **143**, 129-137, doi:10.1016/j.elspec.2004.09.023 (2005).

70 Hage, F. S. *et al.* Topologically induced confinement of collective modes in multilayer graphene nanocones measured by momentum-resolved STEM-VEELS. *Physical Review B* **88**, 12, doi:10.1103/PhysRevB.88.155408 (2013).

71 Schattschneider, P. Exchange of angular momentum in EMCD experiments. *Ultramicroscopy* **109**, 91-95, doi:10.1016/j.ultramic.2008.08.004 (2008).

72 Wang, Y. Y. *et al.* EVOLUTION OF THE LOW-ENERGY EXCITATIONS AND DIELECTRIC FUNCTION OF BA1-XKXBIO3(0-LESS-THAN-OR-EQUAL-TO-X-LESS-THAN-OR-EQUAL-TO-0.50). *Physical Review B* **47**, 14503-14509, doi:10.1103/PhysRevB.47.14503 (1993).

73 Liou, S. C. *et al.* Plasmons dispersion and nonvertical interband transitions in single crystal Bi2Se3 investigated by electron energy-loss spectroscopy. *Physical Review B* **87**, 6, doi:10.1103/PhysRevB.87.085126 (2013).

74 Schuster, R., Kraus, R., Knupfer, M., Berger, H. & Buchner, B. Negative plasmon dispersion in the transition-metal dichalcogenide 2H-TaSe2. *Physical Review B* **79**, 5, doi:10.1103/PhysRevB.79.045134 (2009).

75 Gloter, A., Chu, M. W., Kociak, M., Chen, C. H. & Colliex, C. Probing non-dipole allowed excitations in highly correlated materials with nanoscale resolution. *Ultramicroscopy* **109**, 1333-1337, doi:10.1016/j.ultramic.2009.06.005 (2009).

76 Senga, R. *et al.* Position and momentum mapping of vibrations in graphene nanostructures. *Nature* **573**, 247-+, doi:10.1038/s41586-019-1477-8 (2019).

77 Egerton, R. F. Limits to the spatial, energy and momentum resolution of electron energy-loss spectroscopy. *Ultramicroscopy* **107**, 575-586, doi:10.1016/j.ultramic.2006.11.005 (2007).

78 Wang, Y. Y., Cheng, S. C., Dravid, V. P. & Zhang, F. C. MOMENTUM-TRANSFER RESOLVED ELECTRON-ENERGY-LOSS SPECTROSCOPY OF SOLIDS - PROBLEMS, SOLUTIONS AND APPLICATIONS. *Ultramicroscopy* **59**, 109-119, doi:10.1016/0304-3991(95)00022-s (1995).

79 Plotkin-Swing, B. *et al.* Hybrid pixel direct detector for electron energy loss spectroscopy. *Ultramicroscopy* **217**, 10, doi:10.1016/j.ultramic.2020.113067 (2020).

80 Pogany, A. P. & Turner, P. S. Reciprocity in Electron Diffraction and Microscopy. *Acta Crystallographica Section A* **24**, 103-109 (1968).

81 Frank, J. Envelope Of Electron-Microscopic Transfer-Functions For Partially Coherent Illumination. *Optik* **38**, 519-536 (1973).

82 Cowley, J. M. Electron nanodiffraction. *Microscopy Research and Technique* **46**, 75-97, doi:10.1002/(SICI)1097-0029(19990715)46:2<75::AID-JEMT2>3.0.CO;2-S (1999).

83 Cowley, J. M. Adjustment of a STEM Instrument by Use of Shadow Images. *Ultramicroscopy* **4**, 413-418 (1979).

84 Cowley, J. M. High-Resolution Electron-Microscopy and Microdiffraction. *Ultramicroscopy* **18**, 11-17 (1985).

85 Cowley, J. M. in *Advances in electronics and electron physics* Vol. 46 1-53 (1978).
86 Ronchi, V. Forty Years of History of a Grating Interferometer. *Applied Optics* **3**, 437-451 (1964).
87 Lupini, A. R. & Pennycook, S. J. Rapid autotuning for crystalline specimens from an inline hologram. *Journal of Electron Microscopy* **57** (2008).
88 Lupini, A. R., Wang, P., Nellist, P. D., Kirkland, A. I. & Pennycook, S. J. Aberration measurement using the Ronchigram contrast transfer function. *Ultramicroscopy* **110**, 891-898 (2010).
89 Sawada, H., Allen, C. S., Wang, S., Warner, J. H. & Kirkland, A. I. Aberration measurement of the probe-forming system of an electron microscope using two-dimensional materials. *Ultramicroscopy* **182**, 195-204, doi:https://doi.org/10.1016/j.ultramic.2017.06.024 (2017).
90 Ramasse, Q. M. Twenty years after: How "Aberration correction in the STEM" truly placed a "A synchrotron in a Microscope". *Ultramicroscopy* **180**, 41-51, doi:https://doi.org/10.1016/j.ultramic.2017.03.016 (2017).
91 Ramasse, Q. M. & Bleloch, A. L. Diagnosis of aberrations from crystalline samples in scanning transmission electron microscopy. *Ultramicroscopy* **106**, 37-56 (2005).
92 Sawada, H. *et al.* Measurement method of aberration from Ronchigram by autocorrelation function. *Ultramicroscopy* **108**, 1467-1475, doi:https://doi.org/10.1016/j.ultramic.2008.04.095 (2008).
93 Lin, J. A. & Cowley, J. M. Calibration of the operating parameters for an HB5 stem instrument. *Ultramicroscopy* **19**, 31-42, doi:https://doi.org/10.1016/0304-3991(86)90005-7 (1986).
94 Dwyer, C., Erni, R. & Etheridge, J. Method to measure spatial coherence of subangstrom electron beams. *Appl. Phys. Lett.* **93**, 3, doi:10.1063/1.2957648 (2008).
95 Dwyer, C., Erni, R. & Etheridge, J. Measurement of effective source distribution and its importance for quantitative interpretation of STEM images. *Ultramicroscopy* **110**, 952-957, doi:10.1016/j.ultramic.2010.01.007 (2010).
96 Lupini, A. R., Chi, M. & Jesse, S. Rapid aberration measurement with pixelated detectors. *Journal of Microscopy* **263**, 43-50, doi:https://doi.org/10.1111/jmi.12372 (2016).
97 Lupini, A. R. *et al.* Fast Aberration Measurement in Multi-Dimensional STEM. *Microscopy and Microanalysis* **22**, 252-253, doi:10.1017/S1431927616002117 (2016).
98 Pennycook, T. J. *et al.* Efficient phase contrast imaging in STEM using a pixelated detector. Part 1: Experimental demonstration at atomic resolution. *Ultramicroscopy* **151**, 160-167, doi:https://doi.org/10.1016/j.ultramic.2014.09.013 (2015).
99 Krajnak, M. & Etheridge, J. A symmetry-derived mechanism for atomic resolution imaging. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 27805-27810, doi:10.1073/pnas.2006975117 (2020).
100 N. H. Dekkers, H. d. L. Differential Phase Contrast in a STEM. *Optik* **41**, 452-456 (1974).
101 Lazić, I., Bosch, E. G. T. & Lazar, S. Phase contrast STEM for thin samples: Integrated differential phase contrast. *Ultramicroscopy* **160**, 265-280, doi:https://doi.org/10.1016/j.ultramic.2015.10.011 (2016).
102 Tsuda, K. & Tanaka, M. Direct observation of the symmetry breaking of the nanometer-scale local structure in the paraelectric cubic phase of BaTiO3 using convergent-beam electron diffraction. *Appl. Phys. Express* **9**, 4, doi:10.7567/apex.9.071501 (2016).
103 Jesse, S. *et al.* Big Data Analytics for Scanning Transmission Electron Microscopy Ptychography. *Scientific Reports* **6**, doi:26348
10.1038/srep26348 (2016).
104 Nord, M. *et al.* Strain Anisotropy and Magnetic Domains in Embedded Nanomagnets. *Small* **15**, 6, doi:10.1002/smll.201904738 (2019).
105 Pekin, T. C., Gammer, C., Ciston, J., Minor, A. M. & Ophus, C. Optimizing disk registration algorithms for nanobeam electron diffraction strain mapping. *Ultramicroscopy* **176**, 170-176, doi:10.1016/j.ultramic.2016.12.021 (2017).

106 Grieb, T. *et al.* Strain analysis from nano-beam electron diffraction: Influence of specimen tilt and beam convergence. *Ultramicroscopy* **190**, 45-57, doi:10.1016/j.ultramic.2018.03.013 (2018).

107 Han, Y. M. *et al.* Strain Mapping of Two-Dimensional Heterostructures with Subpicometer Precision. *Nano Lett.* **18**, 3746-3751, doi:10.1021/acs.nanolett.8b00952 (2018).

108 Nord, M. *et al.* Three-dimensional subnanoscale imaging of unit cell doubling due to octahedral tilting and cation modulation in strained perovskite thin films. *Phys. Rev. Mater.* **3**, 7, doi:10.1103/PhysRevMaterials.3.063605 (2019).

109 Hoppe, W. Trace structure analysis, ptychography, phase tomography. *Ultramicroscopy* **10**, 187-198, doi:https://doi.org/10.1016/0304-3991(82)90038-9 (1982).

110 R. W. Gerchberg, W. O. S. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* **35**, 227-246 (1972).

111 Cowley, J. M. Comments on ultra-high-resolution STEM. *Ultramicroscopy* **87**, 1-4 (2001).

112 Nellist, P. D. & Rodenburg, J. M. Beyond the Conventional Information Limit - the Relevant Coherence Function. *Ultramicroscopy* **54**, 61-74 (1994).

113 Rodenburg, J. M. & Bates, R. H. T. The theory of superresolution electron-microscopy via Wigner-distribution deconvolution. *Philosophical Transactions of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* **339**, 521-553 (1992).

114 Lupini, A. R., Oxley, M. P. & Kalinin, S. V. Pushing the limits of electron ptychography. *Science* **362**, 399, doi:10.1126/science.aau7620 (2018).

115 Yang, H. *et al.* Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures. *Nature Communications* **7**, 12532, doi:10.1038/ncomms12532 (2016).

116 Van Dyck, D., Jinschek, J. R. & Chen, F.-R. 'Big Bang' tomography as a new route to atomic-resolution electron tomography. *Nature* **486**, 243-246, doi:10.1038/nature11074 (2012).

117 Verbeeck, J., Tian, H. & Schattschneider, P. Production and application of electron vortex beams. *Nature* **467**, 301-304, doi:10.1038/nature09366 (2010).

118 McMorran, B. J. *et al.* Electron Vortex Beams with High Quanta of Orbital Angular Momentum. *Science* **331**, 192-195, doi:10.1126/science.1198804 (2011).

119 Grillo, V. *et al.* Generation of Nondiffracting Electron Bessel Beams. *Phys. Rev. X* **4**, 7, doi:10.1103/PhysRevX.4.011013 (2014).

120 Uchida, M. & Tonomura, A. Generation of electron beams carrying orbital angular momentum. *Nature* **464**, 737-739 (2010).

121 Rotunno, E. *et al.* Electron-Beam Shaping in the Transmission Electron Microscope: Control of Electron-Beam Propagation Along Atomic Columns. *Phys. Rev. Appl.* **11**, 10, doi:10.1103/PhysRevApplied.11.044072 (2019).

122 Nguyen, D. T., Findlay, S. D. & Etheridge, J. A menu of electron probes for optimising information from scanning transmission electron microscopy. *Ultramicroscopy* **184**, 143-155, doi:10.1016/j.ultramic.2017.08.020 (2018).

123 Zheng, C. L. *et al.* Axicon Lens for Electrons Using a Magnetic Vortex: The Efficient Generation of a Bessel Beam. *Physical Review Letters* **119**, 6, doi:10.1103/PhysRevLett.119.174801 (2017).

