



CHALMERS
UNIVERSITY OF TECHNOLOGY

Simulation study of the performance of neural network-enhanced PACBED for characterizing atomic-scale deformations in 2D van der Waals materials

Downloaded from: <https://research.chalmers.se>, 2025-11-09 06:46 UTC

Citation for the original published paper (version of record):

Yankovich, A., Röding, M., Olsson, E. (2026). Simulation study of the performance of neural network-enhanced PACBED for characterizing atomic-scale deformations in 2D van der Waals materials. *Ultramicroscopy*, 279. <http://dx.doi.org/10.1016/j.ultramic.2025.114246>

N.B. When citing this work, cite the original published paper.



Simulation study of the performance of neural network-enhanced PACBED for characterizing atomic-scale deformations in 2D van der Waals materials

Andrew B. Yankovich ^{a, , *, 1}, Magnus Röding ^{b, c, d, 1}, Eva Olsson ^a

^a Department of Physics, Chalmers University of Technology, 41296 Göteborg, Sweden

^b RISE Research Institutes of Sweden, Bioeconomy and Health, Agriculture and Food, 41276 Göteborg, Sweden

^c Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 41296 Göteborg, Sweden

^d Sustainable Innovation & Transformational Excellence, Pharmaceutical Technology & Development, Operations, AstraZeneca Gothenburg, 43183 Mölndal, Sweden

ARTICLE INFO

Keywords:

2D materials
Van der Waals materials
Strain
Position averaged convergent beam electron diffraction
Convolutional neural networks
Machine learning

ABSTRACT

Two dimensional (2D) van der Waals (vdW) materials have attractive mechanical, electronic, optical, and catalytic properties that are highly tunable especially when they are thin. However, they are rarely perfect and flat, and their properties are strongly influenced by local crystal lattice deformations that include the 2D strain tensor, in-plane rotation and corrugation, where the latter is manifested as local sample tilt. Therefore, to gain more control over their properties, a detailed understanding of these deformations is needed. Position averaged convergent beam electron diffraction (PACBED) is a powerful technique for providing information about local atomic structure. In this work, we perform a comprehensive simulation study of the performance of PACBED in combination with convolutional neural networks (CNNs) for prediction of deformations of 2D materials. We generate around 100,000 simulated PACBED patterns from 2H MoS₂ for thicknesses from 1 to 20 atomic layers where strain, rotation, and tilt parameters are varied. Five convergence angles are explored which vary from conventional nano beam electron diffraction (6.35 mrad) to atomic resolution conditions (32.94 mrad). From this simulated PACBED library, we train regression CNNs to simultaneously predict the 2D strain tensor, in-plane rotation, and tilt of the sample. For different convergence angles and thicknesses, we study the prediction performance for each of the deformation parameters. We find that there is a trade-off between better prediction performance (small convergence angles) and probe size (large convergence angles). For smaller convergence angles like those used for conventional NBED conditions, the strain prediction error can be as low as 0.0003 %, while for larger convergence angles like those used for atomic resolution probes, the strain error increases to 0.001 - 0.003 %. The impressive prediction performance even for large convergence angles suggests that PACBED combined with CNNs is a feasible method for predicting deformation parameters using atomic resolution electron probes. Further, we conclude that the prediction can be difficult for monolayers, and suggest two remedies: excluding tilt from the predictions and performing nonlinear intensity rescaling of the training data. This work contributes to the optimal design of PACBED experiments for characterization of local crystal deformations and, therefore, to an improved understanding of how 2D vdW materials respond to imperfections.

1. Introduction

Understanding how crystalline materials deform at the nanoscale is crucial for realizing the full potential of many material systems and devices. Because many material properties are a direct consequence of its precise atomic structure, local atomic lattice deformations have a direct effect on local properties and device performance. Indeed, deformation can have profound effects on for example mechanical, chemical, electronic, optical, and catalytic properties. Deformation can

be induced by applying external forces, such as mechanical, thermal or electrical, or it can arise from discontinuities in an otherwise perfect atomic crystal structure, such as from surfaces, interfaces, point defects, defect clusters, dislocations, grain boundaries, and composition variation. These types of deformations typically create complex and non-intuitive strain gradients that can spatially vary on the atomic length scale. Therefore, measuring the complete strain state, ideally with atomic spatial resolution, is fundamental for understanding and tailoring a wide range of material properties.

* Corresponding author.

E-mail address: andrew.yankovich@chalmers.se (A.B. Yankovich).

¹ Equal contribution.

Since the pioneering work on graphene [1], two dimensional (2D) van der Waals (vdW) materials have attracted significant attention, partially due to their vast range of properties. 2D vdW materials feature a planar layered crystalline structure where the layers are held together by weak vdW interactions and the atoms within the layers are held together by stronger intralayer (covalent or ionic) bonds. This layered structure provides the ability to isolate vdW materials with constant atomic layer thickness over large areas, even down to single atomic monolayers, as well as the ability to synthesize designer heterostructures by mix-and-match stacking of different component layers to create exotic devices with tailored properties [2–5]. Transition metal dichalcogenide (TMD) materials, which were first discovered more than a century ago by Linus Pauling et al. [6], have emerged as a particularly interesting group of vdW materials. TMD materials are composed of transition metal atoms (e.g. Mo, W, Re) and chalcogenide atoms (e.g. S, Se, Te) [7]. Their exceptional mechanical, electronic, excitonic and optical properties have attracted their use in various nanotechnology applications [8–13]. TMDs, and other vdW materials, are often imagined as being perfectly flat. In contrast, micro- and nanoscopic corrugation are often an intrinsic feature of isolated layers and can directly affect their properties. For example, the discovery of intrinsic corrugation in suspended graphene was critical for understanding its unique electronic transport properties [14–16]. Because TMDs have a thin flake morphology, crystal deformation frequently causes flake corrugation, and flake corrugation frequently causes crystal deformation. Improving our ability to measure and spatially map this deformation phenomenon is critical for obtaining a more complete understanding of the atomic structure and properties of vdW materials.

Scanning/ transmission electron microscopy (S/TEM) is an important characterization tool that provides a plethora of materials information with impressive spatial resolution down to 50 pm [17–22]. Numerous S/TEM methods have been developed to map local crystal deformations and these can be categorized as either image-based or diffraction-based techniques, each with its strengths and weaknesses [23,24].

Image-based techniques require resolving the crystalline lattice in an atomic resolution TEM or STEM image, followed by quantification of the strain, using for example Geometric Phase Analysis (GPA) [25,26], Dark Field Electron Holography (DFEH) [27] or direct peak analysis [28–32]. TEM has been used to directly image strain within crystalline materials [33–37]. However, due to the phase-contrast nature, TEM images are typically not directly interpretable without using more advanced methods [38,39] and special care is required to reliably measure strain [26,40–45]. On the contrary, STEM images are directly interpretable with mass-thickness Z-contrast [46], and sub-picometer precision in locating atom columns is possible [31] which enables strain mapping in nanoparticles with < 1 % precision [32,47]. Image-based methods are powerful, but hold some limitations. GPA and peak analysis methods require that each atomic column position in the image is clearly resolved to precisely locate its position. This introduces a limitation to the real space field of view (FOV) (typically < 50–100 nm) of the resulting strain maps because the images have a limited number of pixels. GPA and DFEH methods are limited to > 1 nm spatial resolution and cannot provide atomic-site-specific strain information like that attainable with direct peak analysis. Furthermore, S/TEM images and the resulting strain maps can be highly sensitive to the electron beam focus and aberration conditions, so these parameters need to be carefully controlled and remain stable during image acquisition.

Diffraction-based techniques rely on momentum-resolved measurements of forward scattered electrons for which there are two convergent beam electron diffraction (CBED) approaches. The first approach measures strain by analyzing position shifts of the higher order Laue zone (HOLZ) lines in CBED patterns [48,49]. Although this approach can achieve nm-scale spatial resolution, it does not work well for all sample thicknesses, particularly thin samples that produce weak

HOLZ information. The second approach, nano beam electron diffraction (NBED), [50–54] uses a near-parallel nanoscale electron beam to record CBED patterns that have non-overlapping Bragg scattering discs. By measuring the displacements of the Bragg discs, the local strain of the sample illuminated by the electron beam can be measured. Furthermore, by recording NBED diffraction patterns over a 2D grid of probe positions on the sample, a four dimensional (4D) STEM data set [55] can be recorded and analyzed to produce strain maps.

Conventional NBED is a powerful strain mapping tool that has advantages and limitations. The FOV of the NBED strain maps is typically limited by the maximum allowed 4D STEM data set size. Current size limitations easily allow for NBED strain mapping over multiple micrometers FOV [54,56], surpassing the limitations of imaging-based methods. The strain precision of NBED can be as low as 0.1% and is typically limited by the ability to accurately measure the position of the Bragg discs. Various methods have been explored to improve disc position measurements, and most are either limited by the number of pixels in the detector or the non-uniform disc intensity caused by dynamical scattering effects [57–59]. These limitations have motivated the measurement of sub-pixel shifts of the diffraction discs and the use of structured probe-forming apertures [56,60–64]. The spatial resolution of NBED is limited by the electron probe size, which is inversely proportional to the convergence angle that defines the Bragg diffraction disc size. In order to reliably determine the positions of the Bragg discs for strain analysis, the discs typically need to be non-overlapping because overlap makes it difficult for conventional methods to determine disc positions. Therefore, the non-overlapping disc criterion limits the maximum convergence angle and the minimum probe size. Conventional NBED uses highly separated Bragg discs and < 1 mrad convergence angles to make disc position measurements easier, resulting in spatial resolutions of 5–10 nm [24]. This can be improved to around 1 nm by increasing the convergence angle so that the Bragg discs are slightly non-overlapping [54], but this can reduce strain precision because the dynamical scattering effects becomes more apparent.

In order to improve the spatial resolution of 4D STEM beyond conventional NBED and into the atomic-scale, larger convergence angles are required that result in substantial Bragg diffraction disc overlap. Additionally, as the probe size is reduced to smaller than the size of a unit cell, the CBED pattern is highly sensitive to the position of the probe within the unit cell. One way to simplify these atomic-scale CBED patterns is to incoherently average them over many probe positions, for example over one unit cell, to produce a position averaged CBED (PACBED) pattern [65]. PACBED patterns are not strongly affected by lens aberrations or focus, but are highly sensitive to sample thickness [65,66], tilt [65], polarity [67], sub-lattice tilts [68], crystal symmetry [69] and composition [70]. Because of the complexity of measuring the positions of individual Bragg discs within highly overlapped CBED or PACBED patterns using conventional methods, strain measurements from these patterns have not been reported.

Recently, there has been a growing interest in applying machine learning (ML) methods to enhance the capabilities of S/TEM [71,72]. Various methods have focused on analyzing PACBED data using convolution neural networks (CNNs) to measure sample thickness, tilt, and crystal phase. [73–75]. ML has also been used to analyze NBED data containing non-overlapping discs to study deformation [76–78]. Shi et al. [77] utilized unsupervised ML to hierarchically cluster NBED data to reveal classes of sample deformation, such as strain, rotation, and corrugation, without the need for a priori structure information. Although this method is good at identifying classes of deformations that are present in the 4D STEM data, it does not directly quantify these deformations. Yuan et al. [76] developed separate CNNs for a known material structure to independently predict the x sample tilt direction, y sample tilt direction, x position of a single Bragg diffraction disc and y position of a single Bragg diffraction disc. This has the implication of requiring multiple trained CNNs and multiple predictions

for different Bragg discs to subsequently calculate tilt and strain information. Their CNN models are trained on simulated diffraction data that are not simultaneously tilted and strained, presumably limiting the ability to disentangle the effects from both tilt and strain diffraction contributions which are common in many real experiments. Munshi et al. [78] developed a neural network model that predicts the structure factor of a material from an NBED pattern. This method does not require a priori knowledge about the material structure because the model was trained on a simulated NBED library from different crystal structures, orientations, thicknesses, and microscope parameters. This model also does not directly quantify the deformations, but requires subsequent use of conventional methods to measure deformations from the resulting structure factors.

