



OPEN

# Deep reinforcement learning for data-driven adaptive scanning in ptychography

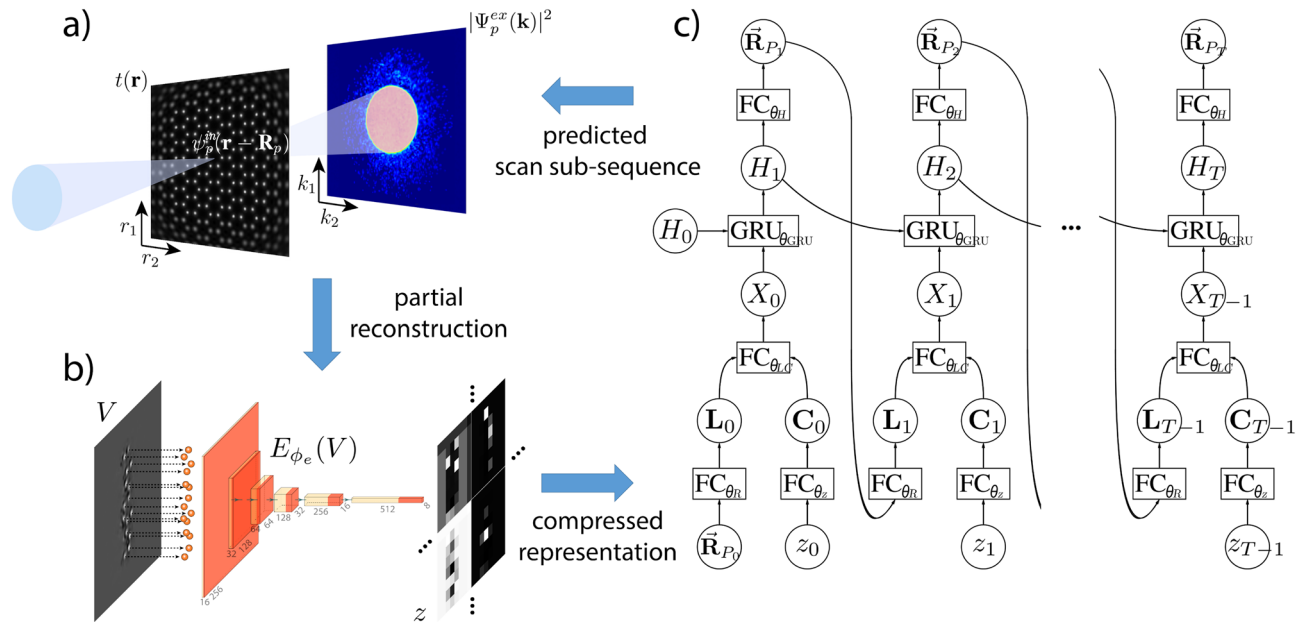
Marcel Schloz<sup>1✉</sup>, Johannes Müller<sup>1</sup>, Thomas C. Pekin<sup>1</sup>, Wouter Van den Broek<sup>1</sup>, Jacob Madsen<sup>2</sup>, Toma Susi<sup>2</sup> & Christoph T. Koch<sup>1</sup>

We present a method that lowers the dose required for an electron ptychographic reconstruction by adaptively scanning the specimen, thereby providing the required spatial information redundancy in the regions of highest importance. The proposed method is built upon a deep learning model that is trained by reinforcement learning, using prior knowledge of the specimen structure from training data sets. We show that using adaptive scanning for electron ptychography outperforms alternative low-dose ptychography experiments in terms of reconstruction resolution and quality.

Ptychography is a coherent diffractive imaging (CDI) method that has found use in light, X-ray and scanning transmission electron microscopies (STEM). The method combines diffraction patterns from spatially overlapping regions to reconstruct the structure of a specimen for arbitrarily large fields of view<sup>1</sup>, with many advantages over other imaging methods<sup>2–5</sup>. The development of new hardware<sup>6,7</sup> and reconstruction algorithms<sup>8,9</sup> has led to ptychography becoming a mature electron microscopy technique<sup>4</sup>. Current research to further improve it is driven by the desire to investigate thick samples<sup>10–14</sup> as well as to lower the required electron dose<sup>15–18</sup>. In order to lower the dose, researchers have tried to vary various experimental parameters while preserving information redundancy through overlapping probes. One approach involves a defocused probe rastered across the specimen with a less dense scan pattern. This uses a lower dose than focused probe ptychography, but introduces additional complications for the reconstruction algorithm due to an increased need to account for partial spatial coherence in the illuminating probe<sup>18</sup>. Another approach is simply to scan faster by lowering the dwell time per probe position, an overall decrease in dose can be realized. However, this comes with its own challenges, as the physical limits of the electron source, microscope, and camera all must be considered. Finally, a third approach is the optimization of the scan pattern, deviating from a raster grid in favour of a generally more efficient pattern<sup>19</sup>. This approach can, however, only yield a limited improvement in reconstruction quality as it is not capable of taking into account the structure of the specimen in the scan pattern.

In this work we present an approach particularly tailored for electron ptychography that enables reduction of the electron dose through adaptive scanning. It is based upon the idea that, at atomic resolution, ptychography requires an increased information redundancy through overlapping illuminating beams only in regions that contain the atomic structure of the scanned specimen. We present here a workflow that scans only the regions with the highest information content in order to strongly improve the ptychographic reconstruction quality while keeping low the total number of scan positions, and therefore the total dose. The scan positions are predicted sequentially during the experiment and the only information required for the prediction process is the diffraction data acquired at previous scan positions. The scan position prediction model of the workflow is a mixture of deep learning models, and the model training is performed with both supervised and reinforcement learning (RL). A schematic of the workflow is given in Fig. 1. The synergy of deep learning and RL has already shown strong performance in various dynamic decision-making problems, such as tasks in robotics<sup>20,21</sup> or visual recognition<sup>22–24</sup>. The success of this approach, despite the complexity of the problems to overcome, can be attributed to the algorithms' ability of learning independently from data. Similarly, the proposed algorithm here solves a sequential decision-making problem by learning from a large amount of simulated or, if available, experimental ptychographic data consisting of hundreds to thousands of diffraction patterns. Here, the focus of the learning is specifically designed to maximize the dynamic range in the reconstruction for each individual scan position. The algorithm then transfers the learned behaviour it developed offline to a realistic experimental environment.