124 Verbeeck, J. *et al.* Demonstration of a 2 x 2 programmable phase plate for electrons. *Ultramicroscopy* **190**, 58-65, doi:10.1016/j.ultramic.2018.03.017 (2018).

125 Tavabi, A. H. *et al.* Generation of electron vortices using nonexact electric fields. *Phys. Rev. Res.* **2**, 5, doi:10.1103/PhysRevResearch.2.013185 (2020).

126 Tavabi, A. H. *et al.* Experimental demonstration of an electrostatic orbital angular momentum sorter for electrons. *arXiv preprint arXiv:.03706* (2019).

127 Grillo, V. *et al.* Observation of nanoscale magnetic fields using twisted electron beams. *Nature Communications* **8**, 6, doi:10.1038/s41467-017-00829-5 (2017).

128 Edstrom, A., Lubk, A. & Rusz, J. Quantum mechanical treatment of atomic-resolution differential phase contrast imaging of magnetic materials. *Physical Review B* **99**, 9, doi:10.1103/PhysRevB.99.174428 (2019).

129 Pohl, D. *et al.* Atom size electron vortex beams with selectable orbital angular momentum. *Scientific Reports* **7**, 8, doi:10.1038/s41598-017-01077-9 (2017).

130 Rusz, J. & Bhowmick, S. Boundaries for Efficient Use of Electron Vortex Beams to Measure Magnetic Properties. *Physical Review Letters* **111**, 5, doi:10.1103/PhysRevLett.111.105504 (2013).

131 Ophus, C. *et al.* Efficient linear phase contrast in scanning transmission electron microscopy with matched illumination and detector interferometry. *Nature Communications* **7**, 7, doi:10.1038/ncomms10719 (2016).

132 Guzzinati, G. *et al.* Probing the symmetry of the potential of localized surface plasmon resonances with phase-shaped electron beams. *Nature Communications* **8**, 8, doi:10.1038/ncomms14999 (2017).

133 Zeltmann, S. E. *et al.* Patterned probes for high precision 4D-STEM bragg measurements. *Ultramicroscopy* **209**, 9, doi:10.1016/j.ultramic.2019.112890 (2020).

134 Guzzinati, G. *et al.* Electron Bessel beam diffraction for precise and accurate nanoscale strain mapping. *Appl. Phys. Lett.* **114**, 5, doi:10.1063/1.5096245 (2019).

135 Stoger-Pollach, M., Schachinger, T., Biedermann, K. & Beyer, V. Valence EELS below the limit of inelastic delocalization using conical dark field EFTEM or Bessel beams. *Ultramicroscopy* **173**, 24-30, doi:10.1016/j.ultramic.2016.11.022 (2017).

136 Nellist, P. D. & Pennycook, S. J. Incoherent imaging using dynamically scattered coherent electrons. *Ultramicroscopy* **78**, 111-124, doi:10.1016/s0304-3991(99)00017-0 (1999).

137 Grillo, V. *et al.* Measuring the orbital angular momentum spectrum of an electron beam. *Nature Communications* **8**, 6, doi:10.1038/ncomms15536 (2017).

138 Rotunno, E. *et al.* Alignment of electron optical beam shaping elements using a convolutional neural network. *arXiv preprint arXiv:.04358* (2021).

139 Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & Freitas, N. d. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE* **104**, 148-175, doi:10.1109/JPROC.2015.2494218 (2016).

140 Silver, D. *et al.* A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362**, 1140-1144 (2018).

141 Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018, doi:10.1038/sdata.2016.18 (2016).

142 Collins, T. J. ImageJ for microscopy. *Biotechniques* **43**, 25-+ (2007).

143 Roels, J. *et al.* An interactive ImageJ plugin for semi-automated image denoising in electron microscopy. *Nature Communications* **11**, 13, doi:10.1038/s41467-020-14529-0 (2020).

144 Somnath, S. *et al.* USID and Pycroscopy--Open frameworks for storing and analyzing spectroscopic and imaging data. *arXiv preprint arXiv:.09515* (2019).

145 Clausen, A. *et al.* LiberTEM: Software platform for scalable multidimensional data processing in transmission electron microscopy. *Journal of Open Source Software* **5**, 2006 (2020).

146 Savitzky, B. H. *et al.* py4DSTEM: Open Source Software for 4D-STEM Data Analysis. *Microscopy and Microanalysis* **25**, 124-125 (2019).

147 Koch, C. T. Determination of core structure periodicity and point defect density along dislocations. *Ph. D. Thesis* (2002).

148 Allen, L. J., D'Alfonso, A. J. & Findlay, S. D. Modelling the inelastic scattering of fast electrons. *Ultramicroscopy* **151**, 11-22, doi:https://doi.org/10.1016/j.ultramic.2014.10.011 (2015).

149 Barthel, J. Dr. Probe: A software for high-resolution STEM image simulation. *Ultramicroscopy* **193**, 1-11, doi:https://doi.org/10.1016/j.ultramic.2018.06.003 (2018).

150 Lobato, I. & Van Dyck, D. MULTEM: A new multislice program to perform accurate and fast electron diffraction and imaging simulations using Graphics Processing Units with CUDA. *Ultramicroscopy* **156**, 9-17, doi:https://doi.org/10.1016/j.ultramic.2015.04.016 (2015).

151 Oelerich, J. O. *et al.* STEMsalabim: A high-performance computing cluster friendly code for scanning transmission electron microscopy image simulations of thin specimens. *Ultramicroscopy* **177**, 91-96, doi:https://doi.org/10.1016/j.ultramic.2017.03.010 (2017).

152 Madsen, J. & Susi, T. abTEM: ab Initio Transmission Electron Microscopy Image Simulation. *Microscopy and Microanalysis* **26**, 448-450, doi:10.1017/S1431927620014701 (2020).

153 Pryor, A., Ophus, C. & Miao, J. A streaming multi-GPU implementation of image simulation algorithms for scanning transmission electron microscopy. *Advanced Structural and Chemical Imaging* **3**, 15, doi:10.1186/s40679-017-0048-z (2017).

154 Mortensen, J. J., Hansen, L. B. & Jacobsen, K. W. Real-space grid implementation of the projector augmented wave method. *Physical Review B* **71**, 035109, doi:10.1103/PhysRevB.71.035109 (2005).

155 Sakaguchi, N., Matsumoto, S., Kunisada, Y. & Ueda, M. Interaction of Localized Surface Plasmons of a Silver Nanosphere Dimer Embedded in a Uniform Medium: Scanning Transmission Electron Microscopy Electron Energy-Loss Spectroscopy and Discrete Dipole Approximation Simulation. *The Journal of Physical Chemistry C* **123**, 6735-6744, doi:10.1021/acs.jpcc.8b11434 (2019).

156 Mendis, B. G. An inelastic multislice simulation method incorporating plasmon energy losses. *Ultramicroscopy* **206**, 112816, doi:https://doi.org/10.1016/j.ultramic.2019.112816 (2019).

157 Brown, H. G., Ciston, J. & Ophus, C. Linear-scaling algorithm for rapid computation of inelastic transitions in the presence of multiple electron scattering. *Phys. Rev. Res.* **1**, 033186, doi:10.1103/PhysRevResearch.1.033186 (2019).

158 Tessmer, J., Singh, S., Gu, Y., El-Awady, J. A. & Graef, M. D. Scanning transmission electron microscopy image simulations of complex dislocation structures generated by discrete dislocation dynamics. *Ultramicroscopy* **219**, 113124, doi:https://doi.org/10.1016/j.ultramic.2020.113124 (2020).

159 Thomas, P. J. & Midgley, P. A. Image-spectroscopy--II. The removal of plural scattering from extended energy-filtered series by Fourier deconvolution. *Ultramicroscopy* **88**, 187-194, doi:10.1016/s0304-3991(01)00078-x (2001).

160 Leapman, R. D., Rez, P. & Mayers, D. F. K, L, and M shell generalized oscillator strengths and ionization cross sections for fast electron collisions. *The Journal of Chemical Physics* **72**, 1232-1243, doi:10.1063/1.439184 (1980).

161 Ahn, C. C. & Rez, P. Inner shell edge profiles in electron energy loss spectroscopy. *Ultramicroscopy* **17**, 105-115, doi:https://doi.org/10.1016/0304-3991(85)90003-8 (1985).

162 Cueva, P., Hovden, R., Mundy, J. A., Xin, H. L. & Muller, D. A. Data Processing for Atomic Resolution Electron Energy Loss Spectroscopy. *Microscopy and Microanalysis* **18**, 667-675, doi:10.1017/S1431927612000244 (2012).

163 Varela, M. *et al.* Atomic-resolution imaging of oxidation states in manganites. *Physical Review B* **79**, 085117 (2009).

164 Verbeeck, J., Van Aert, S. & Bertoni, G. Model-based quantification of EELS spectra: Including the fine structure. *Ultramicroscopy* **106**, 976-980 (2006).

165 Bosman, M. *et al.* Two-dimensional mapping of chemical information at atomic resolution. *Physical Review Letters* **99**, 086102 (2007).

166 Bosman, M., Watanabe, M., Alexander, D. & Keast, V. Mapping chemical and bonding information using multivariate analysis of electron energy-loss spectrum images. *Ultramicroscopy* **106**, 1024-1032 (2006).

167 Thersleff, T. *et al.* Chemically and morphologically distinct grain boundaries in Ge-doped Cu2ZnSnSe4 solar cells revealed with STEM-EELS. *Materials & Design* **122**, 102-109, doi:https://doi.org/10.1016/j.matdes.2017.02.077 (2017).