In this study, we perform a comprehensive simulation study to explore the viability and performance of individual CNNs that are trained to directly and simultaneously predict various deformation parameters from single PACBED patterns of a 2D material. This method does not require the need to explicitly measure disc positions to access strain information (like with conventional NBED) but instead is trained to directly output the full 2D strain tensor simultaneously with the in-plane rotation and tilt. Our CNNs are trained on a large library of simulated PACBED patterns that are calculated using combinations of the strain tensor, rotation, tilt, thickness and electron probe convergence angle in order to disentangle the effects from all. Simulated PACBED libraries are necessary for this work in order to have a large enough labeled training library that precisely and systematically varies all the relevant parameters (e.g. strain tensor, tilt, thickness). Labeled experimental data is not a viable option. This work specifically investigates the CNN prediction performance as a function of electron probe convergence angle and thickness, a systematic investigation of which is lacking. In particular, this work explores the possibility of determining deformations using atomic-scale electron probes that create significantly overlapping Bragg discs. This method assumes that the 2D material composition and crystal phase are known prior information, but these assumptions could be alleviated with future work aimed at making the CNNs more generalizable. Additionally, these CNNs are trained on deformed single crystalline models without the introduction of any crystalline defects. Future work could explore training CNNs to predict deformations even in the presence of defects.

We find that CNNs can effectively predict strain, rotation, and tilt across a wide range of convergence angles and sample thicknesses, even for larger convergence angles that are suitable for atomic resolution STEM imaging. However, there is a trade-off, and to gain the best performance in deformation prediction, smaller convergence angles are required. The method is robust to the sample thicknesses explored here, except when the sample is ultra thin, i.e. monolayers. For experiments specifically aimed at measuring deformations in monolayers, we suggest excluding tilt which we demonstrate improves strain prediction performance. These results will help guide the optimization of diffraction experiments aimed at measuring 2D material deformations by unveiling the expected CNN performance for different experiment and sample conditions. Furthermore, this method, in combination with modern 4D STEM experiments, enables future investigations of fast and robust deformation mapping of 2D materials over large FOVs using atomic resolution probes.

2. Methods

We begin by introducing the material structure which has to be specified for simulating PACBED, and then derive the expressions governing the deformations and the strain tensor. After that, we provide details of how the PACBED simulations are performed, and illustrate the impact of convergence angle, thickness, and deformations on the PACBED patterns. Further, we describe how the training data is produced. Finally, we provide the specification of the CNN and how it is trained.

2.1. Material structure

In this work, we study the trigonal prismatic 2H crystal phase of molybdenum disulfide (2H MoS₂), which is a prototypical TMD material. We use the smallest possible orthogonal unit cell of 2H MoS₂, instead of the nonorthogonal primitive unit cell. The orthogonal unit cell has dimensions $L_x = 3.1613 \text{ \AA}$, $L_y = 5.4755 \text{ \AA}$, $L_z = 12.3063 \text{ \AA}$. The x , y and z directions are aligned with the zigzag $[1\bar{1}00]$, armchair $[11\bar{2}0]$ and $[0001]$ crystallographic directions, respectively. The unit cell contains four molybdenum atoms and eight sulfur atoms that are organized into two monolayers with the AB stacking sequence in the z direction. Hence, further unit cell repetition in the z direction maintains the AB stacking sequence of the 2H crystal phase.

2.2. Deformation model

We consider three unique contributions to the deformations: an in-plane x - y oriented 2D strain tensor (ϵ), a rotation of the lattice about the z direction (γ) (i.e. a rotation in the x - y plane), and a tilt of the sample relative to the direction of the electron beam (ϕ_x and ϕ_y) (i.e. relative to the z direction).

To model strain deformations, we consider strain only in the x - y plane because there are much fewer interactions between the layers in the out-of-plane z direction due to the weak vdW bonding and larger interatomic distances. To derive the 2D strain tensor ϵ , our point of departure is to define the principal stretches, λ_1 and λ_2 , which are scale factors that contract ($\lambda_i < 1$) or expand ($\lambda_i > 1$) the structure along the principal directions. The principal directions are defined by the unit vectors $e_1 = (\cos \theta, \sin \theta)^T$ and $e_2 = (-\sin \theta, \cos \theta)^T$ and provide a coordinate system that is rotated by an angle θ counterclockwise relative to the x and y directions. By convention, $\lambda_1 \geq \lambda_2$, and they constitute the largest and smallest stretches in any direction. Given a counterclockwise rotation matrix R ,

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \quad (1)$$

we calculate the position of a deformed point (x', y') from an original point (x, y) by applying (1) a clockwise rotation R^T , (2) principal stretches λ_1 and λ_2 , and (3) a counterclockwise rotation R so that

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = R \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} R^T \begin{pmatrix} x \\ y \end{pmatrix}. \quad (2)$$

To arrive at the usual zero-centered strain measures, we define the principal strains $\epsilon_1 = \lambda_1 - 1$ and $\epsilon_2 = \lambda_2 - 1$ (that are zero for no deformation). From this, we arrive at the symmetric 2D strain tensor

$$\epsilon = \begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} \\ \epsilon_{xy} & \epsilon_{yy} \end{pmatrix} = R^T \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix} R, \quad (3)$$

where ϵ_{xx} and ϵ_{yy} are the normal strains in the x and y directions, respectively, and ϵ_{xy} is the shear strain.

To model the rotation deformations of the lattice about the z direction, we directly rotate the simulated PACBED patterns. Given our model for strain deformations, the strain tensor is always symmetric and therefore does not account for lattice rotation. Further, a rotation of the lattice by an angle γ corresponds to a rotation of the resulting PACBED pattern by the same angle. Therefore, rotation is easiest and least computationally intensive to model by directly applying a rotation with an angle γ to the simulated PACBED patterns.

To model corrugation, we include a tilt between the sample and the electron beam so that the electron beam incidence is not perfectly normal to the sample surface. This is accomplished using a built-in function in the PACBED simulation software (see below) which provides control over this parameter in two directions, ϕ_x and ϕ_y . Note that the tilts are defined relative to the x - y coordinate system of the undeformed unit cell regardless of strain and rotation.

In this study we set ranges for the modeled deformations to represent expected maximum deviations for typical freestanding TMD flakes.

To simulate strain, the two principal stretches are selected randomly (uniformly) in the range $[0.95, 1.05]$ (ascertaining that $\lambda_1 \geq \lambda_2$). The angle θ is selected randomly in the range $[0, \pi]$ radians. The elements of the strain tensor, ϵ_{xx} , ϵ_{yy} , and ϵ_{xy} are computed using Eq. (3), and all three are consequently distributed non-uniformly in the range $[-0.05, 0.05]$. Note that we provide strain values in % from here on (hence in the range $[-5, 5]$ %). The rotation angle γ is selected randomly in the range $[-10, 10]^\circ$. The sample tilt angles ϕ_x and ϕ_y are selected randomly in the range $[-5, 5]$ mrad. These parameter ranges could be easily extended if needed, but this would incur a larger computational cost in order to sample the deformation parameter space sufficiently densely.

2.3. PACBED simulations

PACBED patterns are simulated using the frozen phonon multislice algorithm implemented in the software Prismatic (version 2.0) [79–81], which accurately captures the dynamical scattering effects in CBED patterns.

A large pre-strained supercell of atomic coordinates are used as the input model for Prismatic. For given values of λ_1 , λ_2 , and θ , a strained unit cell is generated by applying the deformation in Eq. (2) to all the x and y atom coordinates of the original orthogonal unit cell. We then periodically replicate the deformed unit cell many times in the x - y plane, and 10 times in the z direction (corresponding to 20 layers). The resulting model is cropped in the x and y directions to generate a supercell of size $400 \text{ \AA} \times 400 \text{ \AA} \times 123 \text{ \AA}$, which is used as the input model for Prismatic.

The selection of probe positions in this case requires special consideration for two reasons. First, the strained supercell will not satisfy periodic boundary conditions in the x and y directions. This implies that any probe positions sufficiently near the supercell edge will produce distorted PACBED patterns because of the periodic boundary conditions imposed by Prismatic. Second, the probe positions should not follow a rectangular grid because they need to be evenly spread across the deformed unit cell geometry in order to not bias the result towards any part of the crystal structure. Therefore, to overcome these challenges, we sample the probe positions in the following manner. The *relative* probe positions (relative to a unit cell) are arranged in a rectangular grid that is deformed along with the unit cell. The grid size is 16×28 probe positions (i.e. 448 probe positions) within the deformed unit cell (ensuring that the maximum inter-probe step size distance is $\leq 0.2 \text{ \AA}$). The *absolute* probe positions, however, are spread out as much as possible by placing each probe position in a randomly selected unit cell within a square center region (of size $200 \text{ \AA} \times 200 \text{ \AA}$) of the supercell. No two probe positions share the same unit cell. This procedure ensures all probe positions are at least 100 \AA from the supercell edge while providing an even distribution of probe positions across the unit cell geometry (to avoid sampling bias), an even distribution across a large part of the supercell (to sample local thermal configurations efficiently), and satisfaction of the Nyquist sampling criterion (for all values of λ_1 and λ_2 that are used).

All simulations are run using an accelerating beam voltage of 80 kV. Each PACBED pattern is computed as an average across 10 frozen phonon configurations. However, because many more than 1 unit cell is sampled by the *absolute* probe positions (see above), each PACBED pattern effectively averages over many more than 10 unit cell phonon configurations. All calculations utilize root-mean-square displacements of 0.07535 \AA and 0.08856 \AA for the Mo and S atoms, respectively. The model thickness along the z direction is 20 layers (10 unit cells) and PACBED patterns are saved after each layer. A potential space sampling of 0.15 \AA is used for all simulations, and given the constant supercell size, the PACBED reciprocal space FOV is constant throughout the study. Five convergence semi-angles (from here on referred to as convergence angles) are investigated, namely 6.35, 12.65, 18.13, 25.44, and 32.94 mrad, corresponding to convenient choices for our

JEOL Mono NEO ARM 200F microscope. It is important to note that because of the unit cell averaging used in this study, the best achievable PACBED spatial resolution for all the convergence angles studied is 1 unit cell. This is true even for the largest convergence angles that produce sub-Ångström sized electron beams. A spatial resolution of 1 unit cell is sufficient for characterizing most material deformations because they typically vary on the unit cell length scale or larger.

Fig. 1 shows simulated PACBED patterns without deformations for models with eight different sample thicknesses (number of layers $l = 1, 2, 3, 4, 5, 10, 15,$ and 20) and for all five electron beam convergence angles. These patterns show that for the smallest convergence angle, the Bragg discs do not overlap, and for increasing convergence angles the Bragg discs become more overlapped. Additionally, for the thinnest samples (i.e. $l = 1 - 5$), the Bragg disc intensity is uniform, while for the thickest samples, the Bragg discs exhibit non-uniform disc intensity caused by dynamical scattering. Both disc overlap and non-uniform disc intensity can make it more difficult for conventional methods to locate disc positions and measure strain, as discussed in the introduction.