<sup>1</sup>Institute of Physics and IRIS Adlershof, Humboldt Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany. <sup>2</sup>Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria. ✉email: schlozma@hu-berlin.de

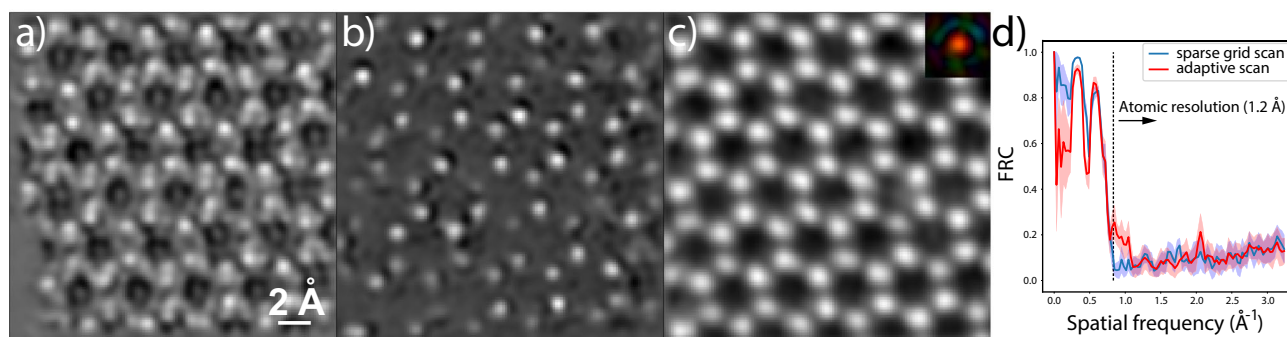


**Figure 1.** Schematic of the adaptive scanning workflow with its three main components. **(a)** Experimental data acquisition in ptychography. At the scan position  $\mathbf{R}_p$  of the scan sub-sequence  $\vec{\mathbf{R}}_{P_t}$ , the beam illuminates a sample, where the incident electron wave  $\psi_p^{in}(\mathbf{r} - \mathbf{R}_p)$  interacts with the transmission function  $t(\mathbf{r})$ . The wave exiting the sample is propagated by a Fourier transform to the detector located in the far field and the intensity  $I_p = |\Psi_p^{ex}(\mathbf{k})|^2$  is recorded as a diffraction pattern. **(b)** A reconstruction  $V$  generated from diffraction patterns of a scan sub-sequence is mapped to the compressed representation  $z$  by using an encoder network  $E_{\phi_e}(V)$ . **(c)** Schematic of the forward propagation process of the RNN model. The RNN consists of GRU units that use the hidden state  $H_t$  from the previous time step and the hybrid input information  $X_t$  to create a new hidden state  $H_{t+1}$ . The hybrid input is the concatenation of the pre-processed information from the sub-sequence of scan positions  $\vec{\mathbf{R}}_{P_t}$  and the corresponding compressed representation of the partial reconstruction  $z_t$ . The output of the GRU cell is used to predict the positions of the next sub-sequence  $\mathbf{R}_{P_{t+1}}$  and is also used as the input for the next GRU cell.

Our approach is conceptually related to the subfield of computer vision that focuses on identifying relevant regions of images or video sequences for the purpose of classification or recognition. However, there are fundamental differences not only in the purpose, but also in the solution strategy for our application in contrast to computer vision tasks. Differences include a lack of direct access to images (updated real space information is only accessible through a highly optimized reconstruction algorithm), non-optimal parameter settings of the reconstruction algorithm and experimental uncertainties such as imprecise scan positioning of the microscope or contamination of the specimen requiring pre-processing of the reconstructed image, and the necessity of a much larger number of measurements requiring methods that improve the performance of the sequential decision making process. Work in adaptive scanning for X-ray fluorescence imaging<sup>25</sup> and for scanning probe microscopy<sup>26</sup> has also recently been reported. However, the work in Ref.<sup>25</sup> is more closely related to previous work in scanning electron microscopy that divides the measurement into a low-dose raster scan and a subsequent high-dose adaptive scan<sup>27</sup>. For the latter work in Ref.<sup>26</sup>, it has been reported that its model suffers in performance as it lacks prior knowledge of the domain structure, which can be compensated by including a deep learning model with domain specific knowledge. Our proposed approach is the first application of adaptive scanning to ptychography, and is further unique in that the scan pattern is predicted using prior knowledge about the sample in the form of a pre-trained deep neural network, thereby improving performance.

## Results

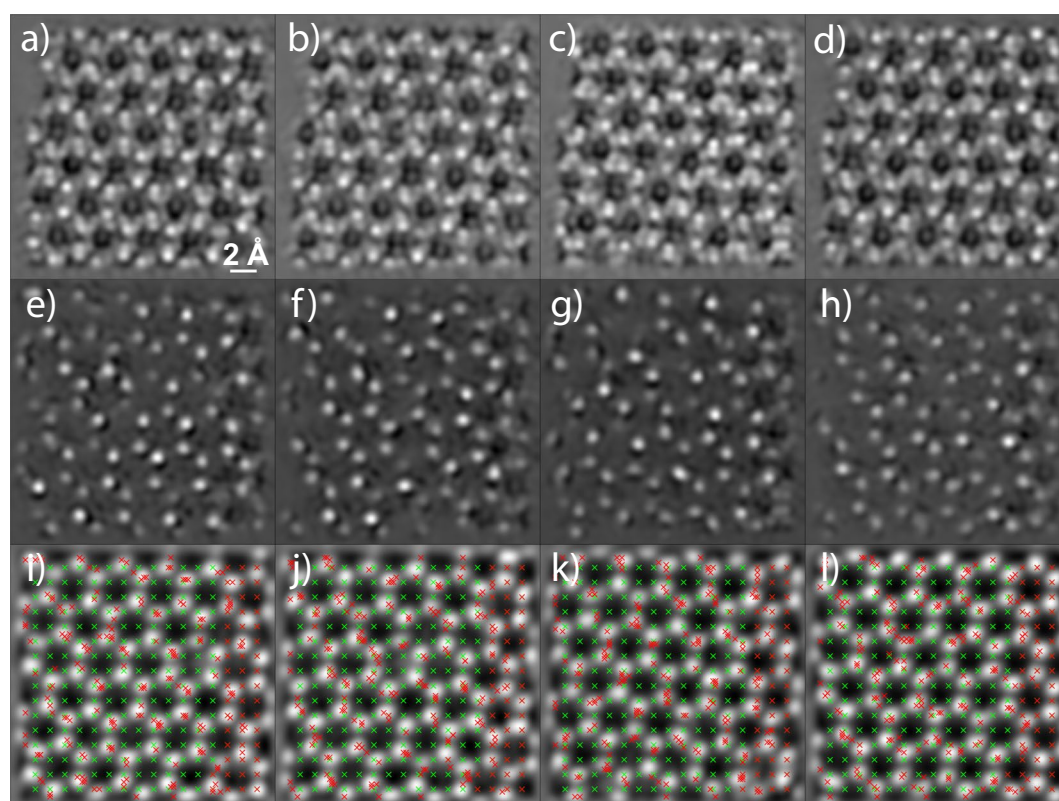
The result of adaptive scanning on experimentally acquired MoS<sub>2</sub> data and its comparison to the result of a sparse grid scanning and the conventional grid scanning procedure is shown in Fig. 2. The data used for the comparison was not part of the training data for the adaptive scanning model. However, the entire data was acquired from the same sample and includes multiple data sets that were recorded from different regions of the sample. This data can be found in Ref.<sup>28</sup>. In our comparison, a ground truth reconstruction is obtained from one of the data sets each consisting of 10,000 diffraction patterns, while only 250 diffraction patterns have been used for the adaptive scanning as well as the sparse grid scanning reconstruction. Figure 2a shows the ptychographic reconstruction when using a sparse grid scanning procedure. The structure of the material is not clearly resolved and/or shows ambiguous features. Figure 2b shows the reconstruction when the scan positions are predicted through adaptive scanning. Although without the same homogeneous reconstruction quality throughout the entire field of view, the structure of the MoS<sub>2</sub> material is now much better resolved and is closer to the ground truth reconstruction of the full data grid scanning procedure, shown in Fig. 2c.