168 Potapov, P. & Lubk, A. Extraction of physically meaningful endmembers from STEM spectrum-images combining geometrical and statistical approaches. *Micron* **145**, 103068, doi:https://doi.org/10.1016/j.micron.2021.103068 (2021).
169 Bobynko, J., MacLaren, I. & Craven, A. J. Spectrum imaging of complex nanostructures using DualEELS: I. digital extraction replicas. *Ultramicroscopy* **149**, 9-20, doi:https://doi.org/10.1016/j.ultramic.2014.10.014 (2015).
170 Craven, A. J., Sala, B., Bobynko, J. & MacLaren, I. Spectrum imaging of complex nanostructures using DualEELS: II. Absolute quantification using standards. *Ultramicroscopy* **186**, 66-81, doi:https://doi.org/10.1016/j.ultramic.2017.12.011 (2018).
171 Shiga, M. *et al.* Sparse modeling of EELS and EDX spectral imaging data by nonnegative matrix factorization. *Ultramicroscopy* **170**, 43-59, doi:https://doi.org/10.1016/j.ultramic.2016.08.006 (2016).
172 Nicoletti, O. *et al.* Three-dimensional imaging of localized surface plasmon resonances of metal nanoparticles. *Nature* **502**, 80-84, doi:10.1038/nature12469 (2013).
173 Lichtert, S. & Verbeeck, J. Statistical consequences of applying a PCA noise filter on EELS spectrum images. *Ultramicroscopy* **125**, 35-42, doi:https://doi.org/10.1016/j.ultramic.2012.10.001 (2013).
174 LeBeau, J. M., Findlay, S. D., Allen, L. J. & Stemmer, S. Quantitative atomic resolution scanning transmission electron microscopy. *Physical Review Letters* **100**, 4, doi:10.1103/PhysRevLett.100.206101 (2008).
175 LeBeau, J. M., Findlay, S. D., Allen, L. J. & Stemmer, S. Standardless Atom Counting in Scanning Transmission Electron Microscopy. *Nano Lett.* **10**, 4405-4408, doi:10.1021/nl102025s (2010).
176 Katz-Boon, H. *et al.* Three-Dimensional Morphology and Crystallography of Gold Nanorods. *Nano Lett.* **11**, 273-278, doi:10.1021/nl103726k (2011).
177 Katz-Boon, H. *et al.* Stability of Crystal Facets in Gold Nanorods. *Nano Lett.* **15**, 1635-1641, doi:10.1021/acs.nanolett.5b00124 (2015).
178 Erni, R., Heinrich, H. & Kostorz, G. Quantitative characterisation of chemical inhomogeneities in (Al)under-bar-Ag using high-resolution Z-ontrast STEM. *Ultramicroscopy* **94**, 125-133, doi:10.1016/s0304-3991(02)00249-8 (2003).
179 Jones, L. & Iop. in *14th European Workshop on Modern Developments and Applications in Microbeam Analysis* Vol. 109 *IOP Conference Series-Materials Science and Engineering* (Iop Publishing Ltd, 2016).
180 Dwyer, C. *et al.* Sub-0.1 nm-resolution quantitative scanning transmission electron microscopy without adjustable parameters. *Appl. Phys. Lett.* **100**, 4, doi:10.1063/1.4711766 (2012).
181 Ishikawa, R., Lupini, A. R., Findlay, S. D. & Pennycook, S. J. Quantitative Annular Dark Field Electron Microscopy Using Single Electron Signals. *Microscopy and Microanalysis* **20**, 99-110, doi:10.1017/s1431927613013664 (2014).
182 Sang, X. H. & LeBeau, J. M. Characterizing the response of a scintillator-based detector to single electrons. *Ultramicroscopy* **161**, 3-9, doi:10.1016/j.ultramic.2015.11.008 (2016).
183 LeBeau, J. M. & Stemmer, S. Experimental quantification of annular dark-field images in scanning transmission electron microscopy. *Ultramicroscopy* **108**, 1653-1658, doi:10.1016/j.ultramic.2008.07.001 (2008).
184 LeBeau, J. M., Findlay, S. D., Allen, L. J. & Stemmer, S. Position averaged convergent beam electron diffraction: Theory and applications. *Ultramicroscopy* **110**, 118-125, doi:10.1016/j.ultramic.2009.10.001 (2010).
185 Verbeeck, J., Beche, A. & Van den Broek, W. A holographic method to measure the source size broadening in STEM. *Ultramicroscopy* **120**, 35-40, doi:10.1016/j.ultramic.2012.05.007 (2012).
186 Maunders, C., Dwyer, C., Tiemeijer, P. C. & Etheridge, J. Practical methods for the measurement of spatial coherence—A comparative study. *Ultramicroscopy* **111**, 1437-1446, doi:https://doi.org/10.1016/j.ultramic.2011.05.011 (2011).

187 Haider, M., Uhlemann, S. & Zach, J. Upper limits for the residual aberrations of a high-resolution aberration-corrected STEM. *Ultramicroscopy* **81**, 163-175, doi:https://doi.org/10.1016/S0304-3991(99)00194-1 (2000).
188 Kuramochi, K. *et al.* Effect of chromatic aberration on atomic-resolved spherical aberration corrected STEM images. *Ultramicroscopy* **110**, 36-42, doi:https://doi.org/10.1016/j.ultramic.2009.09.003 (2009).
189 Zheng, C. L. & Etheridge, J. Measurement of chromatic aberration in STEM and SCEM by coherent convergent beam electron diffraction. *Ultramicroscopy* **125**, 49-58, doi:https://doi.org/10.1016/j.ultramic.2012.10.002 (2013).
190 Voyles, P. M., Grazul, J. L. & Muller, D. A. Imaging individual atoms inside crystals with ADF-STEM. *Ultramicroscopy* **96**, 251-273, doi:10.1016/s0304-3991(03)00092-5 (2003).
191 Sang, X. H., Oni, A. A. & LeBeau, J. M. Atom Column Indexing: Atomic Resolution Image Analysis Through a Matrix Representation. *Microscopy and Microanalysis* **20**, 1764-1771, doi:10.1017/s1431927614013506 (2014).
192 Yankovich, A. B. *et al.* Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts. *Nature Communications* **5**, doi:4155
10.1038/ncomms5155 (2014).
193 Voyles, P. M., Muller, D. A., Grazul, J. L., Citrin, P. H. & Gossmann, H. J. L. Atomic-scale imaging of individual dopant atoms and clusters in highly n-type bulk Si. *Nature* **416**, 826-829, doi:10.1038/416826a (2002).
194 Ishikawa, R., Lupini, A. R., Findlay, S. D., Taniguchi, T. & Pennycook, S. J. Three-Dimensional Location of a Single Dopant with Atomic Precision by Aberration-Corrected Scanning Transmission Electron Microscopy. *Nano Lett.* **14**, 1903-1908 (2014).
195 Mevenkamp, N. *et al.* Poisson noise removal from high-resolution STEM images based on periodic block matching. *Advanced Structural and Chemical Imaging* **1**, 19, doi:10.1186/s40679-015-0004-8 (2015).
196 Bals, S., Van Aert, S., Van Tendeloo, G. & Avila-Brande, D. Statistical estimation of atomic positions from exit wave reconstruction with a precision in the picometer range. *Physical Review Letters* **96**, 4, doi:10.1103/PhysRevLett.96.096106 (2006).
197 De Backer, A., van den Bos, K. H. W., Van den Broek, W., Sijbers, J. & Van Aert, S. StatSTEM: An efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images. *Ultramicroscopy* **171**, 104-116, doi:10.1016/j.ultramic.2016.08.018 (2016).
198 So, Y. G. & Kimoto, K. Effect of specimen misalignment on local structure analysis using annular dark-field imaging. *Journal of Electron Microscopy* **61**, 207-215, doi:10.1093/jmicro/dfs045 (2012).
199 Schmid, H., Okunishi, E., Oikawa, T. & Mader, W. Structural and elemental analysis of iron and indium doped zinc oxide by spectroscopic imaging in Cs-corrected STEM. *Micron* **43**, 49-56, doi:10.1016/j.micron.2011.05.011 (2012).
200 Fan, L., Zhang, F., Fan, H. & Zhang, C. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art* **2**, 7, doi:10.1186/s42492-019-0016-7 (2019).
201 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, doi:10.1021/acsnano.7b07504 (2017).
202 Ziatdinov, M., Nelson, C., Vasudevan, R. K., Chen, D. Y. & Kalinin, S. V. Building ferroelectric from the bottom up: The machine learning analysis of the atomic-scale ferroelectric distortions. *Appl. Phys. Lett.* **115**, 052902, doi:10.1063/1.5109520 (2019).
203 Lin, R., Zhang, R., Wang, C., Yang, X.-Q. & Xin, H. L. TEMImageNet and AtomSegNet Deep Learning Training Library and Models for High-Precision Atom Segmentation, Localization, Denoising, and Super-resolution Processing of Atom-Resolution Scanning TEM Images. *arXiv preprint arXiv:.09093* (2020).

204 Ziatdinov, M., Dyck, O., Jesse, S. & Kalinin, S. V. Atomic mechanisms for the Si atom dynamics in graphene: chemical transformations at the edge and in the bulk. *Advanced Functional Materials* **29**, 1904480 (2019).

205 Ziatdinov, M. *et al.* Building and exploring libraries of atomic defects in graphene: Scanning transmission electron and scanning tunneling microscopy study. *Sci. Adv.* **5**, eaaw8989 (2019).

206 Lee, C. H. *et al.* Deep Learning Enabled Strain Mapping of Single-Atom Defects in Two-Dimensional Transition Metal Dichalcogenides with Sub-Picometer Precision. *Nano Lett.* **20**, 3369-3377, doi:10.1021/acs.nanolett.0c00269 (2020).

207 Maksov, A. *et al.* Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS2. *npj Comput. Mater.* **5**, 8, doi:10.1038/s41524-019-0152-9 (2019).

208 Hinton, G. E. & Salakhutdinov, R. R. Reducing the Dimensionality of Data with Neural Networks. *Science* **313**, 504, doi:10.1126/science.1127647 (2006).

209 Roccapriore, K. M., Ziatdinov, M., Cho, S. H., Hachtel, J. A. & Kalinin, S. V. Predictability of Localized Plasmonic Responses in Nanoparticle Assemblies. *Small* **n/a**, 2100181, doi:https://doi.org/10.1002/smll.202100181 (2021).

210 Kingma, D. P. & Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

211 Kingma, D. P. & Welling, M. An Introduction to Variational Autoencoders. *Foundations and Trends® in Machine Learning* **12**, 307-392, doi:10.1561/2200000056 (2019).

212 Blei, D. M., Kucukelbir, A. & McAuliffe, J. D. Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association* **112**, 859-877, doi:10.1080/01621459.2017.1285773 (2017).

213 Muller, D. A. & Grazul, J. Optimizing the environment for sub-0.2 nm scanning transmission electron microscopy. *Journal of Electron Microscopy* **50**, 219-226 (2001).

214 Vonharrach, H. S. INSTRUMENTAL FACTORS IN HIGH-RESOLUTION FEG STEM. *Ultramicroscopy* **58**, 1-5, doi:10.1016/0304-3991(94)00172-j (1995).

215 Couillard, M., Radtke, G. & Botton, G. A. Strain fields around dislocation arrays in a sigma 9 silicon bicrystal measured by scanning transmission electron microscopy. *Philos. Mag.* **93**, 1250-1267, doi:10.1080/14786435.2013.778428 (2013).

216 Isakozawa, S., Tomonaga, S., Hashimoto, T. & Baba, N. High-precision image-drift-correction method for EM images with a low signal-to-noise ratio. *Microscopy* **63**, 301-312, doi:10.1093/jmicro/dfu016 (2014).

217 Sang, X. H. & LeBeau, J. M. Revolving scanning transmission electron microscopy: Correcting sample drift distortion without prior knowledge. *Ultramicroscopy* **138**, 28-35, doi:10.1016/j.ultramic.2013.12.004 (2014).

218 Dycus, J. H. *et al.* Accurate Nanoscale Crystallography in Real-Space Using Scanning Transmission Electron Microscopy. *Microscopy and Microanalysis* **21**, 946-952, doi:10.1017/s1431927615013732 (2015).

219 Jones, L. *et al.* Smart Align-a new tool for robust non-rigid registration of scanning microscope data. *Advanced Structural and Chemical Imaging* **1**, 16, doi:10.1186/s40679-015-0008-4 (2015).

220 Ophus, C., Ciston, J. & Nelson, C. T. Correcting nonlinear drift distortion of scanning probe and scanning transmission electron microscopies from image pairs with orthogonal scan directions. *Ultramicroscopy* **162**, 1-9, doi:10.1016/j.ultramic.2015.12.002 (2016).

221 Berkels, B. & Liebscher, C. H. Joint non-rigid image registration and reconstruction for quantitative atomic resolution scanning transmission electron microscopy. *Ultramicroscopy* **198**, 49-57, doi:10.1016/j.ultramic.2018.12.016 (2019).

222 Yankovich, A. B. *et al.* Non-rigid registration and non-local principle component analysis to improve electron microscopy spectrum images. *Nanotechnology* **27**, 14, doi:10.1088/0957-4484/27/36/364001 (2016).