Fig. 2 shows unit cell models and their associated simulated PACBED patterns for different strain tensors and rotations, but without tilts. This is shown for a sample thickness of 10 layers and two convergence angles, 6.35 and 18.13 mrad. The illustrated strain conditions are for pure normal strain ϵ_{xx} and ϵ_{yy} , pure shear strain ϵ_{xy} , and pure rotation γ . For all the deformed cases, obvious differences in the PACBED patterns are observed compared to the unstrained case. For pure normal expansive strains, such as Fig. 2(b), one can see that expansive strain in real space causes a decrease in the Bragg scattering angles in that particular direction of the PACBED pattern. This brings the Bragg discs closer together along that particular direction, even though the Bragg disc size remains the same because disc size is determined by the convergence angle. In contrast, for pure normal compressive strains, such as Fig. 2(c), one can see that compressive strain in real space causes an increase in the Bragg scattering angle in that particular direction of the PACBED pattern. For pure shear strains, such as Fig. 2(d), one can see that shear strain in real space causes an associated shearing of the Bragg scattering angles that is less intuitive to identify by eye, but cannot be recreated by a combination of normal strains. For pure rotation, such as Fig. 2(e), one can see that a rotation in real space causes the same rotation of the PACBED pattern. All of these effects can be seen both for the smaller convergence angle where discs are well-separated and for the larger convergence angle where significant disc overlap occurs. In addition to changes in the position of the Bragg discs, the specific strain tensor can also change the intensity within the Bragg discs, as can be seen more easily for the small convergence case by comparing the relative intensities of the center disc and Bragg disc. This effect does not contribute to conventional NBED strain measurements that rely only on Bragg disc positions.

Fig. 3 shows simulated PACBED patterns of models without strain and rotation but with different tilts for the same two convergence angles, 6.35 and 18.13 mrad. For the 10 layer thick sample (Fig. 3(a)–(d)), tilts produce obvious changes in the intensities of the Bragg discs but do not shift the locations of the Bragg discs. These changes in Bragg disc intensity are directed along the tilt direction. This effect is intuitive for small convergence angles where there is an overall shift of bright intensity within the Bragg discs in the tilt direction. However, while tilting is still obvious in the patterns for larger convergence angles, its effect on disc intensity is more complex due to overlapping discs. For the monolayer (Fig. 3(e)), tilts do not produce obvious changes in the disc intensities. This indicates that for very thin samples, tilts have very little effect on PACBED patterns and that prediction of tilts will be difficult.

By inspecting strain, rotation, and tilt separately, it is easier to identify how each deformation source uniquely affects PACBED patterns. However, if all of these deformations are simultaneously present in the sample with different magnitudes, it would be more difficult

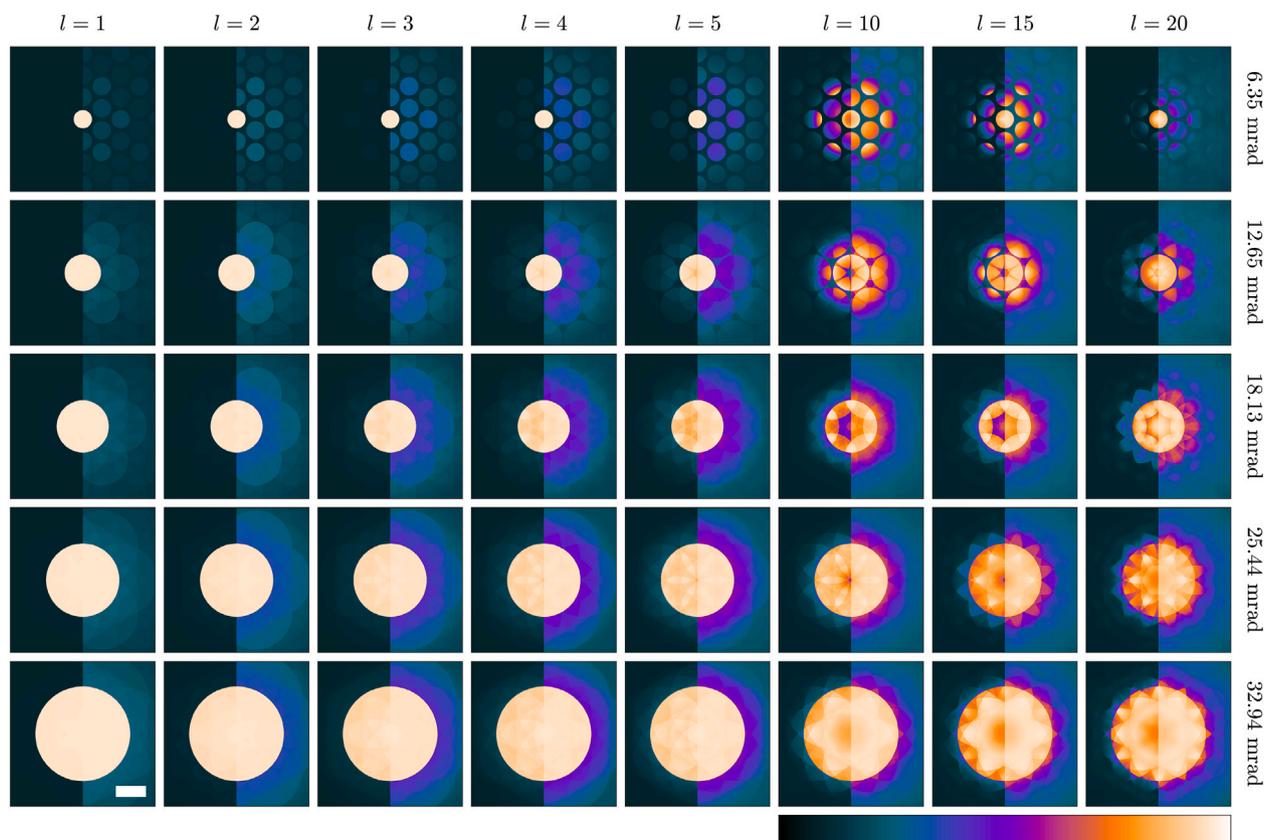


Fig. 1. Simulated PACBED patterns with no deformations for various sample thicknesses ($l = 1, 2, 3, 4, 5, 10, 15,$ and 20 layers) and five convergence angles ($6.35, 12.65, 18.13, 25.44,$ and 32.94 mrad). All the patterns are split images with the left half in linear intensity scale and the right half in cubic root intensity scale. The cubic root intensity scaling enhances the low intensity details, which is particularly important for the thin samples. The scale bar in the lower left panel corresponds to 20 mrad. The FOV is the same for every panel (100 mrad). A qualitative depiction of the colormap is provided at the bottom right.

to disentangle the contributions from each deformation source using conventional approaches.

For each convergence angle, we run 1000 PACBED simulations with unique strain and tilt conditions. The final simulated PACBED patterns are stored in a resolution of 1336^2 pixels with one pattern for each of the 20 layers. PACBED simulations are run on dual Intel Xeon Gold 6130 CPUs (32 cores). Each simulation has an average execution time of ~ 3.8 h and produces ~ 140 MB of data. The code calling Prismatic is implemented in Matlab (Mathworks, Natick, MA, US).

2.4. ML dataset generation

Datasets for ML are generated separately for eight individual thicknesses (number of layers $l = 1, 2, 3, 4, 5, 10, 15,$ and 20) and five convergence angles ($6.35, 12.65, 18.13, 25.44,$ and 32.94 mrad). Each of the datasets are used to train separate CNN models in order to study the effect of varying both the thickness and convergence angle. For each individual case, the 1,000 total PACBED simulations with varied deformations are split into training (500 simulations), validation (250 simulations), and test data subsets (250 simulations). This ensures that the three data subsets have no deformation parameter values in common, avoiding data leakage and reducing the risk of overfitting.

Each input pattern is generated by randomly selecting a pattern from the simulation dataset, and performing random transformations that include rotations, crops, rescaling, translations, noise, and intensity normalization. As discussed previously, the patterns are rotated by an angle γ , uniformly distributed in the range $[-10, 10]^\circ$. The FOV is randomly varied by $\pm 6.25\%$ by cropping the patterns to resolution $(768 \pm 48)^2$ pixels, resulting in a FOV of 100 ± 6.25 mrad, followed by rescaling to 256^2 pixels. The patterns are translated with displacements

in both x and y directions, uniformly distributed in the range $[-10, 10]$ pixels (relative to the final scale and resolution of the patterns). To add noise to the patterns, we assume that the noise is Poisson distributed and that the number of detected electrons is proportional to the pattern intensity in each pixel. To simulate varying exposure times and beam intensities, the number of simulated electrons is sampled log-uniformly in the range $[10^6, 50 \times 10^6]$. To approximate the noise behavior of the fiber coupled CMOS cameras equipped on our JEOL Mono NEO ARM 200F microscope, we approximate the Poisson distribution using a lognormal distribution with the same mean and standard deviation. Future studies could utilize different noise models, such as a pure Poisson distribution that can be observed with modern direct electron detectors. Fig. 4 shows simulated PACBED patterns that indicate the minimum and maximum noise levels used in this study. As a final step, random intensity normalization is applied so that the mean pattern intensity is uniformly distributed in the range $[0.20, 0.30]$ (the individual pixel intensities are then approximately in the range $[0, 1]$). In some cases, as will be elaborated upon below, we compute the square root of the individual pixel intensities before the final intensity normalization step.

After augmentations, ML dataset sizes are 128,000 patterns (training), 32,000 patterns (validation), and 16,000 patterns (test). The CNN inputs are 256^2 pixel PACBED patterns. The CNN outputs are the corresponding deformation parameter vectors, either six-dimensional vectors $(\epsilon_{xx}, \epsilon_{yy}, \epsilon_{xy}, \gamma, \phi_x, \phi_y)$, or, as will be elaborated upon below, four-dimensional vectors $(\epsilon_{xx}, \epsilon_{yy}, \epsilon_{xy}, \gamma)$. The code is implemented in Matlab (Mathworks, Natick, MA, US).