**Figure 2.** Ptychographic reconstruction of a MoS<sub>2</sub> data set with a scanning procedure that follows (a) a sparse grid scan, (b) an adaptive scan and (c) the conventional grid scan. While only 250 diffraction patterns are used in (a) and (b), the full data set with 10,000 diffraction patterns is used in (c). The inset in (c) illustrates the illumination probe used for the reconstruction with an estimated size of 0.93 Å. The pixel size is identical to the one used in the reconstruction and its magnitude is represented by the intensity, and the phase is represented by the HSV colorwheel. (d) FRC of the sparse grid scan and the adaptive scan averaged over 25 data sets and where the standard deviation is illustrated by the shaded area.

Further examples of reconstructions and their corresponding scan sequences are shown in Fig. 3. The results suggest that probe delocalization due to scattering plays an important role as to why an improved ptychographic reconstruction can be achieved by distributing the scan positions predominantly on the atoms of the specimen. This implies that similar results could be achieved by using RL with a reward function that specifically emphasizes the scattered electrons in the recorded diffraction patterns, which is an interesting area for future research.

The final point of our experimental investigation into adaptive scanning in ptychography evaluates the performance of the method for various prediction settings. We compare the Fourier ring correlation (FRC)<sup>29</sup> as well as



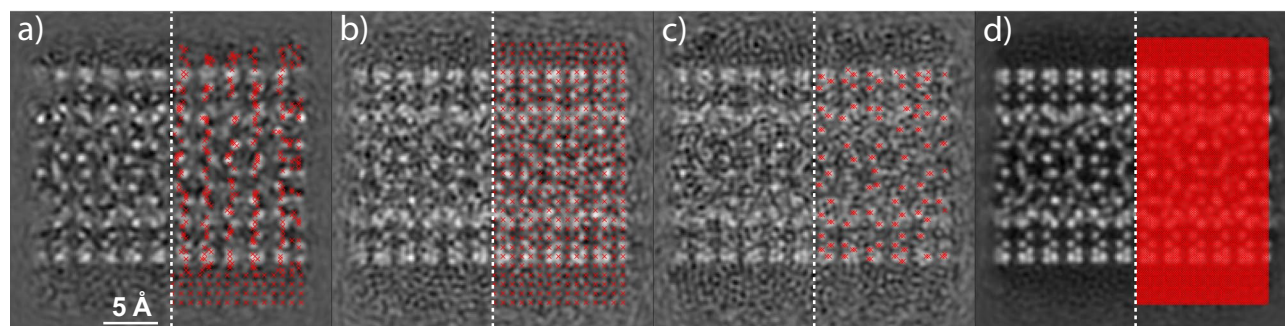
**Figure 3.** Ptychographic reconstructions of different MoS<sub>2</sub> data sets and with different scanning procedures. Reconstruction from 250 diffraction patterns of a data set that correspond to scan positions which follow (a–d) a sparse grid scanning sequence and (e–h) an adaptively predicted sequence. (i–l) Ground truth reconstruction of the full data set with 10,000 diffraction patterns shown with the scan positions used for the corresponding reconstructions (a–d) in green and (e–h) in red.

the structural similarity index measure (SSIM)<sup>30</sup> between the reconstruction obtained from the reduced data and the ground truth reconstruction obtained from the full data to quantify the improvement of the effective image resolution and image quality when using adaptive scanning. In the first comparison, shown in Fig. 2d, we apply the FRC to the sparse grid scan and adaptive scan averaged over 25 data sets, respectively. For both cases, there is a sharp frequency cut off in the proximity of atomic resolution (1.2 Å). However, while the cross-correlation value almost disappears in the case of the sparse grid scan, it plateaus at a value of about 0.2 in the case of the adaptive scan. This result indicates the benefit in terms of achievable resolution of adaptive scanning in contrast to other low dose alternatives. In the latter comparison, SSIM<sub>a</sub> and SSIM<sub>s</sub> refer to reconstructions of reduced data obtained with the adaptive scanning and the sparse grid scanning procedure, respectively. Table 1 shows the relative reconstruction quality improvement  $Q_{\text{SSIM}} = (\text{SSIM}_a - \text{SSIM}_s) / (\text{SSIM}_s)$  for different experimental settings averaged over the same 25 data sets as used before. In the case of 250 scan positions, which corresponds to a dose reduction by a factor of 40 with respect to the original data, tests were performed for different total numbers  $T$  of sub-sequences and therefore different amounts of scan positions included in each sub-sequence  $\bar{R}_p$ . The quality improvement ranges from 9.89% to 15.84% for 2 to 5 sub-sequences, respectively. Note that the scan positions of the first sub-sequence  $\bar{R}_p$ , provided to the RNN as part of the initial input, follow the sparse grid sequence and that the scan positions of each sub-sequence only cover a part of the sample, as can be seen in Fig. 6b. Further tests were performed using a larger number of total scan positions and 5 sub-sequences. However, the difference in quality between the reconstruction generated with the positions of the adaptive scan and the sparse grid scan decreases with the total number of positions used, as can be expected, since the sparse grid sampling covers the sampled area in an increasingly complete manner. These results indicate that the reconstruction quality improves with the frequency by which the positions are predicted, and that low dose experiments benefit the most from the adaptive scanning scheme.

In Fig. 4, we compare the results of various scanning procedures using simulated double-walled carbon nanotube (DWCNT) data. This data is publicly available<sup>28</sup>. Figure 4a shows a ptychography reconstruction that considers 840 diffraction patterns that have been selected through the adaptive scanning procedure. Most of the unit cells of the structure can be resolved with a high contrast and therefore the configuration of the DWCNT can be easily deduced. The predicted scan positions coincide to a high degree with the structure of the DWCNT. Note that the initial scan sub-sequence visualized at the bottom of the reconstruction follows a sparse grid scan sequence. Figure 4b shows the reconstruction when using 840 diffraction patterns obtained from a sparse grid

# Pos.	$N_k/N_u$	Dose ( $e^-/\text{Å}^{-2}$ )	# Sub-seq.	$Q_{\text{SSIM}}$
250	8.21	1.34E5	2	9.89 ± 6.80%
			3	12.38 ± 10.72%
			4	15.71 ± 7.60%
			5	15.84 ± 7.23%
335	10.83	1.79e5	5	10.03 ± 8.17%
420	13.43	2.25e5	5	8.60 ± 5.30%
500	15.86	2.68e5	5	8.08 ± 7.60%

**Table 1.** Performance of adaptive scanning for various experimental settings that differ in the number of scan positions and the total number of sub-sequences. For each setting, the oversampling ratio  $N_k/N_u$ , which is calculated following Ref.<sup>17</sup>, and the electron dose is given.