223 Wang, Y. *et al.* Correcting the linear and nonlinear distortions for atomically resolved STEM spectrum and diffraction imaging. *Microscopy* **67**, 9, doi:10.1093/jmicro/dfy002 (2018).
224 Barcena-Gonzalez, G. *et al.* CDrift: An Algorithm to Correct Linear Drift From A Single High-Resolution STEM Image. *Microscopy and Microanalysis* **26**, 913-920, doi:10.1017/s1431927620001774 (2020).
225 Sanchez, A. M. *et al.* An approach to the systematic distortion correction in aberration-corrected HAADF images. *Journal of Microscopy* **221**, 1-7, doi:10.1111/j.1365-2818.2006.01533.x (2006).
226 Braidy, N., Le Bouar, Y., Lazar, S. & Ricolleau, C. Correcting scanning instabilities from images of periodic structures. *Ultramicroscopy* **118**, 67-76, doi:10.1016/j.ultramic.2012.04.001 (2012).
227 Ning, S. C. *et al.* Scanning distortion correction in STEM images. *Ultramicroscopy* **184**, 274-283, doi:10.1016/j.ultramic.2017.09.003 (2018).
228 Roccapriore, K. M., Ziatdinov, M. & Kalinin, S. V. Identification and Correction of Temporal and Spatial Distortions in Scanning Transmission Electron Microscopy. *arXiv preprint arXiv:.05635* (2020).
229 Sang, X. H. *et al.* Dynamic scan control in STEM: spiral scans. *Advanced Structural and Chemical Imaging* **2**, doi:6
10.1186/s40679-016-0020-3 (2016).
230 Li, X., Dyck, O., Kalinin, S. V. & Jesse, S. Compressed Sensing of Scanning Transmission Electron Microscopy (STEM) With Nonrectangular Scans. *Microscopy and Microanalysis* **24**, 623-633, doi:10.1017/s143192761801543x (2018).
231 Stevens, A. *et al.* A sub-sampled approach to extremely low-dose STEM. *Appl. Phys. Lett.* **112**, 5, doi:10.1063/1.5016192 (2018).
232 Zobelli, A. *et al.* Spatial and spectral dynamics in STEM hyperspectral imaging using random scan patterns. *Ultramicroscopy* **212**, 112912 (2020).
233 Beche, A., Goris, B., Freitag, B. & Verbeeck, J. Development of a fast electromagnetic beam blanker for compressed sensing in scanning transmission electron microscopy. *Appl. Phys. Lett.* **108**, 5, doi:10.1063/1.4943086 (2016).
234 Velazco, A., Nord, M., Beche, A. & Verbeeck, J. Evaluation of different rectangular scan strategies for STEM imaging. *Ultramicroscopy* **215**, 10, doi:10.1016/j.ultramic.2020.113021 (2020).
235 Dai, G. L. *et al.* Reference nano-dimensional metrology by scanning transmission electron microscopy. *Meas. Sci. Technol.* **24**, 9, doi:10.1088/0957-0233/24/8/085001 (2013).
236 Radaelli, P. G. & Cheong, S. W. Structural phenomena associated with the spin-state transition in LaCoO(3). *Physical Review B* **66**, 9, doi:10.1103/PhysRevB.66.094408 (2002).
237 Zhou, J. S. & Goodenough, J. B. Chemical bonding and electronic structure of RNiO3 (R=rare earth). *Physical Review B* **69**, 4, doi:10.1103/PhysRevB.69.153105 (2004).
238 Zhou, J. S. & Goodenough, J. B. Universal octahedral-site distortion in orthorhombic perovskite oxides. *Physical Review Letters* **94**, 4, doi:10.1103/PhysRevLett.94.065501 (2005).
239 Taraci, J. L. *et al.* Strain mapping in nanowires. *Nanotechnology* **16**, 2365-2371, doi:10.1088/0957-4484/16/10/062 (2005).
240 Hytch, M. J., Putaux, J. L. & Thibault, J. Stress and strain around grain-boundary dislocations measured by high-resolution electron microscopy. *Philos. Mag.* **86**, 4641-4656, doi:10.1080/14786430600743876 (2006).
241 Chu, M. W., Szafraniak, I., Hesse, D., Alexe, M. & Gosele, U. Elastic coupling between 90 degrees twin walls and interfacial dislocations in epitaxial ferroelectric perovskites: A quantitative high-resolution transmission electron microscopy study. *Physical Review B* **72**, doi:174112
10.1103/PhydRevB.72.174112 (2005).

242 Zhao, C. W., Xing, Y. M., Zhou, C. E. & Bai, P. C. Experimental examination of displacement and strain fields in an edge dislocation core. *Acta Mater.* **56**, 2570-2575, doi:10.1016/j.actamat.2008.01.045 (2008).
243 Pennycook, S. J. & Nellist, P. D. *Scanning transmission electron microscopy: imaging and analysis*. (Springer Science & Business Media, 2011).
244 Sousa, A. A. & Leapman, R. D. Development and application of STEM for the biological sciences. *Ultramicroscopy* **123**, 38-49, doi:https://doi.org/10.1016/j.ultramic.2012.04.005 (2012).
245 A.K. Tagantsev, L. E. C., and J. Fousek. *Domains in Ferroic Crystals and Thin Films*. (Springer, 2010).
246 Salje, E. K. H. Ferroelasticity. *Contemp. Phys.* **41**, 79-91, doi:10.1080/001075100181196 (2000).
247 Pan, X. Q., Kaplan, W. D., Ruhle, M. & Newnham, R. E. Quantitative comparison of transmission electron microscopy techniques for the study of localized ordering on a nanoscale. *J. Am. Ceram. Soc.* **81**, 597-605 (1998).
248 Jia, C. L. *et al.* Unit-cell scale mapping of ferroelectricity and tetragonality in epitaxial ultrathin ferroelectric films. *Nature Materials* **6**, 64-69, doi:10.1038/nmat1808 (2007).
249 Jia, C. L. *et al.* Atomic-scale study of electric dipoles near charged and uncharged domain walls in ferroelectric films. *Nature Materials* **7**, 57-61, doi:10.1038/nmat2080 (2008).
250 Chisholm, M. F., Luo, W. D., Oxley, M. P., Pantelides, S. T. & Lee, H. N. Atomic-Scale Compensation Phenomena at Polar Interfaces. *Physical Review Letters* **105**, doi:197602 10.1103/PhysRevLett.105.197602 (2010).
251 Jia, C. L. *et al.* Oxygen octahedron reconstruction in the SrTiO(3)/LaAlO(3) heterointerfaces investigated using aberration-corrected ultrahigh-resolution transmission electron microscopy. *Physical Review B* **79**, doi:081405 10.1103/PhysRevB.79.081405 (2009).
252 Borisevich, A. *et al.* Mapping Octahedral Tilts and Polarization Across a Domain Wall in BiFeO(3) from Z-Contrast Scanning Transmission Electron Microscopy Image Atomic Column Shape Analysis. *ACS Nano* **4**, 6071-6079, doi:10.1021/nn1011539 (2010).
253 Kim, Y. M. *et al.* Interplay of Octahedral Tilts and Polar Order in BiFeO3 Films. *Adv. Mater.* **25**, 2497-2504, doi:10.1002/adma.201204584 (2013).
254 He, Q. *et al.* Towards 3D Mapping of BO6 Octahedron Rotations at Perovskite Heterointerfaces, Unit Cell by Unit Cell. *ACS Nano* **9**, 8412-8419, doi:10.1021/acsnano.5b03232 (2015).
255 Kim, Y. M. *et al.* Direct observation of ferroelectric field effect and vacancy-controlled screening at the BiFeO3/LaxSr1-xMnO3 interface. *Nature Materials* **13**, 1019-1025, doi:10.1038/nmat4058 (2014).
256 Borisevich, A. Y. *et al.* Interface dipole between two metallic oxides caused by localized oxygen vacancies. *Physical Review B* **86**, doi:140102 10.1103/PhysRevB.86.140102 (2012).
257 Borisevich, A. Y. *et al.* Exploring Mesoscopic Physics of Vacancy-Ordered Systems through Atomic Scale Observations of Topological Defects. *Physical Review Letters* **109**, doi:065702 10.1103/PhysRevLett.109.065702 (2012).
258 Li, Q. *et al.* Quantification of flexoelectricity in PbTiO3/SrTiO3 superlattice polar vortices using machine learning and phase-field modeling. *Nature Communications* **8**, doi:1468 10.1038/s41467-017-01733-8 (2017).
259 Nelson, C. T. *et al.* Exploring physics of ferroelectric domain walls via Bayesian analysis of atomically resolved STEM data. *Nature Communications* **11**, 12, doi:10.1038/s41467-020-19907-2 (2020).
260 Rossell, M. D. *et al.* Atomic Structure of Highly Strained BiFeO3 Thin Films. *Physical Review Letters* **108**, 5, doi:10.1103/PhysRevLett.108.047601 (2012).

261 Campanini, M., Erni, R., Yang, C. H., Ramesh, R. & Rossell, M. D. Periodic Giant Polarization Gradients in Doped BiFeO3 Thin Films. *Nano Lett.* **18**, 717-724, doi:10.1021/acs.nanolett.7b03817 (2018).

262 Campanini, M. *et al.* Atomic-resolution differential phase contrast STEM on ferroelectric materials: A mean-field approach. *Physical Review B* **101**, 12, doi:10.1103/PhysRevB.101.184116 (2020).

263 Midgley, P. A. & Dunin-Borkowski, R. E. Electron tomography and holography in materials science. *Nature Materials* **8**, 271-280, doi:10.1038/nmat2406 (2009).

264 Welborn, M., Tang, W. J., Ryu, J., Petkov, V. & Henkelman, G. A combined density functional and x-ray diffraction study of Pt nanoparticle structure. *J. Chem. Phys.* **135**, 5, doi:10.1063/1.3607967 (2011).

265 Meredig, B. & Wolverton, C. A hybrid computational-experimental approach for automated crystal structure solution. *Nature Materials* **12**, 123-127, doi:10.1038/nmat3490 (2013).

266 Alkauskas, A., McCluskey, M. D. & Van de Walle, C. G. Tutorial: Defects in semiconductors-Combining experiment and theory. *Journal of Applied Physics* **119**, 11, doi:10.1063/1.4948245 (2016).

267 Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547-555, doi:10.1038/s41586-018-0337-2 (2018).

268 Kolb, B., Lentz, L. C. & Kolpak, A. M. Discovering charge density functionals and structure-property relationships with PROPhet: A general framework for coupling machine learning and first-principles methods. *Scientific Reports* **7**, 9, doi:10.1038/s41598-017-01251-z (2017).

269 Schuett, O. & VandeVondele, J. Machine Learning Adaptive Basis Sets for Efficient Large Scale Density Functional Theory Simulation. *J. Chem. Theory Comput.* **14**, 4168-4175, doi:10.1021/acs.jctc.8b00378 (2018).

270 Sen, F. G. *et al.* in *2015 Ieee 42nd Photovoltaic Specialist Conference IEEE Photovoltaic Specialists Conference* (Ieee, 2015).

271 Schwenker, E. *et al.* An Autonomous Microscopy Workflow for Structure Determination from Atomic-Resolution Images. *Microscopy and Microanalysis* **24**, 510-511, doi:10.1017/S1431927618003045 (2018).

272 Wei, J. *et al.* Direct imaging of atomistic grain boundary migration. *Nature Materials*, doi:10.1038/s41563-020-00879-z (2021).

273 Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**, 600-612, doi:10.1109/tip.2003.819861 (2004).

274 Liu, T.-J., Lin, Y.-C., Lin, W. & Kuo, C. C. J. Visual quality assessment: recent developments, coding applications and future trends. *APSIPA Transactions on Signal and Information Processing* **2**, e4, doi:10.1017/ATSIP.2013.5 (2013).

275 Lin, J. Y., Liu, T., Wu, E. C. & Kuo, C. J. in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific.* 1-5.

276 Guo, J. L. *et al.* Effect of selenium and chlorine co-passivation in polycrystalline CdSeTe devices. *Appl. Phys. Lett.* **115**, 5, doi:10.1063/1.5123169 (2019).

277 Krivanek, O. L. *et al.* Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy. *Nature* **464**, 571-574, doi:10.1038/nature08879 (2010).

278 Treacy, M. M. J. Z Dependence of Electron Scattering by Single Atoms into Annular Dark-Field Detectors. *Microscopy and Microanalysis* **17**, 847-858, doi:10.1017/s1431927611012074 (2011).

279 Yamashita, S. *et al.* Atomic number dependence of Z contrast in scanning transmission electron microscopy. *Scientific Reports* **8**, 7, doi:10.1038/s41598-018-30941-5 (2018).