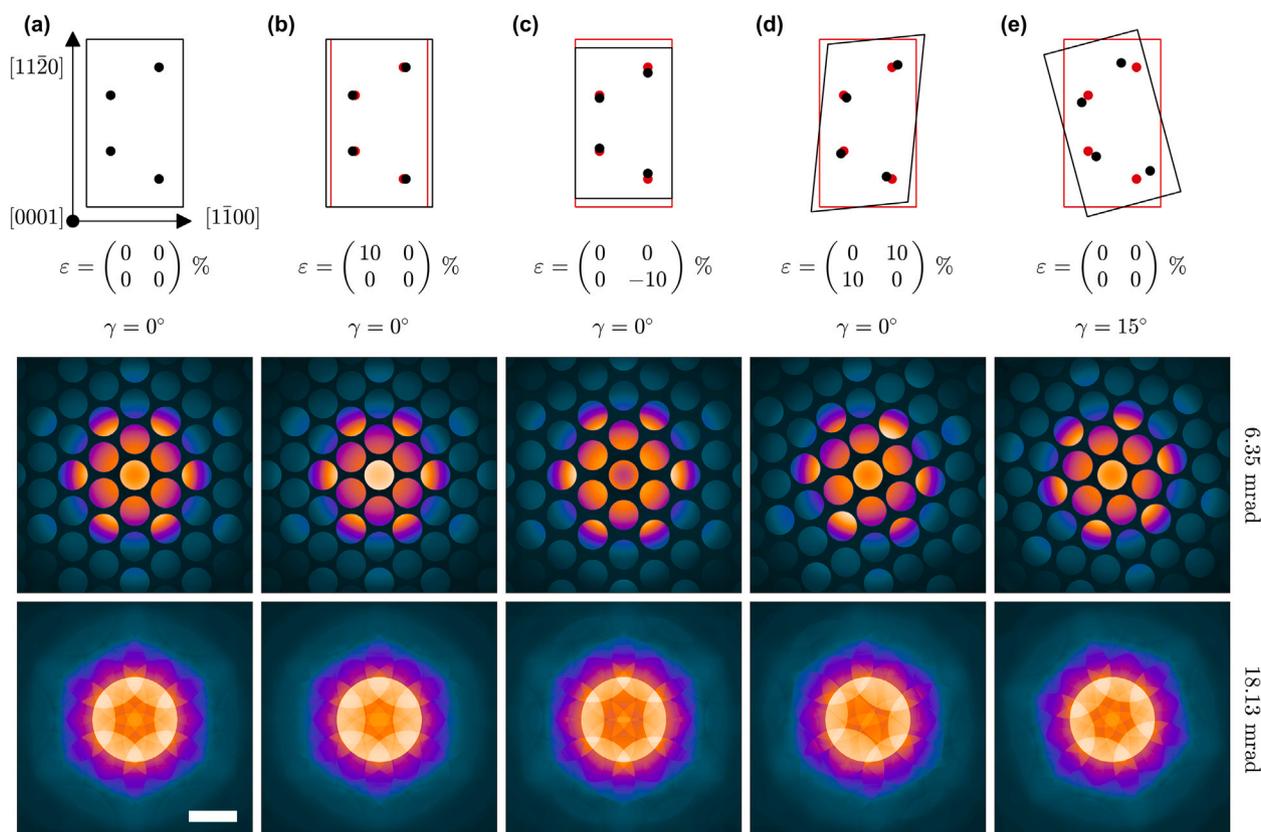


Fig. 2. Atomic models and simulated PACBED patterns for five different strain tensor and rotation cases without tilts. The sample is 10 layers thick. For each case, the top panel illustrates the deformed unit cell (black) relative to the original unit cell (red) and specifies the strain tensor and rotation angle. The crystallographic directions of the unit cell are provided in (a). The middle and bottom panels show the resulting PACBED patterns for the 6.35 and 18.13 mrad convergence angles, respectively. The cases are (a) no deformation, (b) 10% pure normal ε_{xx} expansive strain, (c) 10% pure normal ε_{yy} compressive strain, (d) 10% pure shear ε_{xy} strain, and (e) 15° pure γ rotation. The patterns are shown in cubic root intensity scale. The scale bar in the lower left panel corresponds to 20 mrad. The FOV is the same for every panel (100 mrad).

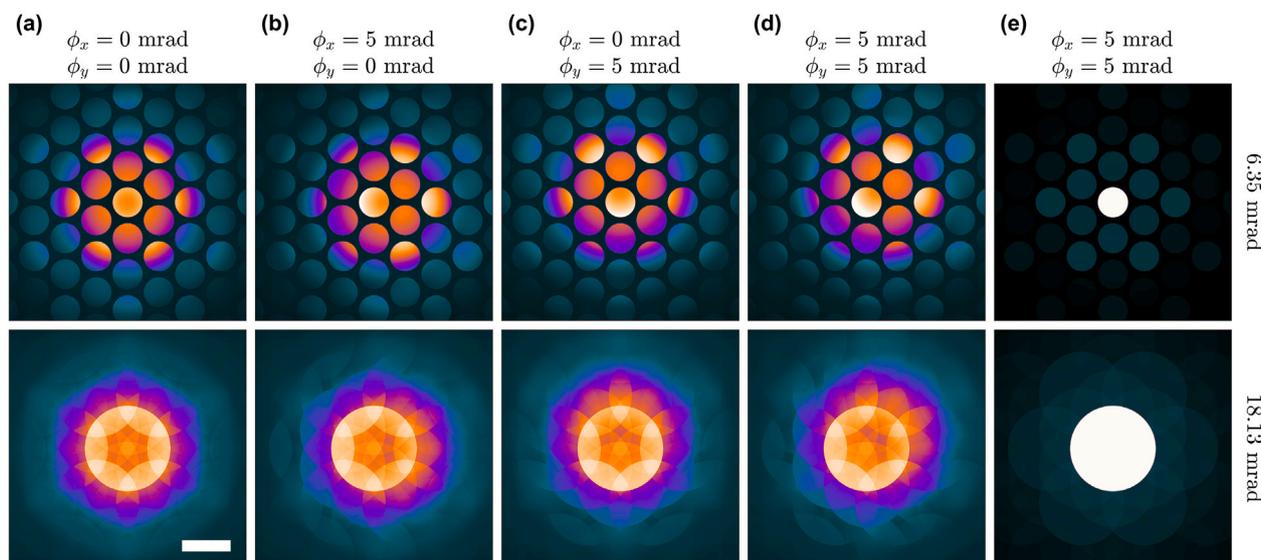


Fig. 3. Simulated PACBED patterns for different sample tilt conditions without deformations. (a)–(d) are for a 10 layer thick sample and (e) is for a monolayer. For each case, the figure illustrates the tilt angles and the resulting PACBED patterns for the 6.35 (upper) and 18.13 (lower) mrad convergence angles. The cases are (a) no tilt, (b) a pure ϕ_x , (c) a pure ϕ_y , and (d) a combined ϕ_x and ϕ_y . The patterns are shown in cubic root intensity scale. The scale bar in the lower left panel corresponds to 20 mrad. The FOV is the same for every panel (100 mrad).

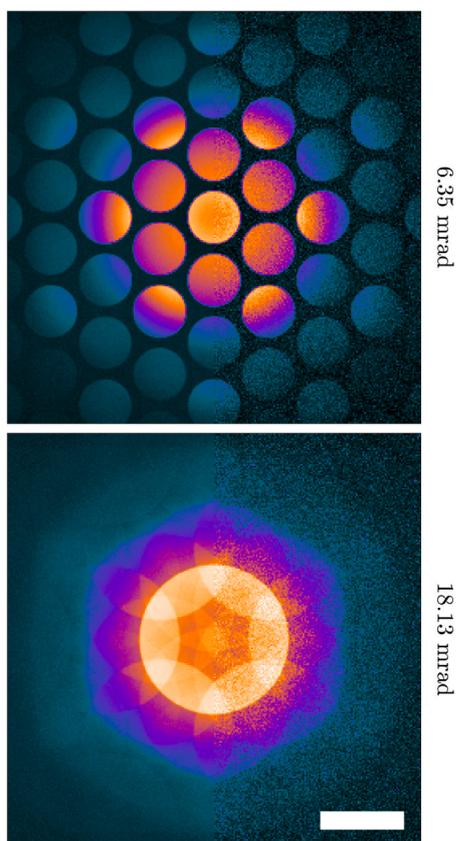


Fig. 4. Simulated PACBED patterns that demonstrate the noise range used in this study for a 10 layer thick sample for convergence angles 6.35 mrad (upper) and 18.13 mrad (lower). The images are split, showing the lowest (left) and the highest (right) noise level. The patterns are shown in cubic root intensity scale. The scale bar in the lower panel corresponds to 20 mrad. The FOV is the same for both panels (100 mrad).

2.5. CNN prediction

We use straightforward convolutional neural networks (CNNs) for predicting the deformation parameters. The CNNs consist of convolutional layers, pooling layers, and fully connected layers. In a convolutional layer, the input is convolved with several convolution kernels to produce an output. In a pooling layer, the input is downsampled to decrease resolution. The output from the convolutional and pooling layers is used by the fully connected layers to compute the final output. The CNNs here are built up of four convolutional blocks, each with two convolutional layers and one pooling layer. The number of convolutional filters applied is 32, 64, 128, and 256 for each convolutional block, with a convolution kernel size of 3×3 . The pooling layers output the average of non-overlapping 2×2 patches in the input. After the convolutional part, four fully connected layers with 512, 384, 256, and 128 nodes are used to compute the final output. After each convolutional and fully connected layer, an Elu activation function [82] is applied to the output to introduce non-linearities. The convolutional blocks can be thought of as feature extractors that extract PACBED pattern information at different scales which are then used by the fully connected layers to produce predictions. The CNN has approximately 3.5M weights. During training the CNN is optimized with respect to minimizing deviations from the target output measured by some loss function [83,84].

The CNNs are trained using stochastic gradient descent (SGD) with momentum 0.9 [85,86] and a batch size of 256. Data augmentation is performed on the training data by varying the image mean uniformly

in the range [0.20, 0.30]. The learning rate is increased stepwise from 10^{-4} to 10^{-3} and then decreased stepwise again to 10^{-4} , comprising a total of 4,000 epochs. The CNN is optimized with respect to mean squared error (MSE) loss,

$$\text{MSE} = \langle (\hat{y} - y)^2 \rangle, \quad (4)$$

where y is the target (true) deformation vector and \hat{y} is the predicted deformation vector. Note that the average in the MSE loss is computed across all samples and all individual parameters in the deformation vector. All parameters are rescaled to the range $[-1, 1]$ for training. Across all epochs, the best-performing model with respect to the MSE of the validation set is selected. Training is performed on single NVIDIA A100 GPUs, and the execution time is approximately 60 h for each run. The networks are implemented in Tensorflow [87].

3. Results and discussion

3.1. Prediction performance

Instead of using the MSE for assessing prediction performance, we use the more intuitive mean absolute error (MAE),

$$\text{MAE}_i = \langle |\hat{y}_i - y_i| \rangle, \quad (5)$$

where y_i and \hat{y}_i are the target (true) and predicted values of the i th deformation parameter, respectively. Due to limited space, we cannot show a performance comparison between training, validation, and test datasets for all thicknesses, convergence angles, and deformation parameters. In all cases, the training, test, and validation MAEs are comparable, indicating that the models generalize reasonably well and overfitting is not a significant problem. Therefore, we assess the prediction performance of the various CNNs using the test datasets.

In order to analyze the prediction performance trends of the CNNs across the studied sample thicknesses, convergence angles and deformation parameters, we show the MAE for all CNNs in Fig. 5 and scatter plots for a selected set of parameters in Fig. 6. The scatter plots show the true vs. predicted values for three of the six parameters, ϵ_{xx} , γ and ϕ_x . We do not show scatter plots of ϵ_{yy} and ϵ_{xy} because they are both nearly identical to ϵ_{xx} . Likewise, we do not show ϕ_y because it is nearly identical to ϕ_x . Further, the scatter plots are only shown for five thicknesses ($l = 1, 3, 5, 10,$ and 20) and three convergence angles (6.35, 18.13, and 32.94 mrad). It is worth pointing out that the prediction error is approximately Gaussian, and hence the scatter plots are denser close to the $y = x$ line than further away. This is not obvious from the scatter plots, making the prediction errors appear larger than they are on average. Fig. 5 and Fig. 6 illustrate five clear trends in the CNN performance.

First, Fig. 5 shows the prediction trends for all the strain tensor parameters (ϵ_{xx} , ϵ_{yy} , ϵ_{xy}) are comparable, and the prediction trends for both tilt parameters (ϕ_x , ϕ_y) are comparable. This is expected behavior and indicates that similar types of deformation parameters are treated similarly by the CNNs and offer similar difficulties for prediction.

Second, for all deformation parameters, the MAE increases with increasing convergence angle, highlighting that CNN prediction is more reliable for smaller convergence angles. This trend is also observed in the scatter plots where there is a wider scatter plot distribution for increasing convergence angle. This is not a surprise considering that, with increasing convergence angle, there is an increasing overlap of the diffraction discs, making it more difficult to identify the details of individual discs. However, it is important to note that all of the deformations can still be predicted with impressive reliability even for the largest convergence angles where conventional strain mapping methods are not possible. This indicates that these CNNs could enable large convergence angle atomic-scale strain mapping, but there is an obvious trade-off between deformation prediction performance and probe size.