**Figure 4.** Ptychographic reconstruction of a DWCNT data set with a scanning procedure that follows (a) an adaptive scan, (b) a sparse grid scan, (c) an alternative low-dose scan and (d) the conventional grid scan. 840 diffraction patterns have been used for the reconstructions in (a) and (b), 13,536 patterns were used for the reconstruction in (c) and the full data set consisting of 13,225 patterns was used for the reconstruction in (d). Half of the corresponding scan positions are illustrated on the right hand side of each reconstruction. Note that the alternative low-dose scan method used for generating (c) consists of a full grid scan at a very low dose and a consecutive scan using 311 scan positions that have been found to match the sample structure based on the result obtained from the previous scan. Only the latter scan positions are visualized in (c).

scanning procedure. The field of view is now much better covered with scan positions, but the periodicity at which the scan positions are spaced seems to be also present in the reconstruction. Hence, the reconstruction shows ambiguous features that make the interpretation of the structure more difficult compared to the previous case. Figure 4c shows the reconstruction of an alternative low-dose scanning procedure which has been described conceptually in Ref.<sup>27</sup>. Here, two consecutive scans are performed. The first scan is a conventional dense grid scan consisting of 13,225 diffraction patterns with an electron dose of  $4e3 e^-/\text{\AA}^{-2}$  compared to  $1e5 e^-/\text{\AA}^{-2}$ . The same scan with the latter dose has also been used for the dense grid scan in Fig. 4d. The second scan of the alternative low-dose scanning procedure was limited to 311 scan positions as to match the total electron dose of the procedures in Fig. 4a and b. An atom finding method was applied to the ptychography reconstruction generated after the first scan to adapt these 311 scan positions to the atomic structure of the DWCNT specimen. Although the predicted scan positions in this approach match the atomic structure almost perfectly, their contribution to the final reconstruction seems to be not sufficient given that most of the total electron dose is required for their optimal prediction. Quantitatively, we obtain a relative reconstruction quality improvement  $Q_{\text{SSIM}}$  of  $2.60 \pm 3.38\%$  and  $16.97 \pm 3.49\%$  when using adaptive scanning with respect to the sparse grid scanning and alternative low-dose scanning procedure, respectively.

## Conclusion

We have presented a method for electron ptychography that reduces the electron dose through adaptive scanning. Sub-sequences of scan positions are predicted by the model within milliseconds, allowing an acquisition rate that theoretically exceeds rates currently achieved in 4D-STEM experiments. The method would therefore have the potential to be applied in real time at the microscope if the used ptychography reconstruction algorithm could generate images sufficiently fast. Future work does therefore require either an improvement of iterative ptychography algorithms in terms of processing speed or the integration of direct ptychography reconstruction methods, such as single-sideband (SSB) ptychography<sup>31</sup>, in the adaptive scanning workflow. We show an improved resolution and reconstruction quality when using an adaptive scanning approach on experimentally acquired monolayer MoS<sub>2</sub> data sets in comparison to another dose reduction scanning approach. These improvements show that adaptive scanning for ptychography is a useful technique to lower the dose needed for the analysis of sensitive samples. It can be provided with simulated or generic experimental training data<sup>32,33</sup> to increase its applicability to a variety of different or less periodic material structures. We have demonstrated the generalizability of our method by applying it to simulated DWCNT data sets and showing that it outperforms other low-dose alternatives. In addition, the proposed workflow can be taken as a blueprint for a broad range of scanning microscopy methods and thus paves the way for future research in machine learning supported, automated and autonomous microscopy<sup>34,35</sup>.

## Methods

**Image formation in ptychography.** The approach described in this paper is compatible with multislice ptychography, but in light of the application to a 2D material we constrain ourselves to single-slice ptychography. Here, ptychography can be expressed by a multiplicative approximation that describes the interaction of a wavefunction  $\psi_p^{\text{in}}(\mathbf{r})$  of an incoming beam with the transmission function  $t(\mathbf{r})$  of a specimen. For each measurement  $p$ , the beam is shifted by  $\mathbf{R}_p$  and a diffraction pattern is acquired with the intensity  $I_p$  that is expressed by

$$I_p = |\Psi_p^{\text{ex}}(\mathbf{k})|^2 = |\mathcal{F} [\psi_p^{\text{in}}(\mathbf{r} - \mathbf{R}_p)t(\mathbf{r})]|^2, \quad (1)$$

where  $\mathcal{F}$  is the Fourier propagator,  $\mathbf{r}$  the real space coordinate,  $\mathbf{k}$  the reciprocal space coordinate and  $\Psi_p^{\text{ex}}(\mathbf{k})$  the exit wavefunction at the detector. The transmission function can be defined as  $t(\mathbf{r}) = e^{i\sigma V(\mathbf{r})}$ , with the interaction constant  $\sigma$  and the complex potential  $V(\mathbf{r})$ . Throughout this treatment,  $\sigma$  is absorbed into  $V(\mathbf{r})$ . X-ray and optical ptychography is mathematically described similarly with the only difference that the transmission function  $t(\mathbf{r})$  is related to the complex refractive index of the specimen. Figure 1a illustrates the experimental configuration of conventional ptychography. The potential of the specimen is recovered from data of experimentally acquired diffraction patterns  $J_p$  using a reconstruction algorithm. Here, we apply a gradient based algorithm<sup>17</sup> with a gradient decent optimization and the potential is retrieved by iteratively minimizing the loss function

$$\mathcal{L}(V) = \frac{1}{P} \sum_{p=1}^P \|I_p(V) - J_p\|_2^2. \quad (2)$$

**Generation of scan sequences.** We use a recurrent neural network (RNN)<sup>36–38</sup> for the generation of scan sequences. Its network architecture is designed to model temporal sequences with recurring input information. Memory cells combine the current input information  $X_t$  with the hidden state  $H_t$  and map it to the next hidden state  $H_{t+1}$ . These hidden states represent the memory gathered from all the previous time steps. Gated recurrent units (GRU)<sup>39</sup>, which allow a computationally fast mapping with a high performance, are used in this work. At every time step  $t$ , an output is generated on the basis of the current hidden state. In the implementation shown here, the output corresponds to multiple scan positions, i.e. a sub-sequence of scan positions, given by a vector of 2D coordinates  $\mathbf{R}_p$ . In principle, the output can be reduced to a single scan position  $\mathbf{R}_p$ , but we do not do so for practical reasons that involve a reduced training performance of the network and also a greatly increased acquisition time due to, e.g., more frequent data transfer and pre-processing of these intermediate data chunks. The sub-sequence is predicted via a fully connected layer (FC) that is parameterized by the layer weights  $\theta_H$ :

$$\tilde{\mathbf{R}}_{P_t} = \text{FC}_{\theta_H}(H_t). \quad (3)$$

At the predicted scan positions  $\tilde{\mathbf{R}}_{P_t}$ , diffraction patterns  $J_{P_t}$  are acquired by the microscope and from these diffraction patterns a potential  $V_t(\mathbf{r})$  is reconstructed minimizing Eq. (2). The intermediate reconstruction  $V_t(\mathbf{r})$  combined with its corresponding sub-sequence of scan positions  $\tilde{\mathbf{R}}_{P_t}$  can then be used for the input information  $X_t$  of the RNN. However, the bandwidth of the information given in  $V_t(\mathbf{r})$  and  $\tilde{\mathbf{R}}_{P_t}$  differs strongly and thus processing is required before the two components can be concatenated and mapped to  $X_t$ . For the processed location information  $\mathbf{L}_t$  based on the sub-sequence  $\tilde{\mathbf{R}}_{P_t}$ , a FC that is parameterized by the weights  $\theta_R$  is used:

$$\mathbf{L}_t = \text{FC}_{\theta_R}(\tilde{\mathbf{R}}_{P_t}). \quad (4)$$

For the processed structure information  $\mathbf{C}_t$  based on the reconstructed potential  $V_t(\mathbf{r})$ , a compressed representation  $z_t$  is generated by using the encoder part of a convolutional autoencoder<sup>40</sup>. This processing step is shown in Fig. 1b and the training of the convolutional autoencoder is described in the Supplementary Information. The compressed representation  $z_t$  is then fed into a FC that is parameterized by the weights  $\theta_z$ :

$$\mathbf{C}_t = \text{FC}_{\theta_z}(z_t). \quad (5)$$

The processed location information  $\mathbf{L}_t$  is subsequently concatenated with the processed structure information  $\mathbf{C}_t$  and mapped to the input information  $X_t$  with a FC that is parameterized by the weights  $\theta_{LC}$ . The whole process of predicting sub-sequences of scan positions and acquiring the corresponding diffraction patterns is repeated until a ptychographic data set of desired size is reached. Finally, backpropagation through time (BPTT) is used to generate the required gradients to update the network weights  $\theta = \{\theta_H, \theta_{GRU}, \theta_{LC}, \theta_R, \theta_z\}$  of the RNN. Figure 1c shows the prediction process modeled by the RNN in full detail.

**Training through reinforcement learning.** A RNN can be combined with RL to provide a formalism for modelling behaviour to solve decision making problems. In the case of adaptive scanning in ptychography, where we seek to predict multiple scan positions at each time step, RL can be formalized by a partially observable stochastic game (POSG) that is described by a 8-tuple,  $(M, \mathcal{S}, \{\mathcal{A}^m\}_{m \in M}, \rho, \{r^m\}_{m \in M}, \{\mathcal{O}^m\}_{m \in M}, \omega, \gamma)$ , with multiple agents  $M$ . At each time-step  $t$  an agent  $m$  selects an action  $a_t^m \in \mathcal{A}^m$  and makes an observation  $o_t^m \in \mathcal{O}^m$  given the state  $s_t \in \mathcal{S}$ . Thus, joint actions  $\mathbf{a}_t = \langle a_t^1, \dots, a_t^M \rangle$  from the joint action space  $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_M$  are executed and joint observations  $\mathbf{o}_t = \langle o_t^1, \dots, o_t^M \rangle$  from the joint observation space  $\mathcal{O} = \mathcal{O}_1 \times \dots \times \mathcal{O}_M$  are received from the environment at every time step. The next state  $s_{t+1}$  is generated according to a transition function  $\rho: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ , the observations  $\mathbf{o}_{t+1}$ , containing incomplete information about the state  $s_{t+1}$ , are generated according to an observation function  $\omega: \mathcal{S} \times \mathcal{S} \times \mathcal{O} \rightarrow [0, 1]$  and each agent receives its immediate reward defined by the reward function  $r^m: \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ . This reward  $r^m$  contributes to the total reward computed at the end of the sequence,  $G^m = \sum_{t=0}^T \gamma^t r^m(\mathbf{a}_t, s_t)$ , also known as the return. The discount factor  $\gamma \in [0, 1]$  controls the emphasis of long-term rewards versus short-term rewards. The entire history of observations and actions up to the current time  $\mathbf{h}_t = \langle \mathbf{o}_1, \mathbf{a}_1, \dots, \mathbf{o}_{t-1}, \mathbf{a}_{t-1}, \mathbf{o}_t \rangle$  is used as basis for optimal or near-optimal decision making. A stochastic policy  $\pi_{\theta^m}(a_t^m | \mathbf{h}_t)$  maps the history of past interactions  $\mathbf{h}_t$  to action probabilities. Given a continuous action space, the policy can be represented by a two-dimensional Gaussian probability distribution:

$$\pi_{\theta^m}(a_t^m | \mathbf{h}_t) = \mathcal{N}(\boldsymbol{\mu}_{\theta^m}(\mathbf{h}_t), \Sigma), \quad (6)$$

with its mean vector  $\boldsymbol{\mu}_{\theta^m}(\mathbf{h}_t)$  corresponding to  $\mathbf{R}_{P_t}$ , where the history  $\mathbf{h}_t$  is summarized in the hidden state  $H_t$  of the RNN and the covariance matrix  $\Sigma$  with fixed variances  $\sigma_x^2 \in [0, 1]$  and  $\sigma_y^2 \in [0, 1]$ . The joint policy of all agents is then defined as  $\boldsymbol{\pi}_{\theta}(\mathbf{a}_t | \mathbf{h}_t) = \prod_{m=1}^M \pi_{\theta^m}(a_t^m | \mathbf{h}_t)$ , with  $\theta = \{\theta^m\}_{m \in M}$ . The goal of RL is now to learn a joint policy that maximizes the expected total reward for each agent  $m$  with respect to its parameters  $\theta^m$ :

$$\mathcal{J}^m(\theta) = \mathbb{E}_{\boldsymbol{\pi}_{\theta}(\boldsymbol{\tau})}[G^m] \approx \frac{1}{N} \sum_{n=1}^N G_n^m, \quad (7)$$

where the expected total reward can be approximated by Monte Carlo sampling with  $N$  samples. In this paper, improvement of the policy is achieved by updating the policy parameters  $\theta^m = \{\theta_H^m, \theta_{GRU}^m, \theta_{LC}^m, \theta_R^m, \theta_z^m\}$  with 'REINFORCE'<sup>41</sup>, a policy gradient method:

$$\begin{aligned} \nabla_{\theta^m} \mathcal{J}^m(\theta) &= \mathbb{E}_{\boldsymbol{\pi}_{\theta}(\boldsymbol{\tau})}[\nabla_{\theta^m} \log \boldsymbol{\pi}_{\theta}(\boldsymbol{\tau}) G^m] \\ &\approx \frac{1}{N} \sum_{n=1}^N \sum_{t=0}^T \nabla_{\theta^m} \log \pi_{\theta^m}(a_{n,t}^m | \mathbf{h}_{n,t}) G_n^m. \end{aligned} \quad (8)$$

The derivation of  $\nabla_{\theta^m} \mathcal{J}^m(\theta)$  is given in the Supplementary Information.

**Learning to adaptively scan in ptychography.** While policy gradient methods are the preferred choice to solve reinforcement learning problems in which the action spaces are continuous<sup>42</sup>, they come with significant problems. Like any gradient based method, policy gradient solutions mainly converge to local, not global, optima<sup>43</sup>. In this paper, we reduce the effect of this problem during training by splitting the training of the RNN into supervised learning and RL. While training in RL can be performed with a policy whose parameters are