280 Li, H. *et al.* Atomic Structure and Dynamics of Single Platinum Atom Interactions with Monolayer MoS2. *ACS Nano* **11**, 3392-3403, doi:10.1021/acsnano.7b00796 (2017).

281 Wang, S. *et al.* Preferential Pt Nanocluster Seeding at Grain Boundary Dislocations in Polycrystalline Monolayer MoS2. *ACS Nano* **12**, 5626-5636, doi:10.1021/acsnano.8b01418 (2018).
282 Zan, R., Bangert, U., Ramasse, Q. & Novoselov, K. S. Metal–Graphene Interaction Studied via Atomic Resolution Scanning Transmission Electron Microscopy. *Nano Lett.* **11**, 1087-1092, doi:10.1021/nl103980h (2011).
283 Wang, S. *et al.* In Situ Atomic-Scale Studies of the Formation of Epitaxial Pt Nanocrystals on Monolayer Molybdenum Disulfide. *ACS Nano* **11**, 9057-9067, doi:10.1021/acsnano.7b03648 (2017).
284 Li, X. *et al.* Precursor Design for High Density Single Pt Atom Sites on MoS2: Enhanced Stability at Elevated Temperatures and Reduced 3D Clustering. *Chemistry of Materials* **32**, 2541-2551, doi:10.1021/acs.chemmater.9b05288 (2020).
285 Tai, K.-L. *et al.* Phase Variations and Layer Epitaxy of 2D PdSe2 Grown on 2D Monolayers by Direct Selenization of Molecular Pd Precursors. *ACS Nano* **14**, 11677-11690, doi:10.1021/acsnano.0c04230 (2020).
286 Sinha, S. *et al.* Atomic structure and defect dynamics of monolayer lead iodide nanodisks with epitaxial alignment on graphene. *Nature Communications* **11**, 823, doi:10.1038/s41467-020-14481-z (2020).
287 Poh, S. M. *et al.* Molecular Beam Epitaxy of Highly Crystalline MoSe2 on Hexagonal Boron Nitride. *ACS Nano* **12**, 7562-7570, doi:10.1021/acsnano.8b04037 (2018).
288 Gallagher, J. C. *et al.* Epitaxial growth of iridate pyrochlore Nd2Ir2O7 films. *Scientific Reports* **6**, 22282, doi:10.1038/srep22282 (2016).
289 Gerkman, M. A. *et al.* Direct Imaging of Individual Molecular Binding to Clean Nanopore Edges in 2D Monolayer MoS2. *ACS Nano* **14**, 153-165, doi:10.1021/acsnano.9b06061 (2020).
290 Cai, C., Han, S., Wang, Q. & Gu, M. Direct Observation of Yolk–Shell Transforming to Gold Single Atoms and Clusters with Superior Oxygen Evolution Reaction Efficiency. *ACS Nano* **13**, 8865-8871, doi:10.1021/acsnano.9b02135 (2019).
291 Sun, Y. *et al.* Subunit cell–level measurement of polarization in an individual polar vortex. *Sci. Adv.* **5**, eaav4355, doi:10.1126/sciadv.aav4355 (2019).
292 Ohyama, J. *et al.* Atomic-scale insight into the structural effect of a supported Au catalyst based on a size-distribution analysis using Cs-STEM and morphological image-processing. *Journal of Catalysis* **335**, 24-35, doi:https://doi.org/10.1016/j.jcat.2015.11.021 (2016).
293 Melzer, D. *et al.* Atomic-Scale Determination of Active Facets on the MoVTeNb Oxide M1 Phase and Their Intrinsic Catalytic Activity for Ethane Oxidative Dehydrogenation. *Angewandte Chemie International Edition* **55**, 8873-8877, doi:https://doi.org/10.1002/anie.201600463 (2016).
294 Qiao, B. *et al.* Single-atom catalysis of CO oxidation using Pt1/FeOx. *Nature Chemistry* **3**, 634-641, doi:10.1038/nchem.1095 (2011).
295 Liu, J. *et al.* In situ tracing of atom migration in Pt/NiPt hollow spheres during catalysis of CO oxidation. *Chemical Communications* **50**, 1804-1807, doi:10.1039/C3CC47772J (2014).
296 Gerkman, M. A., Sinha, S., Warner, J. H. & Han, G. G. D. Direct Imaging of Photoswitching Molecular Conformations Using Individual Metal Atom Markers. *ACS Nano* **13**, 87-96, doi:10.1021/acsnano.8b08441 (2019).
297 Lee, J. K. *et al.* Metal Atom Markers for Imaging Epitaxial Molecular Self-Assembly on Graphene by Scanning Transmission Electron Microscopy. *ACS Nano* **13**, 7252-7260, doi:10.1021/acsnano.9b02906 (2019).
298 Kengmana, E. S., Lee, J. K., Li, X., Warner, J. H. & Han, G. G. D. Self-Assembly of Bowlic Supramolecules on Graphene Imaged at the Individual Molecular Level using Heavy Atom Tagging. *Small* **16**, 2002860, doi:https://doi.org/10.1002/smll.202002860 (2020).

299 Henninen, T. R., Bon, M., Wang, F., Passerone, D. & Erni, R. The Structure of Sub-nm Platinum Clusters at Elevated Temperatures. *Angew. Chem.-Int. Edit.* **59**, 839-845, doi:10.1002/anie.201911068 (2020).
300 Furnival, T., Leary, R. K. & Midgley, P. A. Denoising time-resolved microscopy image sequences with singular value thresholding. *Ultramicroscopy* **178**, 112-124, doi:10.1016/j.ultramic.2016.05.005 (2017).
301 Wang, F., Henninen, T. R., Keller, D. & Erni, R. Noise2Atom: unsupervised denoising for scanning transmission electron microscopy images. *Applied Microscopy* **50**, 23, doi:10.1186/s42649-020-00041-8 (2020).
302 Mirman, R. *Point groups, space groups, crystals, molecules*. (World Scientific, 1999).
303 Glinchuk, M. D. & Stephanovich, V. A. Dynamic properties of relaxor ferroelectrics. *Journal of Applied Physics* **85**, 1722-1726, doi:10.1063/1.369316 (1999).
304 Vugmeister, B. E. Polarization dynamics and formation of polar nanoregions in relaxor ferroelectrics. *Physical Review B* **73**, doi:174117
10.1103/PhysRevB.73.174117 (2006).
305 Takenaka, H., Grinberg, I., Liu, S. & Rappe, A. M. Slush-like polar structures in single-crystal relaxors. *Nature* **546**, 391-+, doi:10.1038/nature22068 (2017).
306 Binder, K. & Reger, J. D. THEORY OF ORIENTATIONAL GLASSES MODELS, CONCEPTS, SIMULATIONS. *Adv. Phys.* **41**, 547-627, doi:10.1080/00018739200101553 (1992).
307 Binder, K. & Young, A. P. SPIN-GLASSES - EXPERIMENTAL FACTS, THEORETICAL CONCEPTS, AND OPEN QUESTIONS. *Rev. Mod. Phys.* **58**, 801-976, doi:10.1103/RevModPhys.58.801 (1986).
308 Dagotto, E. Complexity in strongly correlated electronic systems. *Science* **309**, 257-262, doi:10.1126/science.1107559 (2005).
309 Dagotto, E., Hotta, T. & Moreo, A. Colossal magnetoresistant materials: The key role of phase separation. *Phys. Rep.-Rev. Sec. Phys. Lett.* **344**, 1-153, doi:10.1016/s0370-1573(00)00121-6 (2001).
310 Blinc, R. *et al.* Local polarization distribution and Edwards-Anderson order parameter of relaxor ferroelectrics. *Physical Review Letters* **83**, 424-427, doi:10.1103/PhysRevLett.83.424 (1999).
311 Cross, L. E. RELAXOR FERROELECTRICS. *Ferroelectrics* **76**, 241-267, doi:10.1080/00150198708016945 (1987).
312 Cliffe, M. J. *et al.* Structural simplicity as a restraint on the structure of amorphous silicon. *Physical Review B* **95**, doi:224108
10.1103/PhysRevB.95.224108 (2017).
313 Keen, D. A. & Goodwin, A. L. The crystallography of correlated disorder. *Nature* **521**, 303-309, doi:10.1038/nature14453 (2015).
314 Cheetham, A. K., Bennett, T. D., Coudert, F. X. & Goodwin, A. L. Defects and disorder in metal organic frameworks. *Dalton Trans.* **45**, 4113-4126, doi:10.1039/c5dt04392a (2016).
315 Dobigeon, N. & Brun, N. Spectral mixture analysis of EELS spectrum-images. *Ultramicroscopy* **120**, 25-34, doi:10.1016/j.ultramic.2012.05.006 (2012).
316 Pennycook, S. J., Varela, M., Lupini, A. R., Oxley, M. P. & Chisholm, M. F. Atomic-resolution spectroscopic imaging: past, present and future. *Journal of Electron Microscopy* **58**, 87-97, doi:10.1093/jmicro/dfn030 (2009).
317 Jiang, Y. *et al.* Electron ptychography of 2D materials to deep sub-angstrom resolution. *Nature* **559**, 343-+, doi:10.1038/s41586-018-0298-5 (2018).
318 Ziatdinov, M., Nelson, C., Vasudevan, R. K., Chen, D. Y. & Kalinin, S. V. Building ferroelectric from the bottom up: The machine learning analysis of the atomic-scale ferroelectric distortions. *Applied Physics Letters* **115**, 5, doi:10.1063/1.5109520 (2019).
319 Belianinov, A. *et al.* Identification of phases, symmetries and defects through local crystallography. *Nature Communications* **6**, doi:7801