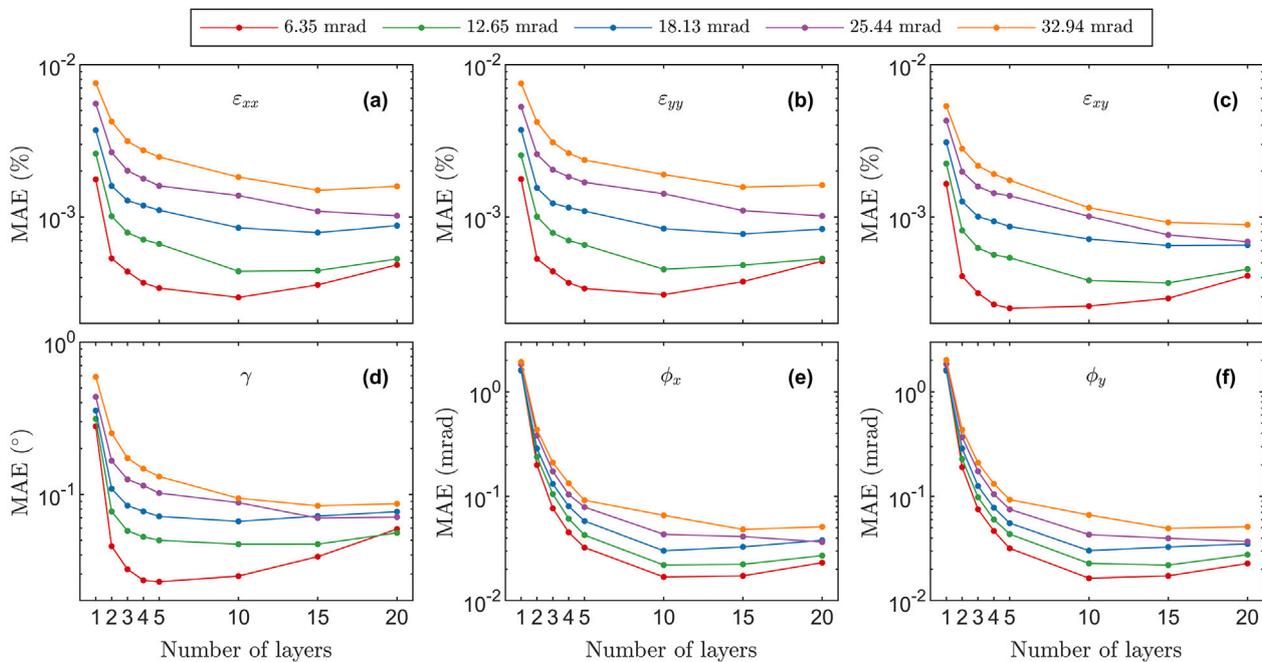


Fig. 5. Prediction performance in terms of mean absolute error (MAE) for all convergence angles and thicknesses for the parameters (a) ε_{xx} , (b) ε_{yy} , (c) ε_{xy} , (d) γ , (e) ϕ_x , and (f) ϕ_y . The y axes is in log scale.

Third, as discussed earlier and shown in Fig. 3(e), there is essentially no tilt information in patterns from very thin samples. This observation is evident in the poor tilt prediction performance for thin samples ($l = 1-5$). For example, the MAE values for tilt prediction of monolayers is around 2 mrad, which is poor considering that this is 20% of the total tilt range (± 5 mrad). The scatter plots for tilt prediction of monolayers (Fig. 6(c)) also reveal poor reliability for all convergence angles. They exhibit the broadest scatter plot distributions and poorest performance across all the CNN models. However, for $l \geq 5$, tilt prediction is reliable for all convergence angles, with $\text{MAE} \leq 0.1$ mrad and narrow scatter plot distributions.

Fourth, for all deformation parameters, the MAE decreases sharply as a function of thickness for small thicknesses. Indeed, the scatter plots show broader distributions for the smallest thicknesses. These observations highlight a significant decrease in CNN prediction performance for thin samples, especially monolayers. There are two sources for the increased errors for thin samples. One source is that for very thin samples, the intensities of the Bragg discs (which hold all the information about deformations) are extremely low (as can be seen in Fig. 1). Hence, for small thicknesses, the Bragg signals can easily be lost in the noise. This problem is clearly alleviated by increasing the thickness (at around 5 layers), which creates more elastic scattering and higher Bragg disc intensities, consequently making deformation prediction easier. A second source is that the poor tilt prediction performance for thin samples (discussed in point 3 above) hinders the prediction of the strain and rotation deformation parameters. The rescaling of all deformation parameters to the range $[-1, 1]$ removes biases in the MSE loss function that would otherwise be present due to vastly different original ranges. However, if there is a substantial variation in the difficulty of predicting different parameters, the MSE loss function will still introduce a ‘bias’ in the sense that the stochastic gradient descent optimizer will focus most of its efforts on reducing the largest errors, which for $l \leq 5$ would generally occur for the tilt parameters. Therefore, the reduced prediction performance of strain and rotation for thin samples is partly due to the difficulty of predicting tilt. Potential solutions to alleviate both of these sources of increased error in thin samples will be discussed below in Section 3.2.

Fifth, deformation prediction performance gets slightly worse as thickness increases (i.e. $l = 15-20$), with increasing MAE and slightly

broader scatter plot distributions. This increase could be attributed to increasing dynamical scattering effects with thickness, which create non-uniform disc intensities and an increased intricacy of the patterns, overshadowing the deformation information. Indeed, when visually inspecting PACBED patterns for even larger thicknesses than shown in this study (up to 164 layers), there are various dynamical scattering features in the patterns that get entangled with the deformation information, suggesting that the deformation prediction errors would likely continue to increase for larger thicknesses.

It is worth discussing the connection between the error metrics used here and those of accuracy (the deviation between the true value and the average of the predicted values for a given true value) and precision (standard deviation or variance of the predicted values for a given true value). The usual error metrics in ML, e.g. MSE and MAE, incorporates both accuracy and precision, meaning that the error cannot easily be interpreted in terms of accuracy and precision individually. However, we observe that in most scatter plots (Fig. 6), the trends mostly follow the line $y = x$, with the notable exception of tilt prediction for monolayers. This indicates that in most cases the prediction performance (Fig. 5) is limited by precision rather than accuracy. Accuracy and precision metrics are defined differently in different studies, and this study analyzes MAE. For this reason we do not attempt to directly compare these metrics across different studies.

3.2. Improving prediction performance for monolayers

We explore two routes towards improving the prediction performance of thin samples ($l \leq 5$) that are motivated by the two sources for the increased errors for thin samples identified above. In Fig. 7, we illustrate the improvement in the prediction performance in terms of MAE of these two routes for the 6.35 mrad convergence angle and for thicknesses $l \leq 5$, and in Fig. 8 we illustrate the improvements from the first route with scatter plots for a monolayer ($l = 1$).

The first route is to train CNNs to only predict strain and rotation deformation parameters and not tilt. This is motivated by the observation that the CNNs have severe difficulty predicting tilts for thin samples, affecting prediction of the strain and rotation deformation parameters as well. Fig. 7 reveals that there is a substantial reduction in error

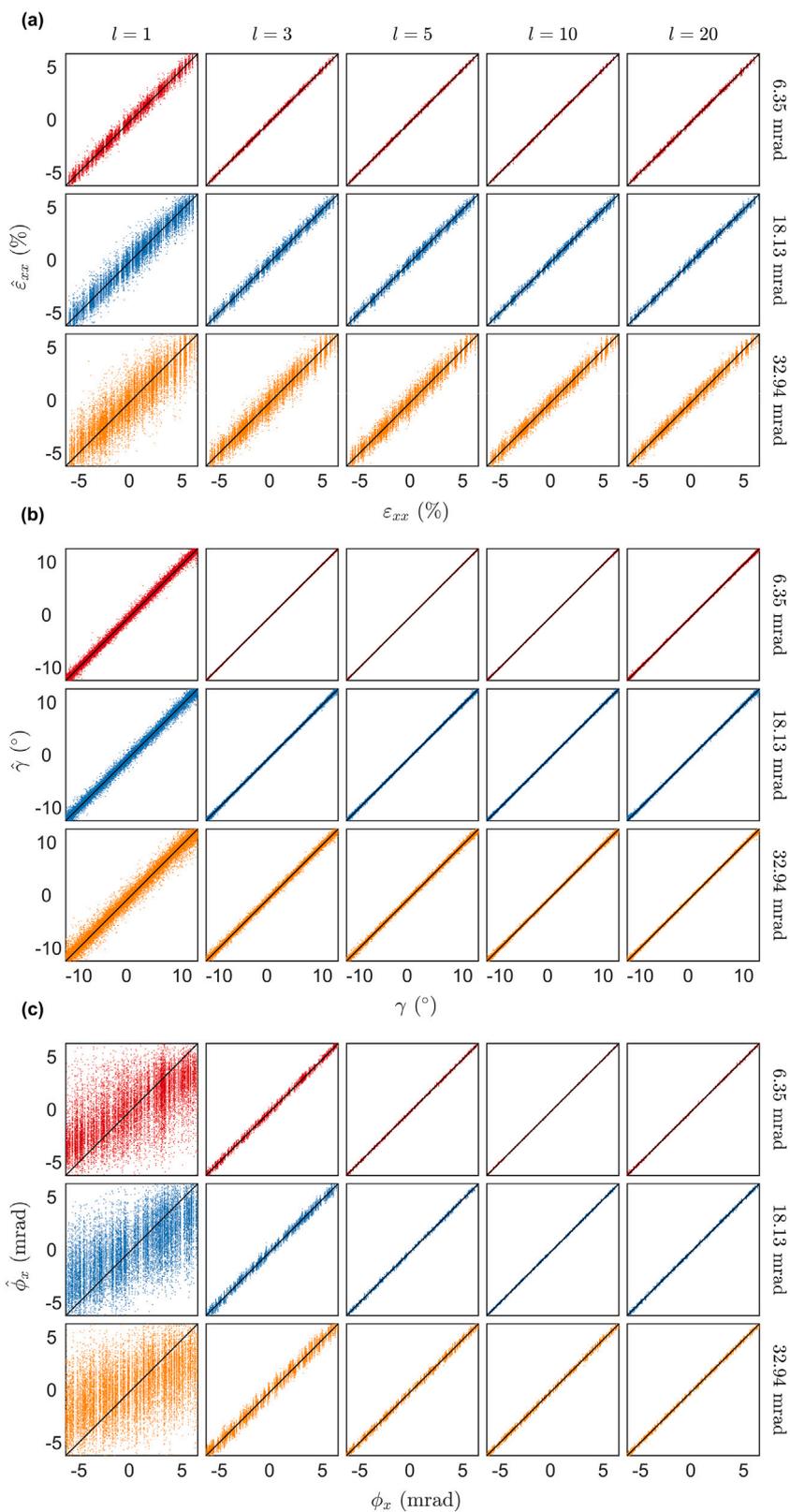


Fig. 6. Scatter plots of the true (x axis) vs. predicted (y axis) values for three of the six deformation parameters: (a) ϵ_{xx} , (b) γ , and (c) ϕ_x . Due to lack of space, we do not show the scatter plots for ϵ_{yy} and ϵ_{xy} because they are both nearly identical to ϵ_{xx} , and we do not show the scatter plot for ϕ_y because it is nearly identical to ϕ_x . Further, results are shown only for a subset of the thicknesses and convergence angles.