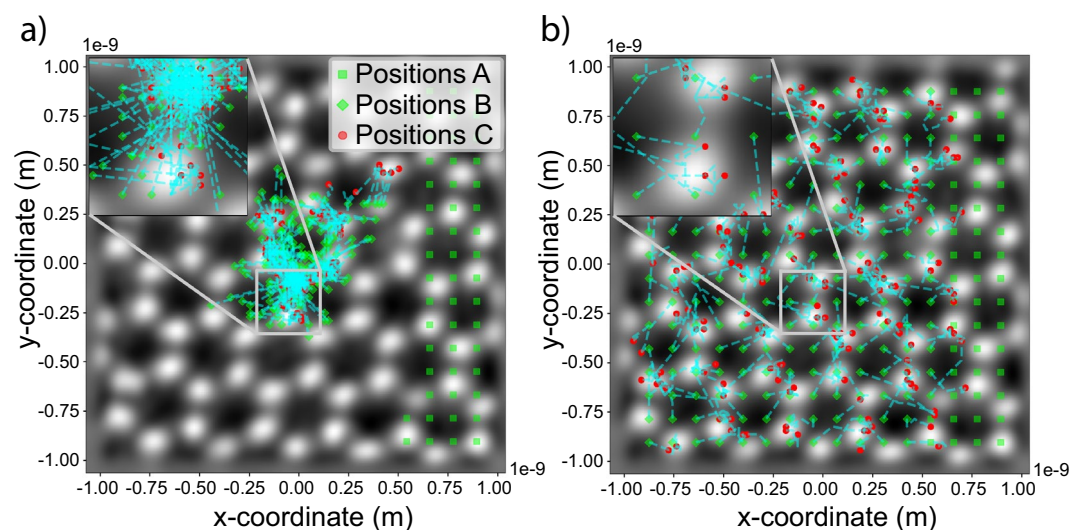
arbitrarily initialized, this is not ideal. Having an adequate initial guess of the policy and using RL subsequently to only fine-tune the policy is a much easier problem to solve. A sparse grid scan sequence is a reasonable initialization as it follows the current scanning convention used in a microscope. Pre-training of the parameterized policy for the RL model can then be performed by supervised learning applied on the RNN such that the discrepancy between the predicted scan positions  $\vec{\mathbf{R}}_{P_t} = \vec{\mu}_\theta(\mathbf{h}_t)$  and the scan positions of the initialization sequence  $\vec{\mathbf{R}}_{P_t}^{\text{init}}$  is minimized:

$$\mathcal{H}(\theta) = \sum_t^T \|\vec{\mathbf{R}}_{P_t} - \vec{\mathbf{R}}_{P_t}^{\text{init}}\|_2^2. \quad (9)$$

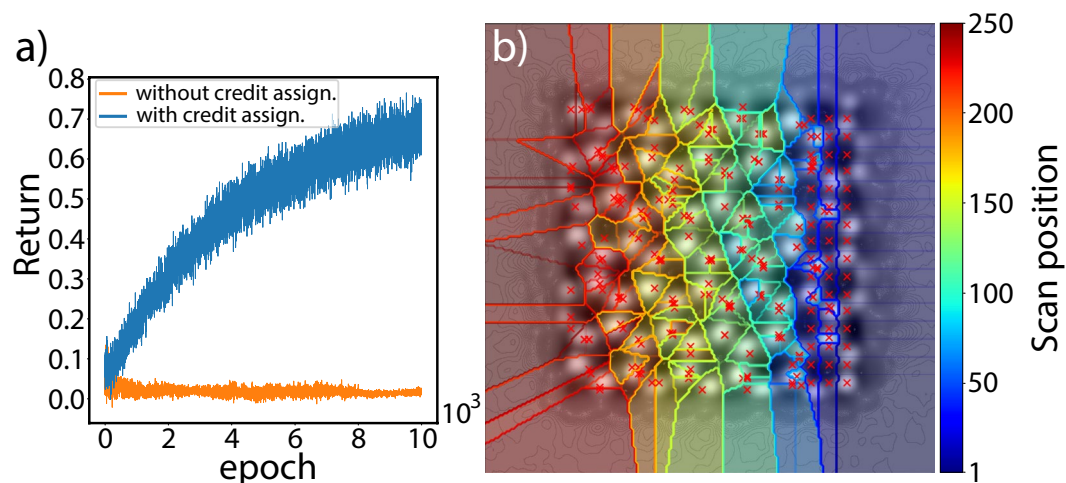
Figure 5 illustrates the scan positions during the fine-tuning of the policy through RL for the first 10,000 iterations when either (a) a policy that has not been initialized via supervised learning or (b) an initialized policy is used. While the scan positions in both cases converge to the atomic structure, the positions predicted by the non-initialized policy are distributed only within a small region of the field of view during the entire training. Note that once the predicted scan sequence mimics the sparse grid scan sequence as a result of the supervised learning based initialization, all further improvements in performance are the result of the subsequent training through RL.

A high variance of gradient estimates is another problem that particularly strongly affects the Monte Carlo policy gradient method<sup>42,44,45</sup>. Due to this, the sampling efficiency is relatively low, which causes a slow convergence to a solution. This makes deep RL applied to ptychography challenging as the image reconstruction itself requires iterative processing. The high variance can be in part attributed to the difficulty of assigning credit from the overall performance to an individual agent's action. Here, we introduce a way to estimate the reward function in order to tackle the credit assignment problem for adaptive scanning in ptychography. The reward function should naturally correspond to the quality of the ptychographic reconstruction. We have found empirically that a high reconstruction quality correlates positively with a high dynamic range in the phase. Therefore, the reward function could intuitively be formalized by  $r^m(\mathbf{a}_t|s_t) = P^{-1} \sum_{\mathbf{r} \in \text{FOV}} V(\mathbf{r})$ , where  $P$  is the total number of scan positions. This formulation, however, does not solve the credit assignment problem and results in an insufficient training performance, as shown in Fig. 6a. To estimate the reward for the actions of each individual agent, we use a tessellation method that partitions the atomic potential into small segments. A Voronoi diagram<sup>46</sup>, where each position corresponds to a seed for one Voronoi cell, enables assignment of only a part of the total phase to each position. More precisely, the Voronoi diagram formed by the predicted scan positions is overlaid with the corresponding ptychographic reconstruction at the end of the prediction process and the summed phase within each Voronoi cell is the reward for that cell's seed position. The reward function can be expressed by  $r^m(\mathbf{a}_t|s_t) = P^{-1} \sum_{\mathbf{r} \in \text{Cell}^m} V(\mathbf{r})$ . Figure 6b shows a Voronoi diagram generated by predicted scan positions.

**Settings.** For the experimental investigation, we acquired multiple ptychographic data sets from a monolayer molybdenum disulfide (MoS<sub>2</sub>) specimen with a Nion HERMES microscope. The microscope was operated with a 60 kV acceleration voltage, a convergence angle of 33 mrad and diffraction patterns with a pixel size



**Figure 5.** Fine-tuning of a policy with RL that (a) has not been initialized and (b) has been initialized via supervised learning. In the latter case, the training starts with a sequence that matches a sparse grid scan sequence. Positions A indicate the scan positions of the first sub-sequence  $\vec{\mathbf{R}}_{P_0}$  that is provided to the RNN as part of the initial input. Positions B and C are the scan positions of all predicted sub-sequences at iteration 0 and 10,000, respectively. The trajectories they form during the optimization process are indicated by dashed blue lines.