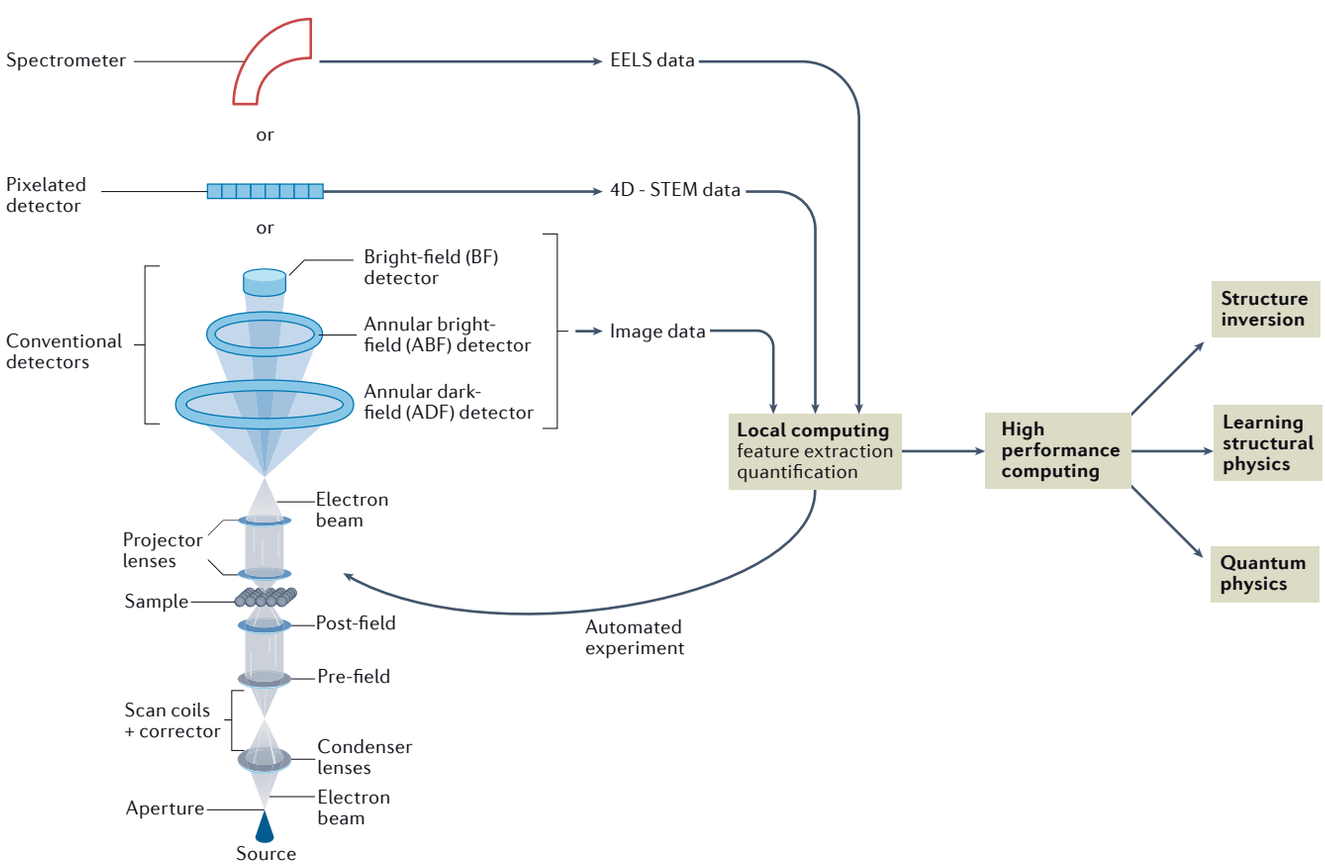
10.1038/ncomms8801 (2015).
320 Ziatdinov, M. *et al.* Atomic-scale observation of structural and electronic orders in the layered compound alpha-RuCl3. *Nature Communications* **7**, doi:13774
10.1038/ncomms13774 (2016).
321 He, Q., Woo, J., Belianinov, A., Guliants, V. V. & Borisevich, A. Y. Better Catalysts through Microscopy: Mesoscale M1/M2 Intergrowth in Molybdenum-Vanadium Based Complex Oxide Catalysts for Propane Ammoxidation. *ACS Nano* **9**, 3470-3478, doi:10.1021/acsnano.5b00271 (2015).
322 Wu, Z. *et al.* A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* **32**, 4-24, doi:10.1109/TNNLS.2020.2978386 (2021).
323 Kalinin, S. V., Dyck, O., Ghosh, A., Sumpter, B. G. & Ziatdinov, M. Unsupervised Machine Learning Discovery of Chemical Transformation Pathways from Atomically-Resolved Imaging Data. *arXiv preprint arXiv:2010.09196* (2020).
324 Wicks, S. *et al.* Collective dynamics in nanostructured polycrystalline ferroelectric thin films using local time-resolved measurements and switching spectroscopy. *Acta Mater.* **58**, 67-75, doi:10.1016/j.actamat.2009.08.057 (2010).
325 Vlcek, L., Maksov, A., Pan, M., Vasudevan, R. K. & Kalinin, S. V. Knowledge Extraction from Atomically Resolved Images. *ACS Nano* **11**, 10313-10320, doi:10.1021/acsnano.7b05036 (2017).
326 Landau, L. D. & Lifshitz, E. M. *Statistical Physics*. (Elsevier Science, 2013).
327 Kitaev, A. Anyons in an exactly solved model and beyond. *Annals of Physics* **321**, 2-111, doi:https://doi.org/10.1016/j.aop.2005.10.005 (2006).
328 Vlcek, L., Yang, S., Ziatdinov, M., Kalinin, S. & Vasudevan, R. Statistical Physics-based Framework and Bayesian Inference for Model Selection and Uncertainty Quantification. *Microscopy and Microanalysis* **25**, 130-131, doi:10.1017/S1431927619001387 (2019).
329 Mani Prudhvi Valleti, S., Vlcek, L., Vasudevan, R. K. & Kalinin, S. V. Inversion of lattice models from the observations of microscopic degrees of freedom: parameter estimation with uncertainty quantification. *arXiv e-prints* (2019). <https://ui.adsabs.harvard.edu/abs/2019arXiv190909244M>.
330 Chakraborty, M. *et al.* Reconstruction of the interatomic forces from dynamic scanning transmission electron microscopy data. *Journal of Applied Physics* **127**, 7, doi:10.1063/5.0009413 (2020).
331 Nobakht, A. Y. *et al.* Reconstruction of effective potential from statistical analysis of dynamic trajectories. *AIP Adv.* **10**, 6, doi:10.1063/5.0006103 (2020).
332 Schepman, A. M. H., Vandervoort, J. A. P., Kramer, J. & Mellema, J. E. GETTING STARTED WITH A SCANNING-TRANSMISSION ELECTRON-MICROSCOPE COUPLED TO A SMALL COMPUTER. *Ultramicroscopy* **3**, 265-269, doi:10.1016/s0304-3991(78)80035-7 (1978).
333 Strahm, M. & Butler, J. H. FAST DIGITAL DATA ACQUISITION AND ONLINE PROCESSING SYSTEM FOR AN HB5 SCANNING-TRANSMISSION ELECTRON-MICROSCOPE. *Rev. Sci. Instrum.* **52**, 840-848, doi:10.1063/1.1136679 (1981).
334 Andree, P. J., Mellema, J. E. & Ruigrok, R. W. H. DISCRIMINATION OF HEAVY AND LIGHT-ELEMENTS IN A SPECIMEN BY USE OF SCANNING-TRANSMISSION ELECTRON-MICROSCOPY. *Ultramicroscopy* **17**, 237-241, doi:10.1016/0304-3991(85)90090-7 (1985).
335 Haider, M., Epstein, A., Jarron, P. & Boulin, C. A VERSATILE, SOFTWARE CONFIGURABLE MULTICHANNEL STEM DETECTOR FOR ANGLE-RESOLVED IMAGING. *Ultramicroscopy* **54**, 41-59, doi:10.1016/0304-3991(94)90091-4 (1994).
336 Yang, Y. S. *et al.* Deciphering chemical order/disorder and material properties at the single-atom level. *Nature* **542**, 75-+, doi:10.1038/nature21042 (2017).
337 Kalinin S.V., S. B. G., Archibald R.K. Big-deep-smart data in imaging for guiding materials design. *Nature Materials* **14**, 973 (2015).

338 Spurgeon, S. R. *et al.* Towards data-driven next-generation transmission electron microscopy. *Nature Materials*, 6, doi:10.1038/s41563-020-00833-z.
339 Vasilevsky, N. A., Minnier, J., Haendel, M. A. & Champieux, R. E. Reproducible and reusable research: are journal data sharing policies meeting the mark? *PeerJ* **5**, 18, doi:10.7717/peerj.3208 (2017).
340 Molloy, J. C. The Open Knowledge Foundation: Open Data Means Better Science. *PLoS. Biol.* **9**, 4, doi:10.1371/journal.pbio.1001195 (2011).
341 Woelfle, M., Olliaro, P. & Todd, M. H. Open science is a research accelerator. *Nature Chemistry* **3**, 745-748, doi:10.1038/nchem.1149 (2011).
342 Hanisch, R. J., Gilmore, I. S. & Plant, A. L. Improving Reproducibility in Research: The Role of Measurement Science. *J. Res. Natl. Inst. Stand. Technol.* **124**, 13, doi:10.6028/jres.124.024 (2019).
343 Blaiszik, B. *et al.* The Materials Data Facility: Data Services to Advance Materials Science Research. *Jom* **68**, 2045-2052, doi:10.1007/s11837-016-2001-3 (2016).
344 Jain, A. *et al.* Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, doi:011002
10.1063/1.4812323 (2013).
345 Iudin, A., Korir, P. K., Salavert-Torres, J., Kleywegt, G. J. & Patwardhan, A. EMPIAR: a public archive for raw electron microscopy image data. *Nat. Methods* **13**, 387-388, doi:10.1038/nmeth.3806 (2016).
346 Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **10**, 980-980, doi:10.1038/nsb1203-980 (2003).
347 Guo, S. *et al.* Direct Mapping of Ion Diffusion Times on LiCoO(2) Surfaces with Nanometer Resolution. *Journal of the Electrochemical Society* **158**, A982-A990, doi:10.1149/1.3604759 (2011).
348 Nikiforov, M. P. *et al.* Temperature-dependent phase transitions in zeptoliter volumes of a complex biological membrane. *Nanotechnology* **22**, doi:055709
10.1088/0957-4484/22/5/055709 (2011).
349 Kalinin, S. V. *et al.* Big, Deep, and Smart Data in Scanning Probe Microscopy. *ACS Nano* **10**, 9068-9086, doi:10.1021/acsnano.6b04212 (2016).
350 Ziatdinov, M., Nelson, C., Vasudevan, R., Chen, D. & Kalinin, S. Building Ferroelectric from the Bottom Up: The Machine Learning Analysis of the Atomic-Scale Ferroelectric Distortions. *ChemRxiv.* *https://doi.org/10.26434/chemrxiv.8001473.v1*, doi:10.26434/chemrxiv.8001473.v1 (2019).
351 Komsa, H. P. *et al.* Two-Dimensional Transition Metal Dichalcogenides under Electron Irradiation: Defect Production and Doping. *Physical Review Letters* **109**, doi:035503
10.1103/PhysRevLett.109.035503 (2012).
352 Mishra, R., Ishikawa, R., Lupini, A. R. & Pennycook, S. J. Single-atom dynamics in scanning transmission electron microscopy. *MRS Bull.* **42**, 644-652, doi:10.1557/mrs.2017.187 (2017).
353 Toma Susi, J. C. M., Jani Kotakoski. Manipulating low-dimensional materials down to the level of single atoms with electron irradiation. *Ultramicroscopy* **180**, 163–172 (2017).
354 Dyck, O. *et al.* Building Structures Atom by Atom via Electron Beam Manipulation. *Small* **14**, doi:1801771
10.1002/smll.201801771 (2018).
355 Dyck, O., Jesse, S. & Kalinin, S. V. A self-driving microscope and the Atomic Forge. *MRS Bull.* **44**, 669-670, doi:10.1557/mrs.2019.211 (2019).
356 Kelley, K. P. *et al.* Fast Scanning Probe Microscopy via Machine Learning: Non-Rectangular Scans with Compressed Sensing and Gaussian Process Optimization. *Small* **16**, 6, doi:10.1002/smll.202002878 (2020).
357 Stevens, A. *et al.* Applying compressive sensing to TEM video: a substantial frame rate increase on any camera. *Advanced Structural and Chemical Imaging* **1**, doi:10