when removing the tilt parameters from the training (black curve vs. red curve), proving that a major difficulty in predicting the strain and rotation parameters for $l = 1$ is simply due to the presence of the tilt

parameters in the same loss function. To a much smaller degree, the same holds for $l = 2 - 5$. Fig. 8 shows scatter plots of the true vs. predicted values for these four deformation parameters when including

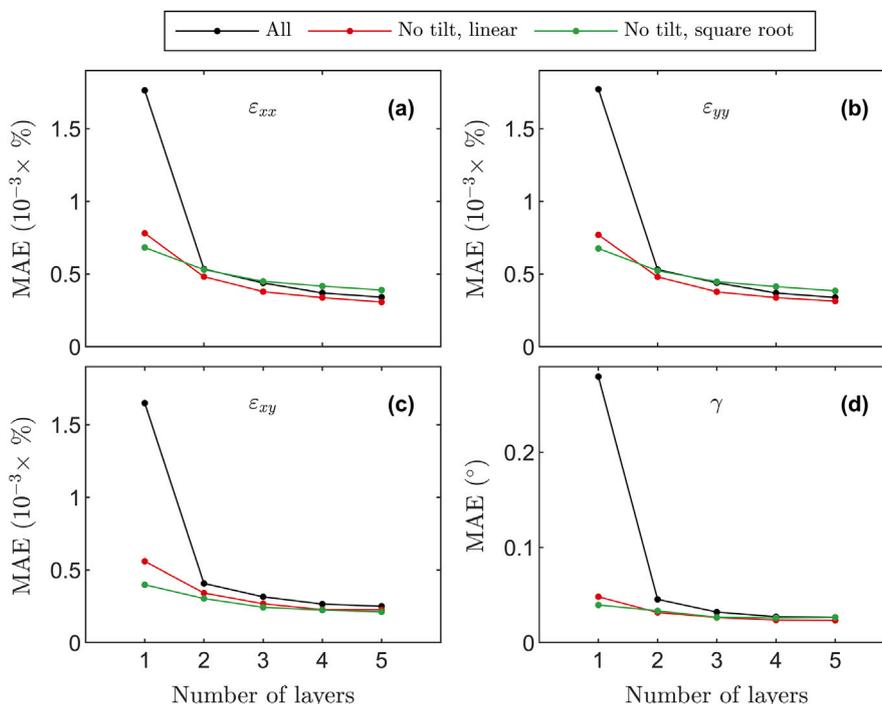


Fig. 7. Prediction performance in terms of mean absolute error (MAE) for convergence angle 6.35 and thicknesses $l \leq 5$ for the parameters (a) ϵ_{xx} , (b) ϵ_{yy} , (c) ϵ_{xy} , and (d) γ . Errors are shown for predictions from the CNNs trained to predict all parameters on a linear intensity scale (black), trained to predict only strain and rotation parameters (no tilt) on a linear intensity scale (red), and the same on a square root intensity scale (green).

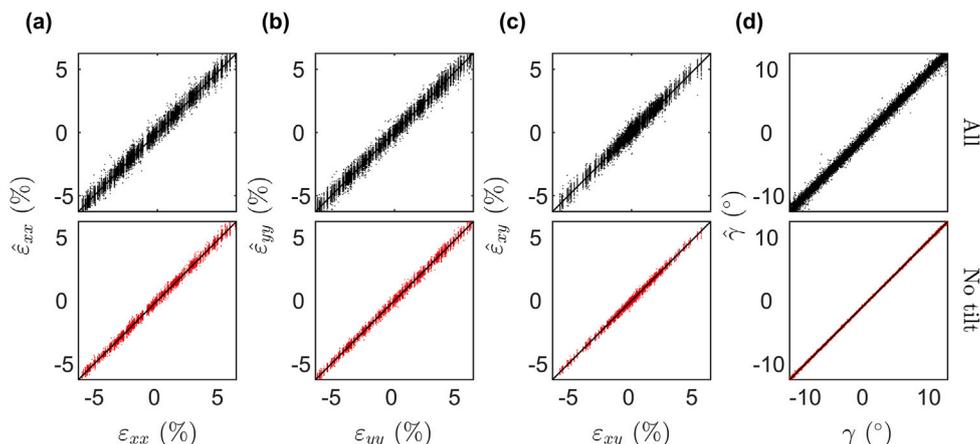


Fig. 8. Scatter plots of the true (x axis) vs predicted (y axis) values for the strain and rotation deformation parameters: (a) ϵ_{xx} , (b) ϵ_{yy} , (c) ϵ_{xy} , and (d) γ . Scatter plots are shown for the CNNs trained to predict all six parameters (top) and the CNNs trained to predict only the four strain and rotation deformation parameters (bottom). Scatter plots are shown for convergence angle 6.35 mrad, only monolayer ($l = 1$), and a linear intensity scale.

(top) and not including (bottom) tilt prediction, revealing a noticeably narrower distribution when not including tilt prediction.

The second route is to train CNNs on PACBED patterns with nonlinear intensity scaling. This is motivated by the observation that for thin samples, the intensities of the Bragg discs are extremely low. As can be seen in Fig. 1, the use of a nonlinear intensity scaling can be used to more easily see low intensity features in the patterns. To this end, we generate training data with a square root intensity scaling, which effectively compresses the pattern intensities by reducing contrast between the Bragg discs and center disc. The idea is that in this manner, CNNs can extract more deformation information. Indeed, as seen in Fig. 7, the square root intensity scaling provides a small improvement in prediction performance for $l = 1$, but has a much smaller impact on prediction performance compared to the first route. For $l = 2-5$, there is no consistent improvement at all. This indicates that nonlinear intensity

scaling is a promising approach to improve prediction performance, but only for monolayers.

3.3. Generalization and extensions

Performing PACBED simulations requires specifying parameters both for the material, such as composition, crystal phase, and thickness, and for the experiment, such as electron beam energy, convergence angle, and diffraction field of view. This requires detailed prior information to train CNN prediction models. In this study, we kept some parameters such as beam voltage, material composition and crystal phase constant to simplify the study and isolate the impact of changing individual parameters on prediction performance. We cannot expect the CNN prediction models to generalize to new electron beam conditions or new materials.

Hence, the CNN prediction models are only as generally applicable as the data they are trained on, and there are two possible solutions to this limitation. First, CNNs could be trained to cope with a larger number and range of different electron beam and material parameters, e.g. convergence angle and thickness. Alternatively, a larger number of separate and more specific prediction models could be trained for different electron beam conditions and materials. Concerning the convergence angle, it is typically well known a priori, so it is feasible to utilize the second solution and train a CNN for a single specific convergence angle. However, sample thickness is generally not known a priori, so both solutions could be utilized. First, a CNN intended for real use could be trained on a range of thicknesses, while recognizing the fact that including a 'difficult' thickness i.e. monolayers might be detrimental to the overall performance. The second solution is to determine the thickness (using e.g. PACBED or electron energy loss spectroscopy) before deformation prediction and use this prior information to select a suitable CNN that was trained on simulated data from the correct thickness.

Regarding the first option, given that PACBED patterns can be simulated for a larger parameter range, a single CNN can, in principle, be trained to predict deformation parameters regardless of this extension. However, because including some of these parameters could considerably compromise prediction performance, the inclusion of more parameters could decrease the overall performance. Furthermore, extending the ranges of the deformation parameters could eventually decrease the overall prediction performance if extreme deformations are included. Regarding the second option, since each of these specialized models would require a set of prior information (i.e. beam and material information) in order to select the suitable model, this solution could be organized into a hierarchical system. It could be composed of operator decisions, additional analysis steps, and other CNNs that provide the required prior knowledge.

In this work, although averaging is performed over unit cell sized regions, exploring the use of sub-Ångström sized probes is still of interest for two reasons. First, most STEM experiments use atomic scale probes for imaging and spectroscopy, and when deformation information is desired, it is often beneficial to collect both atomic resolution images and NBED data. Furthermore, it is tedious to switch between atomic resolution STEM imaging conditions and conventional NBED deformation mapping conditions. Additionally, it is ideal to collect registered atomic resolution images and diffraction deformation maps from precisely the same areas. This work demonstrates that the same atomic scale probe can be used to perform standard atomically resolved imaging and diffraction deformation mapping. Furthermore, in principle, the two can be done simultaneously by post-processing a single 4D STEM data set. Second, sub-unit cell PACBED averaging is an intriguing method for extracting sub-unit cell information, as has recently been explored for thickness measurements [74]. Measuring deformation information from sub-unit cell regions (e.g. individual atomic columns) could be a powerful method for measuring very local deformations around e.g. point defects and surfaces. This work of using atomic scale probes for diffraction deformation measurements constitutes an important step towards this idea.

There are numerous directions to pursue for future work. First, the study would benefit from investigating the prediction performance for different diffraction fields of view. Second, it would be of interest to train separate networks for strain, rotation, and tilt to likely obtain better overall performance. To explore this while reducing the need for computational resources, it could be efficient to train a network on predicting all parameters. Then, using a transfer learning approach, let several instances of that network specialize towards predicting individual parameters. Third, given the indication that nonlinear intensity scaling might be beneficial, it could be beneficial to implement a more general parameterization of a nonlinear scaling as a first layer in the CNN. This training would then learn to optimize the scaling parameters. Fourth, the CNNs in this study were trained on deformed single

crystalline models without the introduction of any crystalline defects. Future work could explore training CNNs to predict deformations even in the presence of defects. Finally, the current investigation is limited to determining 2D strain tensors. Extending this capability to explore prediction of the 3D strain tensor is relevant to many materials.

4. Conclusion

We have performed a comprehensive simulation study on the performance of PACBED combined with CNNs for predicting the local 2D strain tensor, rotation, and tilt of thin 2D materials. We study 2H MoS₂ as an example. For five convergence angles ranging from conventional NBED conditions (6.35 mrad) to atomic resolution conditions (32.94 mrad), and eight thicknesses ranging from 1 to 20 layers, we have simulated 100,000 PACBED patterns with varying strain, rotation and tilt parameters. We train CNN models to simultaneously predict all these deformation parameters and study the prediction performance as a function of convergence angle and thickness. In general, we find excellent prediction performance across all conditions, but there is a trade-off between better prediction performance (small convergence angles) and probe size (large convergence angles). For smaller convergence angles like those used for conventional NBED conditions, the strain prediction error can be as low as 0.0003%, while for larger convergence angles like those used for atomic resolution probes, the strain error increases to 0.001–0.003%. The better prediction performance for small convergence angles is likely due to reduced overlap of discs in the PACBED pattern. The impressive prediction performance even for large convergence angles suggests that PACBED combined with CNNs is a feasible method for predicting deformation parameters using atomic resolution electron probes. For monolayers, we show that the prediction of strain and rotation improves considerably when excluding tilt prediction in the same model. This is due to the CNNs focusing on the parameters that are most difficult to predict, which for thin samples are the tilt parameters. For the tilt ranges we study here, there is negligible effect of tilt on PACBED intensity for monolayers. To a lesser extent, we demonstrate that for monolayers it can be beneficial to use a nonlinear intensity scaling of the images in the training data. In our case, we used square root scaling, which gave a consistent improvement across parameters for monolayers. This work demonstrates that PACBED combined with CNNs is a viable method for pushing the limits of measuring deformations with low errors and high spatial resolution for a wide range of convergence angles and thicknesses. Furthermore, this study will help guide the design of PACBED experiments for measuring deformations in 2D vdW materials.