**Figure 6.** (a) Learning curves for the first 10,000 iterations of RL with multiple agents and without credit assignment or with credit assignment, illustrated in orange and blue, respectively. (b) A Voronoi diagram is used to assign a unique reward to each scan position of the predicted sequence. The scan positions are shown as red dots, where the first 50 positions are distributed on the right side within the dark blue area. For visualization purpose, the ground truth reconstruction is included in the diagram.

of 0.84 mrad were acquired using a Dectris ELA direct-electron detector mounted at the electron energy-loss spectroscopy (EELS) camera port. Distortions induced by the EEL spectrometer were corrected with in-house developed software. For the ptychographic data set acquisition, a conventional grid scan with a scanning step size of 0.02 nm was used. From the experimentally acquired data sets we created 200 smaller data sets, each with 10,000 diffraction patterns and located at different regions of the sample. 175 of these small data sets were dedicated to the training of the network model, while the remaining 25 data sets were used to test the trained model and to generate the results shown here. The diffraction patterns were binned by a factor of 2 to  $64 \times 64$  pixels. The adaptive scanning model was trained on the small data sets with the goal of predicting optimal scan sequences of 250 to 500 probe positions, out of the possible 10,000, which corresponds to a dose reduction by a factor of 40 to 20. Each sub-sequence contains 50 to 100 positions, where the initially given first sub-sequence follows a sparse grid scanning sequence.

We conducted a second investigation of the model's performance on simulated data sets of a DWCNT and compared it to the performance of alternative low-dose data acquisition methods in ptychography. The data sets were generated using the simulation tool *abTEM*<sup>47</sup> where we set the acceleration voltage to 60 kV, the convergence angle to 40 mrad and the scanning step size to 0.02 nm. The size of the diffraction patterns is  $86 \times 86$  pixels with a pixel size of 0.91 mrad. 962 data sets have been used to train the network model and from the 13,225 diffraction patterns in each data set only 840 were chosen within the prediction process of the adaptive scanning workflow. 25 data sets were used to compare the different scanning procedures and analyse their performance using the  $Q_{SSIM}$  metric. The simulated DWCNT consisted of an inner and an outer nanotube with a diameter of 9.78 Å and 16.44 Å and a chiral angle of  $44^\circ$  and  $60^\circ$ , respectively. For each data set, a unique rotation between the two nanotubes and a translation of the entire DWCNT within the field of view was applied.

The ptychographic reconstructions were performed with an optimized version of ROP<sup>17,28</sup> that allows simultaneous reconstruction from a batch of different data sets and therefore the parallel hardware architecture of a NVIDIA V100 GPU could be optimally used to efficiently train the model. For a batch size of 24, reconstructions were retrieved in about 35 s. A gradient descent step size  $\alpha_{ROP}$  of 525 was chosen and the potential was retrieved at iteration 5. In the experimental investigation, the reconstructed potential was  $200 \times 200$  pixels with a pixel size of 0.0154 nm, for a field of view of  $2 \times 2$  nm, while for the simulation, the reconstructed potential was  $200 \times 200$  pixels with a pixel size of 0.0140 nm, for a field of view of  $2.3 \times 2.3$  nm. For the generation of the reward function, Voronoi diagrams were generated with the Jump Flooding Algorithm<sup>48</sup> and for the implementation of the network models, PyTorch<sup>49</sup> was used. For the compression of structure information, we used a convolutional autoencoder consisting of 6 convolutional layers with kernels of dimension 3, a stride of 1 and channels that ranged from 16 to 512 for the encoder and decoder part, respectively. The input of the autoencoder had a dimension of 512 with a pixel size of 0.0064 nm and thus a scaling and an interpolation was required before the potential generated by ROP could be compressed. In addition, the value of the potential  $V_i$  at each pixel  $i$  was transformed to zero mean and unit variance. For the prediction of the scan sequences, pre-training and fine-tuning was performed with a RNN model composed of 2 stacked GRU layers with hidden states  $H_i$  of size 2048, the Adam optimizer<sup>50</sup> with a learning rate  $\alpha_{RNN}$  of  $1e-5$  and a batch size of 24. For the fine-tuning, a policy with variances of  $\sigma_x^2 = \sigma_y^2 = 0.0125^2$  was chosen and a myopic behavior was enforced by setting the discount factor for the return,  $G$ , to  $\gamma = 0$ . In the case of the experimental investigation, training of the autoencoder involved 100,000 iterations, while the training of the RNN with supervised learning and RL has been performed with 800 and 20,000 iterations, respectively. In the second investigation with the simulated data, the training of the



autoencoder has been done with 30,000 iterations and for the training of the RNN with supervised learning and RL, 200 and 2800 iterations were used, respectively.

### Code availability

The source code of the presented work is available at Gitlab<sup>51</sup> and the experimental and simulated data can be found in Ref.<sup>28</sup>. All other code is available from the corresponding author on reasonable request.