10.1186/s40679-015-0009-3 (2015).
358 Kovarik, L., Stevens, A., Liyu, A. & Browning, N. D. Implementing an accurate and rapid sparse sampling approach for low-dose atomic resolution STEM imaging. *Appl. Phys. Lett.* **109**, doi:164102
10.1063/1.4965720 (2016).
359 Jesse, S. *et al.* Direct atomic fabrication and dopant positioning in Si using electron beams with active real-time image-based feedback. *Nanotechnology* **29**, doi:255303
10.1088/1361-6528/aabb79 (2018).
360 Gulrajani, I. & Lopez-Paz, D. In search of lost domain generalization. *arXiv preprint arXiv:2007.01434* (2020).
361 Kalinin, S. V. *et al.* Automated and Autonomous Experiments in Electron and Scanning Probe Microscopy. *ACS Nano* **15**, 12604-12627, doi:10.1021/acsnano.1c02104 (2021).
362 Wilson, A. G., Hu, Z., Salakhutdinov, R. & Xing, E. P. in *Artificial intelligence and statistics.* 370-378 (PMLR).
363 Zheng, H. M. *et al.* Observation of Transient Structural-Transformation Dynamics in a Cu2S Nanorod. *Science* **333**, 206-209, doi:10.1126/science.1204713 (2011).
364 Klie, R. F., Ito, Y., Stemmer, S. & Browning, N. S. Observation of oxygen vacancy ordering and segregation in Perovskite oxides. *Ultramicroscopy* **86**, 289-302, doi:10.1016/s0304-3991(00)00120-0 (2001).
365 Jang, J. H. *et al.* In Situ Observation of Oxygen Vacancy Dynamics and Ordering in the Epitaxial LaCoO3 System. *ACS Nano* **11**, 6942-6949, doi:10.1021/acsnano.7b02188 (2017).
366 Hoehl, D. *et al.* TEMPERATURE-DEPENDENCE OF ELECTRON-BEAM INDUCED EPITAXIAL CRYSTALLIZATION OF SILICON. *Phys. Status Solidi A-Appl. Res.* **122**, K35-K37, doi:10.1002/pssa.2211220149 (1990).
367 Jencic, I., Bench, M. W., Robertson, I. M. & Kirk, M. A. ELECTRON-BEAM-INDUCED CRYSTALLIZATION OF ISOLATED AMORPHOUS REGIONS IN SI, GE, GAP, AND GAAS. *Journal of Applied Physics* **78**, 974-982 (1995).
368 Remeika, M. & Bezryadin, A. Sub-10 nanometre fabrication: molecular templating, electron-beam sculpting and crystallization of metallic nanowires. *Nanotechnology* **16**, 1172-1176, doi:10.1088/0957-4484/16/8/032 (2005).
369 Bae, I. T., Zhang, Y. W., Weber, W. J., Higuchi, M. & Giannuzzi, L. A. Electron-beam induced recrystallization in amorphous apatite. *Appl. Phys. Lett.* **90**, doi:021912
10.1063/1.2430779 (2007).
370 Zhang, Y. *et al.* Ion-induced damage accumulation and electron-beam-enhanced recrystallization in SrTiO(3). *Physical Review B* **72**, doi:094112
10.1103/PhysRevB.72.094112 (2005).
371 Huang, P. Y. *et al.* Imaging Atomic Rearrangements in Two-Dimensional Silica Glass: Watching Silica's Dance. *Science* **342**, 224-227, doi:10.1126/science.1242248 (2013).
372 Kurasch, S. *et al.* Atom-by-Atom Observation of Grain Boundary Migration in Graphene. *Nano Lett.* **12**, 3168-3173, doi:10.1021/nl301141g (2012).
373 Yang, Z. Q. *et al.* Direct Observation of Atomic Dynamics and Silicon Doping at a Topological Defect in Graphene. *Angew. Chem.-Int. Edit.* **53**, 8908-8912, doi:10.1002/anie.201403382 (2014).
374 Krasheninnikov, A. V. & Banhart, F. Engineering of nanostructured carbon materials with electron or ion beams. *Nature Materials* **6**, 723-733, doi:10.1038/nmat1996 (2007).
375 Lingerfelt, D. B. *et al.* Nonadiabatic Effects on Defect Diffusion in Silicon-Doped Nanographenes. *Nano Lett.* **21**, 236-242, doi:10.1021/acs.nanolett.0c03587 (2021).
376 Susi, T., Meyer, J. C. & Kotakoski, J. Quantifying transmission electron microscopy irradiation effects using two-dimensional materials. *Nature Reviews Physics* **1**, 397-405, doi:10.1038/s42254-019-0058-y (2019).

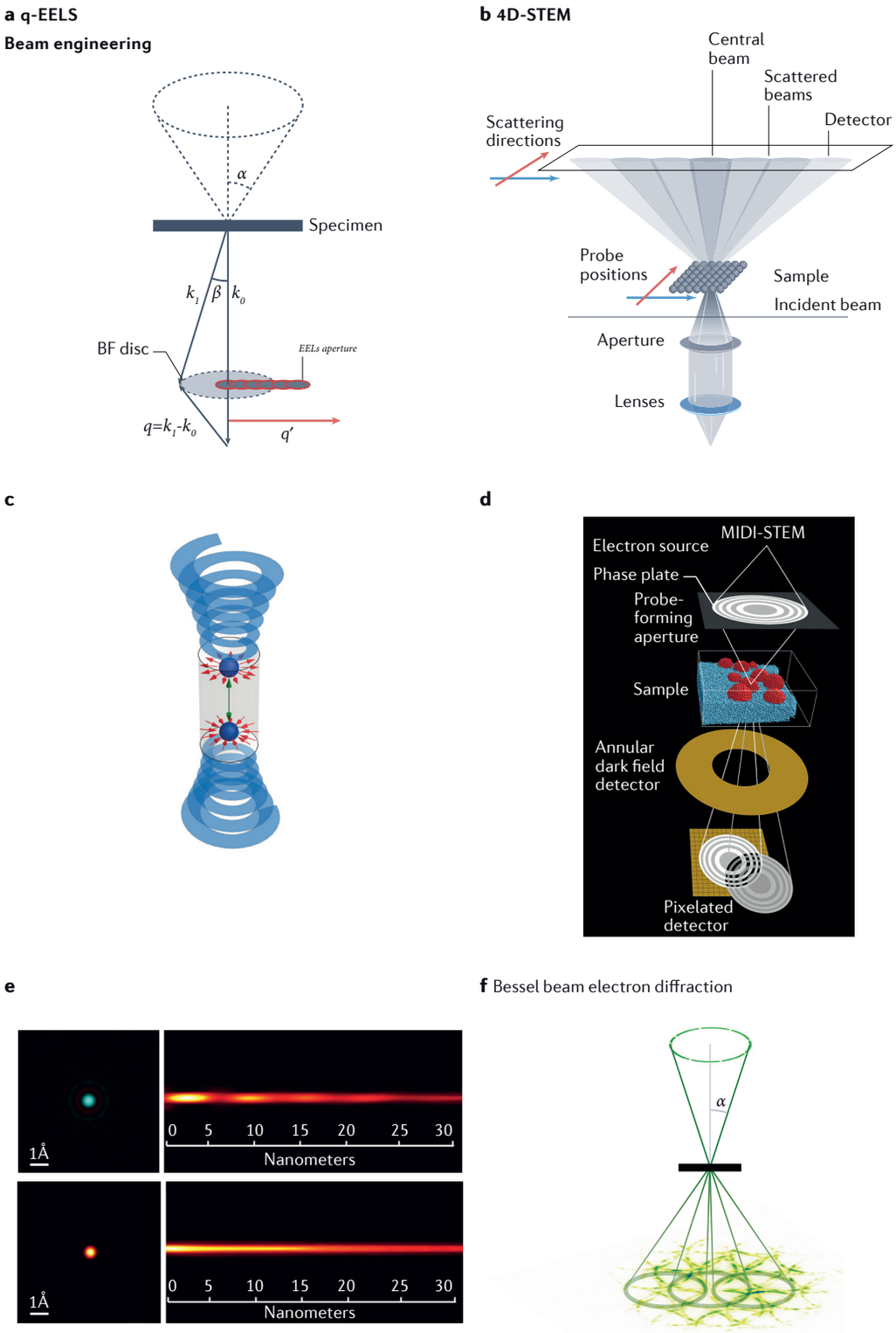
377 Krasheninnikov, A. V. & Nordlund, K. Ion and electron irradiation-induced effects in nanostructured materials. *Journal of Applied Physics* **107**, doi:071301 10.1063/1.3318261 (2010).
378 Egerton, R. F. Radiation damage to organic and inorganic specimens in the TEM. *Micron* **119**, 72-87, doi:10.1016/j.micron.2019.01.005 (2019).
379 Jesse, S. *et al.* Atomic-Level Sculpting of Crystalline Oxides: Toward Bulk Nanofabrication with Single Atomic Plane Precision. *Small* **11**, 5895-5900 (2015).
380 Mousavi, S. S., Schukat, M. & Howley, E. in *Proceedings of SAI Intelligent Systems Conference.* 426-440 (Springer).
381 Schrittwieser, J. *et al.* Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature* **588**, 604-+, doi:10.1038/s41586-020-03051-4 (2020).
382 Wagner, T. & Raunser, S. The evolution of SPHIRE-crYOLO particle picking and its application in automated cryo-EM processing workflows. *Communications biology* **3**, 1-5 (2020).
383 Wang, S. J. & Summers, R. M. Machine learning and radiology. *Med. Image Anal.* **16**, 933-951, doi:10.1016/j.media.2012.02.005 (2012).

Fig 1


Spectrometer
EELS data
or
Pixelated detector
4D - STEM data
or
Conventional detectors
Bright-field (BF) detector
Annular bright-field (ABF) detector
Annular dark-field (ADF) detector
Image data
Electron beam
Projector lenses
Sample
Post-field
Pre-field
Scan coils + corrector
Condenser lenses
Aperture
Electron beam
Source
Automated experiment
Local computing
feature extraction
quantification
High performance computing
Structure inversion
Learning structural physics
Quantum physics

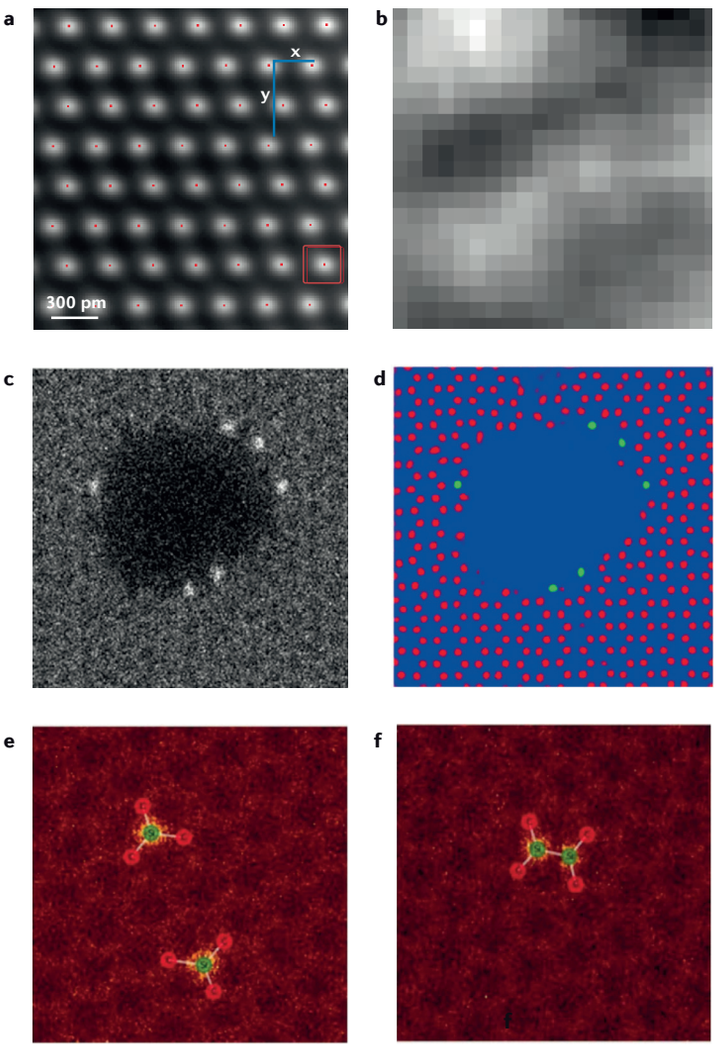
Manuscript number NRMP-20-063_STEM_Kalinin Primer 07|01|22