CRedit authorship contribution statement

Andrew B. Yankovich: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Magnus Röding:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Eva Olsson:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andrew Yankovich reports financial support was provided by the Swedish Research Council and the Swedish Foundation for Strategic Research. Eva Olsson reports financial support was provided by the Knut and Alice Wallenberg Foundation. Magnus Röding reports financial support was provided by the Swedish Research Council. Magnus

Röding reports financial support was provided by the Swedish Research Council Formas. Magnus Röding reports a relationship with Astra Zeneca that includes: employment and equity or stocks. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

A.B.Y. and E.O. acknowledge funding from the Knut and Alice Wallenberg Foundation, Sweden (KAW, grant No. 2019.0140). A.B.Y. acknowledges funding from the Swedish Research Council (VR starting, grant No. 2020-04986) and the Swedish Research Council and Swedish Foundation for Strategic Research for access to ARTEMI, the Swedish National Infrastructure in Advanced Electron Microscopy (grant No. 2021-00171 and RIF21-0026). M.R. acknowledges the financial support from the Swedish Research Council for Sustainable Development (grant No. 2019-01295) and the Swedish Research Council (grant No. 2023-04248). The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council (grant No. 2022-06725). A.B.Y. and E.O. declare no conflicts of interest. M.R. is an employee of AstraZeneca and has stock ownership and/or stock options or interests in the company.

Data availability

Data will be made available on request.

References

- [1] Kostya S. Novoselov, Andre K. Geim, Sergei V. Morozov, De-eng Jiang, Yanshui Zhang, Sergey V. Dubonos, Irina V. Grigorieva, Alexandr A. Firsov, Electric field effect in atomically thin carbon films, *Science* 306 (2004) 666–669.
- [2] Andre K. Geim, Irina V. Grigorieva, Van der Waals heterostructures, *Nature* 499 (2013) 419–425.
- [3] Yuan Liu, Nathan O. Weiss, Xidong Duan, Hung-Chieh Cheng, Yu Huang, Xiangfeng Duan, Van der Waals heterostructures and devices, *Nat. Rev. Mater.* 1 (2016) 1–17.
- [4] Rafi Bistritzer, Allan H. MacDonald, Moiré bands in twisted double-layer graphene, *Proc. Natl. Acad. Sci.* 108 (2011) 12233–12237.
- [5] Yuan Cao, Valla Fatemi, Ahmet Demir, Shiang Fang, Spencer L. Tomarken, Jason Y. Luo, Javier D. Sanchez-Yamagishi, Kenji Watanabe, Takashi Taniguchi, Efthimos Kaxiras, et al., Correlated insulator behaviour at half-filling in magic-angle graphene superlattices, *Nature* 556 (2018) 80–84.
- [6] Roscoe G. Dickinson, Linus Pauling, The crystal structure of molybdenite, *J. Am. Chem. Soc.* 45 (1923) 1466–1471.
- [7] J.A. Wilson, A.D. Yoffe, The transition metal dichalcogenides discussion and interpretation of the observed optical, electrical and structural properties, *Adv. Phys.* 18 (1969) 193–335.
- [8] Branimir Radisavljevic, Aleksandra Radenovic, Jacopo Brivio, Valentina Giacometti, Andras Kis, Single-layer MoS₂ transistors, *Nature Nanotechnology* 6 (2011) 147–150.
- [9] Xiaodong Xu, Wang Yao, Di Xiao, Tony F. Heinz, Spin and pseudospins in layered transition metal dichalcogenides, *Nat. Phys.* 10 (2014) 343–350.
- [10] F.H.L. Koppens, Thomas Mueller, Ph Avouris, A.C. Ferrari, M. Serena Vitiello, M. Polini, Photodetectors based on graphene, other two-dimensional materials and hybrid systems, *Nature Nanotechnology* 9 (2014) 780–793.
- [11] Kin Fai Mak, Jie Shan, Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides, *Nat. Photonics* 10 (2016) 216–226.
- [12] Sajedah Manzeli, Dmitry Ovchinnikov, Diego Pasquier, Oleg V. Yazyev, Andras Kis, 2D transition metal dichalcogenides, *Nat. Rev. Mater.* 2 (2017) 1–15.
- [13] Chiara Trovatiello, Andrea Marini, Xinyi Xu, Changhwan Lee, Fang Liu, Nicola Curreli, Cristian Manzoni, Stefano Dal Conte, Kaiyuan Yao, Alessandro Ciattoni, et al., Optical parametric amplification by monolayer transition metal dichalcogenides, *Nat. Photonics* 15 (2021) 6–10.
- [14] Jannik C. Meyer, Andre K. Geim, Mikhail I. Katsnelson, Konstantin S. Novoselov, Tim J. Booth, Siegmund Roth, The structure of suspended graphene sheets, *Nature* 446 (2007) 60–63.
- [15] Annalisa Fasolino, J.H. Los, Mikhail I. Katsnelson, Intrinsic ripples in graphene, *Nat. Mater.* 6 (2007) 858–861.
- [16] Ke Xu, Peigen Cao, James R. Heath, Scanning tunneling microscopy characterization of the electrical properties of wrinkles in exfoliated graphene monolayers, *Nano Lett.* 9 (2009) 4446–4451.
- [17] Maximilian Haider, Stephan Uhlemann, Eugen Schwan, Harald Rose, Bernd Kabius, Knut Urban, Electron microscopy image enhanced, *Nature* 392 (1998) 768–769.
- [18] O.L. Krivanek, N. Dellby, A.R. Lupini, Towards sub-Å electron beams, *Ultramicroscopy* 78 (1999) 1–11.
- [19] O.L. Krivanek, P.D. Nellist, N. Dellby, M.F. Murfitt, Z. Szilagy, Towards sub-0.5 Å electron beams, *Ultramicroscopy* 96 (2003) 229–237.
- [20] Christian Kisielowski, B. Freitag, M. Bischoff, H. Van Lin, S. Lazar, Georg Knippels, Peter Tiemeijer, Maarten van der Stam, Sebastian von Harrach, Michael Stekelenburg, et al., Detection of single atoms and buried defects in three dimensions by aberration-corrected electron microscope with 0.5-Å information limit, *Microsc. Microanal.* 14 (2008) 469–477.
- [21] Rolf Erni, Marta D. Rossell, Christian Kisielowski, Ulrich Dahmen, Atomic-resolution imaging with a sub-50-pm electron probe, *Phys. Rev. Lett.* 102 (2009) 096101.
- [22] Hidetaka Sawada, Yasumasa Tanishiro, Nobuhiro Ohashi, Takeshi Tomita, Fumio Hosokawa, Toshiyuki Kaneyama, Yukihito Kondo, Kunio Takayanagi, STEM imaging of 47-pm-separated atomic columns by a spherical aberration-corrected electron microscope with a 300-kV cold field emission gun, *J. Electron Microsc.* 58 (2009) 357–361.
- [23] Martin J. Hÿtch, Andrew M. Minor, Observing and measuring strain in nanostructures and devices with transmission electron microscopy, *MRS Bull.* 39 (2014) 138–146.
- [24] David Cooper, Thibaud Denneulin, Nicolas Bernier, Armand Béch e, Jean-Luc Rouvi re, Strain mapping of semiconductor specimens with nm-scale resolution in a transmission electron microscope, *Micron* 80 (2016) 145–165.
- [25] M.J. Hÿtch, E. Snoeck, R. Kilaas, Quantitative measurement of displacement and strain fields from HREM micrographs, *Ultramicroscopy* 74 (1998) 131–146.
- [26] M.J. Hÿtch, T. Plamann, Imaging conditions for reliable measurement of displacement and strain in high-resolution electron microscopy, *Ultramicroscopy* 87 (2001) 199–212.
- [27] Martin Hÿtch, Florent Houdellier, Florian H e, Etienne Snoeck, Nanoscale holographic interferometry for strain measurements in electronic devices, *Nature* 453 (2008) 1086–1089.
- [28] R. Bierwolf, M. Hohenstein, F. Phillipp, O. Brandt, G.E. Crook, K. Ploog, Direct measurement of local lattice distortions in strained layer structures by HREM, *Ultramicroscopy* 49 (1993) 273–285.
- [29] Pedro L. Galindo, Slawomir Kret, Ana M. Sanchez, Jean-Yves Laval, Andres Yanez, Joaquin Pizarro, Elisa Guerrero, Teresa Ben, Sergio I. Molina, The peak pairs algorithm for strain mapping from HRTEM images, *Ultramicroscopy* 107 (2007) 1186–1193.
- [30] Jian-Min Zuo, Amish B. Shah, Honggyu Kim, Yifei Meng, Wenpei Gao, Jean-Luc Rouvi re, Lattice and strain analysis of atomic resolution Z-contrast images based on template matching, *Ultramicroscopy* 136 (2014) 50–60.
- [31] Andrew B. Yankovich, Benjamin Berkels, Wolfgang Dahmen, Peter Binev, Sergio I. Sanchez, Steven A. Bradley, Ao Li, Izabela Szlufarska, Paul M. Voyles, Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts, *Nat. Commun.* 5 (2014) 4155.
- [32] Torben Nilsson Pingel, Mikkel Jørgensen, Andrew B. Yankovich, Henrik Gr nbeck, Eva Olsson, Influence of atomic site-specific strain on catalytic activity of supported nanoparticles, *Nat. Commun.* 9 (2018) 2722.
- [33] L.D. Marks, Direct atomic imaging of solid surfaces: I. Image simulation and interpretation, *Surf. Sci.* 139 (1984) 281–298.
- [34] M. Jos e Yacam n, J.A. Ascencio, G. Canizal, Observation of surface relaxation surface steps and surface reconstruction in gold nanorods, *Surf. Sci.* 486 (2001) L449–L453.
- [35] K. Du, F. Ernst, M.C. Pelsozy, J. Barthel, K. Tillmann, Expansion of interatomic distances in platinum catalyst nanoparticles, *Acta Mater.* 58 (2010) 836–845.
- [36] Takeshi Fujita, Pengfei Guan, Keith McKenna, Xingyuo Lang, Akihiko Hirata, Ling Zhang, Tomoharu Tokunaga, Shigeo Arai, Yuta Yamamoto, Nobuo Tanaka, et al., Atomic origins of the high catalytic activity of nanoporous gold, *Nat. Mater.* 11 (2012) 775–780.
- [37] Miguel Lopez-Haro, Kenta Yoshida, Eloy Del Rio, Jose A. Perez-Omil, Edward D. Boyes, Susana Trasobares, Jian-Min Zuo, Pratibha L. Gai, Jose J. Calvino, Strain field in ultrasmall gold nanoparticles supported on cerium-based mixed oxides. Key influence of the support redox state, *Langmuir* 32 (2016) 4313–4322.
- [38] G nter M bus, R. Schweinfest, T. Gemming, T. Wagner, M. R hle, Iterative structure retrieval techniques in HREM: a comparative study and a modular program package, *J. Microsc.* 190 (1998) 109–130.
- [39] Sara Bals, Sandra Van Aert, Gustaaf Van Tendeloo, David Avila-Brandee, Statistical estimation of atomic positions from exit wave reconstruction with a precision in the picometer range, *Phys. Rev. Lett.* 96 (2006) 096106.
- [40] K. Tillmann, M. Lentzen, R. Rosenfeld, Impact of column bending in high-resolution transmission electron microscopy on the strain evaluation of GaAs/InAs/GaAs heterostructures, *Ultramicroscopy* 83 (2000) 111–128.
- [41] Andreas Rosenauer, D. Gerthsen, V. Potin, Strain state analysis of InGaN/GaN-sources of error and optimized imaging conditions, *Phys. Status Solidi (A)* 203 (2006) 176–184.