Received: 20 December 2022; Accepted: 23 May 2023

Published online: 30 May 2023

### References

- Rodenburg, J. M. Ptychography and related diffractive imaging methods. *Adv. Imaging Electron Phys.* **150**, 87–184 (2008).
- Humphry, M., Kraus, B., Hurst, A., Maiden, A. & Rodenburg, J. Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging. *Nat. Commun.* **3**, 1–7 (2012).
- Lozano, J. G., Martinez, G. T., Jin, L., Nellist, P. D. & Bruce, P. G. Low-dose aberration-free imaging of Li-rich cathode materials at various states of charge using electron ptychography. *Nano Lett.* **18**, 6850–6855 (2018).
- Jiang, Y. *et al.* Electron ptychography of 2D materials to deep sub-ångström resolution. *Nature* **559**, 343–349 (2018).
- Zhou, L. *et al.* Low-dose phase retrieval of biological specimens using cryo-electron ptychography. *Nat. Commun.* **11**, 1–9 (2020).
- McMullan, G., Faruqi, A. & Henderson, R. Direct electron detectors. *Methods Enzymol.* **579**, 1–17 (2016).
- Tate, M. W. *et al.* High dynamic range pixel array detector for scanning transmission electron microscopy. *Microsc. Microanal.* **22**, 237–249 (2016).
- Thibault, P. *et al.* High-resolution scanning x-ray diffraction microscopy. *Science* **321**, 379–382 (2008).
- Maiden, A. M. & Rodenburg, J. M. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy* **109**, 1256–1262 (2009).
- Van den Broek, W. & Koch, C. T. Method for retrieval of the three-dimensional object potential by inversion of dynamical electron scattering. *Phys. Rev. Lett.* **109**, 245502 (2012).
- Maiden, A. M., Humphry, M. J. & Rodenburg, J. Ptychographic transmission microscopy in three dimensions using a multi-slice approach. *JOSA A* **29**, 1606–1614 (2012).
- Van den Broek, W. & Koch, C. T. General framework for quantitative three-dimensional reconstruction from arbitrary detection geometries in TEM. *Phys. Rev. B* **87**, 184108 (2013).
- Tsai, E. H., Usov, I., Diaz, A., Menzel, A. & Guizar-Sicairos, M. X-ray ptychography with extended depth of field. *Opt. Express* **24**, 29089–29108 (2016).
- Jiang, Y. *et al.* Breaking the Rayleigh limit in thick samples with multi-slice ptychography. *Microsc. Microanal.* **24**, 192–193 (2018).
- Pelz, P. M., Qiu, W. X., Bücke, R., Kassier, G. & Miller, D. R. J. Low-dose cryo electron ptychography via non-convex Bayesian optimization. *Sci. Rep.* **7**, 9883 (2017).
- Song, J. *et al.* Atomic resolution defocused electron ptychography at low dose with a fast, direct electron detector. *Sci. Rep.* **9**, 1–8 (2019).
- Schloz, M. *et al.* Overcoming information reduced data and experimentally uncertain parameters in ptychography with regularized optimization. *Opt. Express* **28**, 28306–28323 (2020).
- Chen, Z. *et al.* Mixed-state electron ptychography enables sub-angstrom resolution imaging with picometer precision at low dose. *Nat. Commun.* **11**, 1–10 (2020).
- Huang, X. *et al.* Optimization of overlap uniformness for ptychography. *Opt. Express* **22**, 12634–12644 (2014).
- Levine, S., Finn, C., Darrell, T. & Abbeel, P. End-to-end training of deep visuomotor policies. *J. Mach. Learn. Res.* **17**, 1334–1373 (2016).
- Andrychowicz, O. M. *et al.* Learning dexterous in-hand manipulation. *Int. J. Robot. Res.* **39**, 3–20 (2020).
- Sharma, S., Kiros, R. & Salakhutdinov, R. Action recognition using visual attention. arXiv preprint [arXiv:1511.04119](https://arxiv.org/abs/1511.04119) (2015).
- Ba, J., Mnih, V. & Kavukcuoglu, K. Multiple object recognition with visual attention. arXiv preprint [arXiv:1412.7755](https://arxiv.org/abs/1412.7755) (2014).
- Mnih, V., Heess, N., Graves, A. *et al.* Recurrent models of visual attention. In *Advances in Neural Information Processing Systems* 2204–2212 (2014).
- Betterton, J.-R., Ratner, D., Webb, S. & Kochenderfer, M. Reinforcement learning for adaptive illumination with x-rays. In *2020 IEEE International Conference on Robotics and Automation (ICRA)* 328–334 (IEEE, 2020).
- Vasudevan, R. K. *et al.* Autonomous experiments in scanning probe microscopy and spectroscopy: Choosing where to explore polarization dynamics in ferroelectrics. arXiv preprint [arXiv:2011.13050](https://arxiv.org/abs/2011.13050) (2020).
- Dahmen, T. *et al.* Feature adaptive sampling for scanning electron microscopy. *Sci. Rep.* **6**, 1–11 (2016).
- Schloz, M., Müller, J., Pekin, T. C., Van den Broek, W. & Koch, C. T. Deep reinforcement learning for data-driven adaptive scanning in ptychography. *Zenodo*<https://doi.org/10.5281/zenodo.7865575> (2023).
- Van Heel, M., Keegstra, W., Schutter, W. & Van Bruggen, E. Arthropod hemocyanin structures studied by image analysis. *Life Chem. Rep. Suppl* **1**, 5 (1982).
- 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 (2004).
- Pelz, P. M., Johnson, I., Ophus, C., Ercius, P. & Scott, M. C. Real-time interactive 4D-stem phase-contrast imaging from electron event representation data: Less computation with the right representation. *IEEE Signal Process. Mag.* **39**, 25–31 (2021).
- Madsen, J. *et al.* Ab initio electrostatic potentials for 4D-stem ptychographic reconstruction. *Microsc. Microanal.* **28**, 392–393 (2022).
- Huang, P. Y. *et al.* Direct imaging of a two-dimensional silica glass on graphene. *Nano Lett.* **12**, 1081–1086 (2012).
- Kalinin, S. V. *et al.* Automated and autonomous experiment in electron and scanning probe microscopy. arXiv preprint [arXiv:2103.12165](https://arxiv.org/abs/2103.12165) (2021).
- Kalinin, S. V. *et al.* Machine learning in scanning transmission electron microscopy. *Nat. Rev. Methods Primers* **2**, 1–28 (2022).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
- Elman, J. L. Finding structure in time. *Cogn. Sci.* **14**, 179–211 (1990).
- Werbos, P. J. Generalization of backpropagation with application to a recurrent gas market model. *Neural Netw.* **1**, 339–356 (1988).
- Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint [arXiv:1412.3555](https://arxiv.org/abs/1412.3555) (2014).
- Masci, J., Meier, U., Cireşan, D. & Schmidhuber, J. Stacked convolutional auto-encoders for hierarchical feature extraction. In *International Conference on Artificial Neural Networks* 52–59 (Springer, 2011).
- Williams, R. J. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Mach. Learn.* **8**, 229–256 (1992).

42. Liu, Y., Zhang, K., Basar, T. & Yin, W. An improved analysis of (variance-reduced) policy gradient and natural policy gradient methods. In *NeurIPS* (2020).
43. Sutton, R. S., McAllester, D. A., Singh, S. P. & Mansour, Y. Policy gradient methods for reinforcement learning with function approximation. In *Advances in Neural Information Processing Systems* 1057–1063 (2000).
44. Wu, C. et al. Variance reduction for policy gradient with action-dependent factorized baselines. arXiv preprint [arXiv:1803.07246](https://arxiv.org/abs/1803.07246) (2018).
45. Feriani, A. & Hossain, E. Single and multi-agent deep reinforcement learning for AI-enabled wireless networks: A tutorial. *IEEE Commun. Surv. Tutor.* **23**, 1226–1252 (2021).
46. Okabe, A. Spatial tessellations. In *International Encyclopedia of Geography: People, the Earth, Environment and Technology: People, the Earth, Environment and Technology* 1–11 (2016).
47. Madsen, J. & Susi, T. The abTEM code: Transmission electron microscopy from first principles. *Open Res. Eur.* **1**, 13015. <https://doi.org/10.12688/openreseurope.13015.1> (2021).
48. Rong, G. & Tan, T.-S. Jump flooding in GPU with applications to Voronoi diagram and distance transform. In *Proceedings of the 2006 Symposium on Interactive 3D Graphics and Games* 109–116 (2006).
49. Paszke, A. et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems* Vol. 32 (eds Wallach, H. et al.) 8024–8035 (Curran Associates, Inc., 2019).
50. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
51. <https://gitlab.com/Schlozma/adaptive-scanning-in-ptychography>.

## Acknowledgements

M.S., T.C.P. and W.V.d.B. acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Grant No. BR 5095/2-1. M.S., J.Mü. and C.T.K. acknowledge support from the DFG—Project-ID 182087777—SFB 951. M.S. and C.T.K. acknowledge financial support from the DFG—Project-ID 414984028—SFB 1404. We thank Prof. Sang Ho Oh and Dr. Jinsol Seo (Korea Institute of Energy Technology, Naju, Korea), as well as Dr. Bumsu Park (CEMES, Toulouse, France) for providing the MoS<sub>2</sub> sample. We acknowledge funding from the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme via Grant agreement No. 756277-ATMEN (J.Ma. and T.S.). M.S. would like to thank Jayesh K. Gupta (Microsoft, Stanford, USA) for his helpful suggestion about the description of the POSG formalism.

## Author contributions

M.S. conceived the idea of the adaptive scanning workflow. J.Mü. carried out the ptychography experiment. M.S. implemented all the neural network applications and modified the ptychography reconstruction algorithm. M.S. wrote up the manuscript. J.Ma. and T.S. developed the atom finding method used for the alternative low-dose procedure. T.C.P., W.V.d.B., T.S. and C.K. commented on the manuscript. W.V.d.B. and C.K. supervised the project. The manuscript reflects the contributions of all authors.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-35740-1>.

**Correspondence** and requests for materials should be addressed to M.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2023