Fig 2



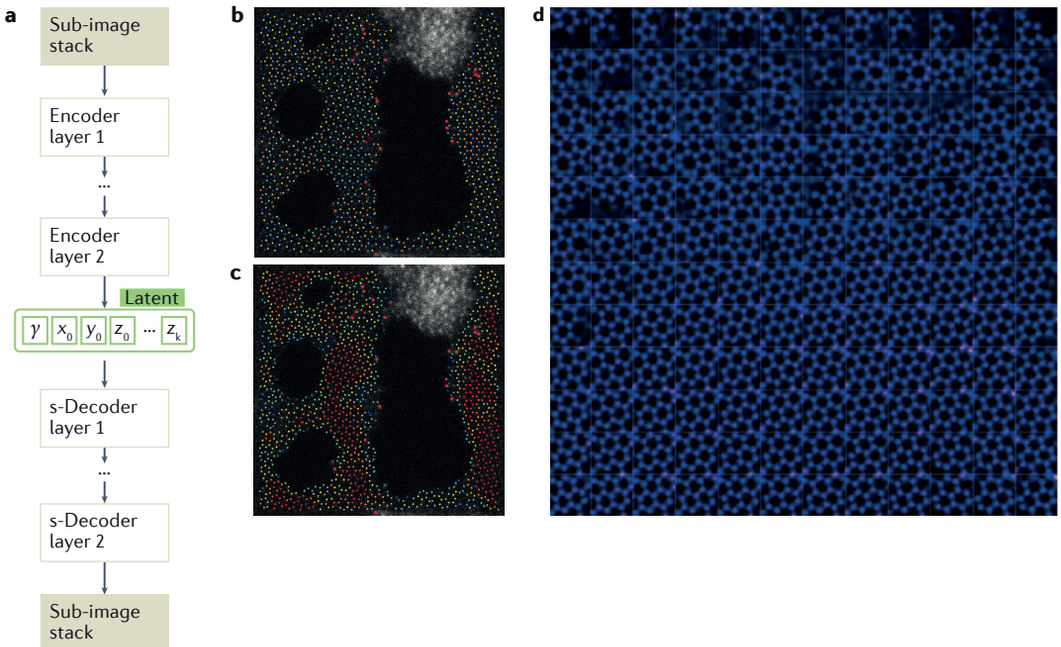
Manuscript number NRMP-20-063_STEM_Kalinin Primer 07|01|22

Fig 3



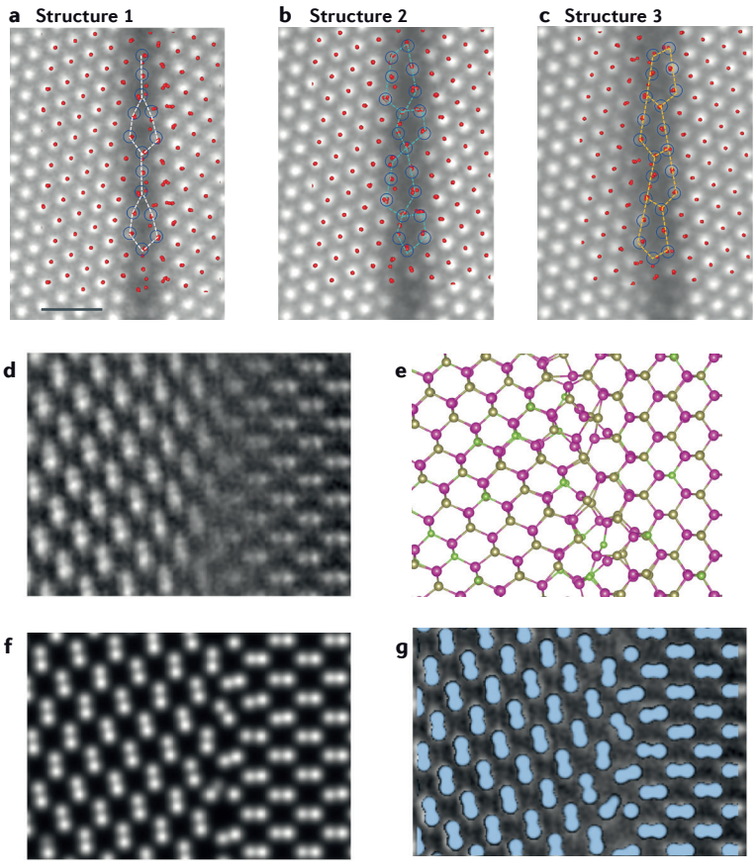
Manuscript number NRMP-20-063_STEM_Kalinin Primer 07|01|22

Fig 4



Manuscript number NRMP-20-063_STEM_Kalinin Primer 07|01|22

Fig 5



Manuscript number NRMP-20-063_STEM_Kalinin Primer 07|01|22

Fig 6

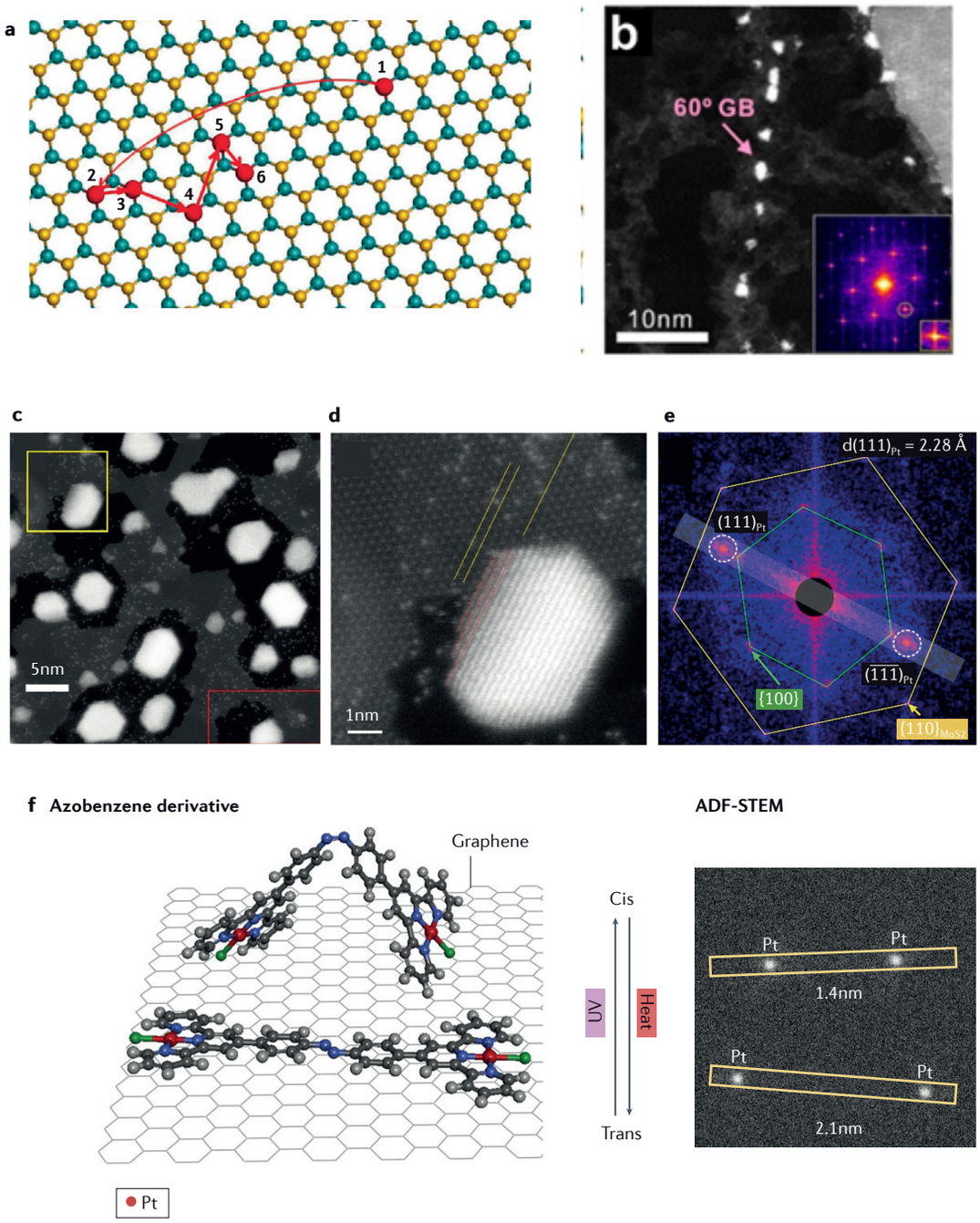
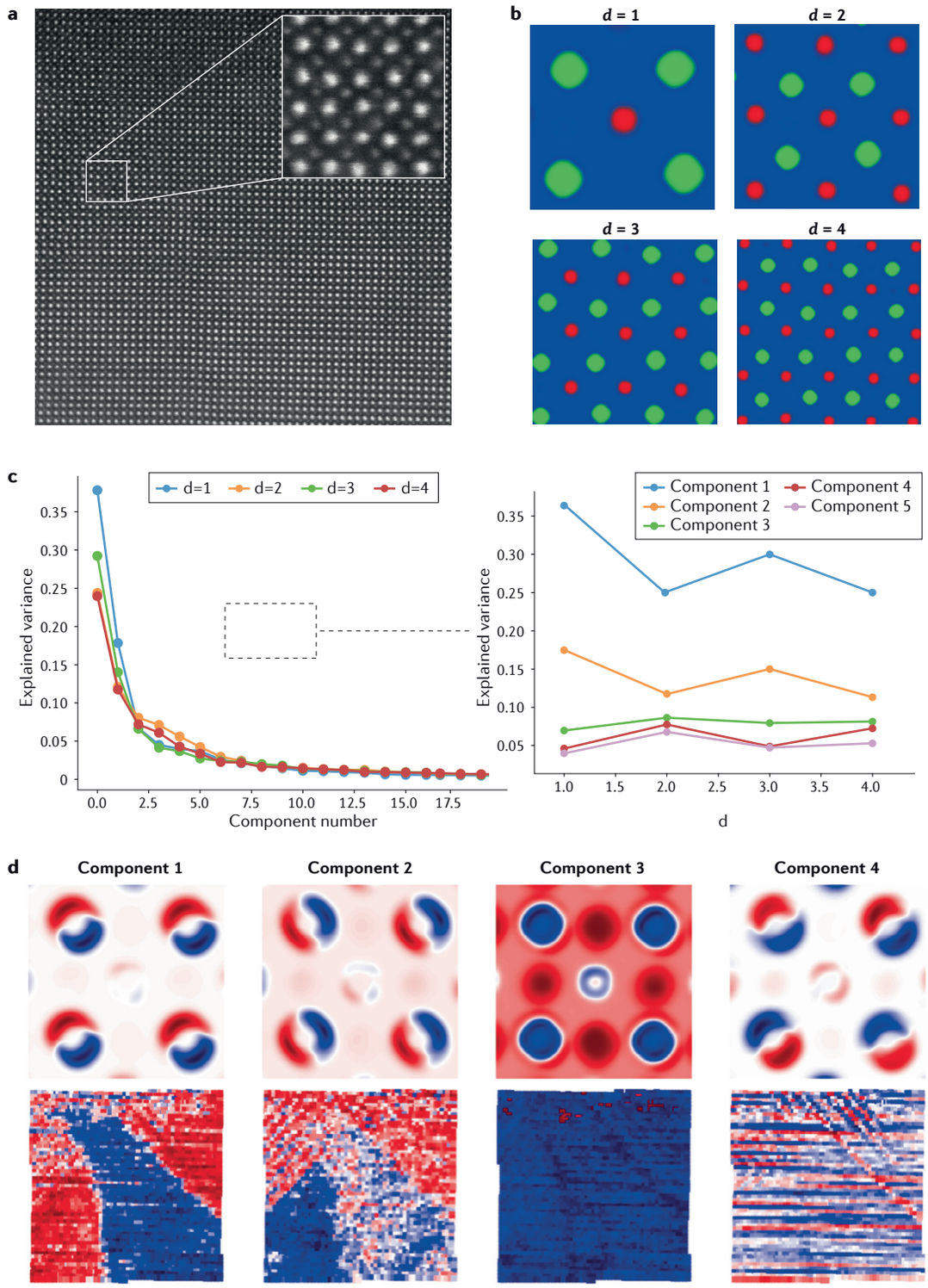


Fig 7


a
b
d = 1
d = 2
d = 3
d = 4
c
d=1
d=2
d=3
d=4
Explained variance
Component number
Component 1
Component 2
Component 3
Component 4
Component 5
Explained variance
d
d
Component 1
Component 2
Component 3
Component 4