- [42] K. Du, F. Philipp, On the accuracy of lattice-distortion analysis directly from high-resolution transmission electron micrographs, *J. Microsc.* 221 (2006) 63–71.
- [43] Elisa Guerrero, Pedro Galindo, Andrés Yáñez, Teresa Ben, Sergio I. Molina, Error quantification in strain mapping methods, *Microsc. Microanal.* 13 (2007) 320–328.
- [44] Jayhoon Chung, Lew Rabenberg, Effects of strain gradients on strain measurements using geometrical phase analysis in the transmission electron microscope, *Ultramicroscopy* 108 (2008) 1595–1602.
- [45] Jacob Madsen, Pei Liu, Jakob B. Wagner, Thomas W. Hansen, Jakob Schiøtz, Accuracy of surface strain measurements from transmission electron microscopy images of nanoparticles, *Adv. Struct. Chem. Imaging* 3 (2017) 1–12.
- [46] S.J. Pennycook, D.E. Jesson, High-resolution Z-contrast imaging of crystals, *Ultramicroscopy* 37 (1991) 14–38.
- [47] Xiaonan Luo, Aakash Varambhia, Weixin Song, Dogan Ozkaya, Sergio Lozano-Perez, Peter D. Nellist, High-precision atomic-scale strain mapping of nanoparticles from STEM images, *Ultramicroscopy* 239 (2022) 113561.
- [48] P.M. Jones, G.M. Rackham, John Wickham Steeds, Higher order laue zone effects in electron diffraction and their use in lattice parameter determination, *Proc. R. Soc. A* 354 (1977) 197–222.
- [49] A. Armigliato, R. Balboni, G.P. Carnevale, G. Pavia, D. Piccolo, Stefano Frabboni, A. Benedetti, A.G. Cullis, Application of convergent beam electron diffraction to two-dimensional strain mapping in silicon devices, *Appl. Phys. Lett.* 82 (2003) 2172–2174.
- [50] Muneyuki Naito, Manabu Ishimaru, Yoshihiko Hirotsu, Masaki Takashima, Local structure analysis of Ge-Sb-Te phase change materials using high-resolution electron microscopy and nanobeam diffraction, *J. Appl. Phys.* 95 (2004) 8130–8135.
- [51] Koji Usuda, Toshinori Numata, Toshifumi Irisawa, Norio Hirashita, Shinichi Takagi, Strain characterization in SOI and strained-Si on SGOI MOSFET channel using nano-beam electron diffraction (NBD), *Mater. Sci. Eng.: B* 124 (2005) 143–147.
- [52] Akio Toda, Hidetatsu Nakamura, Toshinori Fukai, Nobuyuki Ikarashi, Channel strain in advanced complementary metal-oxide-semiconductor field effect transistors measured using nano-beam electron diffraction, *Japan. J. Appl. Phys.* 47 (2008) 2496.
- [53] A. Béché, J.L. Rouvière, L. Clément, J.M. Hartmann, Improved precision in strain measurement using nanobeam electron diffraction, *Appl. Phys. Lett.* 95 (2009).
- [54] V.B. Ozdol, C. Gammer, X.G. Jin, P. Ercius, C. Ophus, J. Ciston, A.M. Minor, Strain mapping at nanometer resolution using advanced nano-beam electron diffraction, *Appl. Phys. Lett.* 106 (2015).
- [55] Colin Ophus, Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond, *Microsc. Microanal.* 25 (2019) 563–582.
- [56] Yimo Han, Kayla Nguyen, Michael Cao, Paul Cueva, Saien Xie, Mark W. Tate, Prafull Purohit, Sol M. Gruner, Jiwoong Park, David A. Muller, Strain mapping of two-dimensional heterostructures with subpicometer precision, *Nano Lett.* 18 (2018) 3746–3751.
- [57] Knut Müller, Andreas Rosenauer, Marco Schowalter, Josef Zweck, Rafael Fritz, Kerstin Volz, Strain measurement in semiconductor heterostructures by scanning transmission electron microscopy, *Microsc. Microanal.* 18 (2012) 995–1009.
- [58] Thomas C. Pekin, Christoph Gammer, Jim Ciston, Andrew M. Minor, Colin Ophus, Optimizing disk registration algorithms for nanobeam electron diffraction strain mapping, *Ultramicroscopy* 176 (2017) 170–176.
- [59] Sihan Wang, Tim B. Eldred, Jacob G. Smith, Wenpei Gao, AutoDisk: Automated diffraction processing and strain mapping in 4D-STEM, *Ultramicroscopy* 236 (2022) 113513.
- [60] Christoph Mahr, Knut Müller-Caspary, Tim Grieb, Marco Schowalter, Thorsten Mehrtens, Florian F. Krause, Dennis Zillmann, Andreas Rosenauer, Theoretical study of precision and accuracy of strain analysis by nano-beam electron diffraction, *Ultramicroscopy* 158 (2015) 38–48.
- [61] Tim Grieb, Florian F. Krause, Marco Schowalter, Dennis Zillmann, Roman Sellin, Knut Müller-Caspary, Christoph Mahr, Thorsten Mehrtens, Dieter Bimberg, Andreas Rosenauer, Strain analysis from nano-beam electron diffraction: Influence of specimen tilt and beam convergence, *Ultramicroscopy* 190 (2018) 45–57.
- [62] Giulio Guzzinati, Wannes Ghielens, Christoph Mahr, Armand Béché, Andreas Rosenauer, Toon Calders, Jo Verbeeck, Electron beam diffraction for precise and accurate nanoscale strain mapping, *Appl. Phys. Lett.* 114 (2019).
- [63] Steven E. Zeltmann, Alexander Müller, Karen C. Bustillo, Benjamin Savitzky, Lauren Hughes, Andrew M. Minor, Colin Ophus, Patterned probes for high precision 4D-STEM bragg measurements, *Ultramicroscopy* 209 (2020) 112890.
- [64] Christoph Mahr, Knut Müller-Caspary, Tim Grieb, Florian F. Krause, Marco Schowalter, Andreas Rosenauer, Accurate measurement of strain at interfaces in 4D-STEM: A comparison of various methods, *Ultramicroscopy* 221 (2021) 113196.
- [65] James M. LeBeau, Scott D. Findlay, Leslie J. Allen, Susanne Stemmer, Position averaged convergent beam electron diffraction: Theory and applications, *Ultramicroscopy* 110 (2010) 118–125.
- [66] J.A. Pollock, M. Weyland, D.J. Taplin, L.J. Allen, S.D. Findlay, Accuracy and precision of thickness determination from position-averaged convergent beam electron diffraction patterns using a single-parameter metric, *Ultramicroscopy* 181 (2017) 86–96.
- [67] James M. LeBeau, Adrian J. D'Alfonso, Nicholas J. Wright, Leslie J. Allen, Susanne Stemmer, Determining ferroelectric polarity at the nanoscale, *Appl. Phys. Lett.* 98 (2011).
- [68] Jinwoo Hwang, Jack Y. Zhang, Junwoo Son, Susanne Stemmer, Nanoscale quantification of octahedral tilts in perovskite films, *Appl. Phys. Lett.* 100 (2012).
- [69] Jack Y. Zhang, Jinwoo Hwang, Santosh Raghavan, Susanne Stemmer, Symmetry lowering in extreme-electron-density perovskite quantum wells, *Phys. Rev. Lett.* 110 (2013) 256401.
- [70] Colin Ophus, Peter Ercius, Mark Huijben, Jim Ciston, Non-spectroscopic composition measurements of SrTiO₃-La_{0.7}Sr_{0.3}MnO₃ multilayers using scanning convergent beam electron diffraction, *Appl. Phys. Lett.* 110 (2017).
- [71] Sergei V. Kalinin, Colin Ophus, Paul M. Voyles, Rolf Erni, Demie Kepaptsoglou, Vincenzo Grillo, Andrew R. Lupini, Mark P. Oxley, Eric Schwenker, Maria K.Y. Chan, et al., Machine learning in scanning transmission electron microscopy, *Nat. Rev. Methods Prim.* 2 (2022) 11.
- [72] Kevin P. Treder, Chen Huang, Judy S. Kim, Angus I. Kirkland, Applications of deep learning in electron microscopy, *Microscopy* 71 (2022) i100–i115.
- [73] W. Xu, J.M. LeBeau, A deep convolutional neural network to analyze position averaged convergent beam electron diffraction patterns, *Ultramicroscopy* 188 (2018) 59–69.
- [74] C. Zhang, J. Feng, L.R. DaCosta, P.M. Voyles, Atomic resolution convergent beam electron diffraction analysis using convolutional neural networks, *Ultramicroscopy* 210 (2020) 112921.
- [75] Young-Hoon Kim, Sang-Hyeok Yang, Myoung-Ho Jeong, Min-Hyoung Jung, Dae-hee Yang, Hyangsook Lee, Taehwan Moon, Jinseong Heo, Hu Young Jeong, Eunha Lee, et al., Hybrid deep learning crystallographic mapping of polymorphic phases in polycrystalline Hf_{0.5}Zr_{0.5}O₂ thin films, *Small* 18 (2022) 2107620.
- [76] Renliang Yuan, Jiong Zhang, Lingfeng He, Jian-Min Zuo, Training artificial neural networks for precision orientation and strain mapping using 4D electron diffraction datasets, *Ultramicroscopy* 231 (2021) 113256.
- [77] Chuqiao Shi, Michael C. Cao, Sarah M. Rehn, Sang-Hoon Bae, Jeehwan Kim, Matthew R. Jones, David A. Muller, Yimo Han, Uncovering material deformations via machine learning combined with four-dimensional scanning transmission electron microscopy, *Npj Comput. Mater.* 8 (2022) 114.
- [78] Joydeep Munshi, Alexander Rakowski, Benjamin H. Savitzky, Steven E. Zeltmann, Jim Ciston, Matthew Henderson, Shreyas Cholia, Andrew M. Minor, Maria K.Y. Chan, Colin Ophus, Disentangling multiple scattering with deep learning: application to strain mapping from electron diffraction patterns, *Npj Comput. Mater.* 8 (2022) 254.
- [79] C. Ophus, A fast image simulation algorithm for scanning transmission electron microscopy, *Adv. Struct. Chem. Imaging* 3 (2017) 13.
- [80] A. Pryor, C. Ophus, J. Miao, A streaming multi-GPU implementation of image simulation algorithms for scanning transmission electron microscopy, *Adv. Struct. Chem. Imaging* 3 (2017) 15.
- [81] L.R. DaCosta, H.G. Brown, P.M. Pelz, A. Rakowski, N. Barber, P. O'Donovan, P. McBean, L. Jones, J. Ciston, M.C. Scott, C. Ophus, Prismatic 2.0 – simulation software for scanning and high resolution transmission electron microscopy (STEM and HRTEM), *Micron* 151 (2021) 103141.
- [82] D.-A. Clevert, T. Unterthiner, S. Hochreiter, Fast and accurate deep network learning by exponential linear units (elus), 2016, arXiv preprint arXiv:1511.07289.
- [83] J. Schmidhuber, Deep learning in neural networks: an overview, *Neural Netw.* 61 (2015) 85–117.
- [84] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: A comprehensive review, *Neural Comput.* 29 (2017) 2352–2449.
- [85] N. Qian, On the momentum term in gradient descent learning algorithms, *Neural Netw.* 12 (1999) 145–151.
- [86] L. Bottou, Large-scale machine learning with stochastic gradient descent, in: *Proceedings of COMPSTAT'2010*, Springer, 2010, pp. 177–186.
- [87] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015, Software available from tensorflow.org, URL <https://www.tensorflow.org/>.