The Advantages of Sub-Sampling and Inpainting for Scanning Transmission Electron Microscopy

Nigel D. Browning[1, 2,*], Jony Castagna[3,4], Angus I. Kirkland[5,6], Amirafshar Moshtaghpour[1,5], Daniel Nicholls[1], Jack Wells[3], Alex W. Robinson[1], Yalin Zheng[3,7]

[1] Mechanical, Materials & Aerospace Eng., University of Liverpool, Liverpool, L69 3GH, UK

[2] Sivananthan Laboratories, 590 Territorial Drive, Bolingbrook, IL 60440, USA

[3] Distributed Algorithms CDT, University of Liverpool, Liverpool, L69 3GH, UK

[4] UKRI-STFC Hartree Centre, Daresbury Laboratory, Warrington, WA4 4AD, UK

[5] Rosalind Franklin Institute, Harwell Science and Innovation Campus, Didcot, OX11 0QS, UK

[6] Department of Materials, University of Oxford, Oxford, OX2 6NN, UK

[7] Department of Eye & Vision Sciences, University of Liverpool, Liverpool, L7 8TX, UK.

*Corresponding author: nigel.browning@liverpool.ac.uk

Abstract

Images and spectra obtained from aberration corrected scanning transmission electron microscopes (STEM) are now used routinely to quantify the morphology, structure, composition, chemistry, bonding, and optical/electronic properties of nanostructures, interfaces and defects in many materials/biological systems. However, obtaining quantitative and reproducible atomic resolution observations from some experiments is actually harder with these ground-breaking instrumental capabilities, as the increase in beam current from using the correctors brings with it the potential for electron beam modification of the specimen during image acquisition. This beam effect is even more acute for in-situ STEM observations, where the desired outcome being investigated is a result of a series of complicated transients, all of which can be modified in unknown ways by the electron beam. The aim in developing and applying new methods in STEM is therefore to focus on more efficient use of the dose that is supplied to the sample and to extract the most information from each image (or set of images). For STEM (and for that matter all electron/ion/photon - scanning systems), one way to achieve this is by sub-sampling the image and using Inpainting algorithms to reconstruct it. By separating final image quality from overall dose in this way, and manipulating the dose distribution to be best for the stability of the sample, images can be acquired both faster and with less beam effects. In this paper the methodology behind sub-sampling and Inpainting is described, and the potential for Inpainting to be applied to novel real time dynamic experiments will be discussed.

1. Introduction to Scanned Imaging

In a standard scanning pattern used for Scanning Transmission Electron Microscopes (STEM) and many other instruments, the scan system works by moving the beam from left to right across a single row with a dwell time for each pixel in that row (Figure 1). At the end of the row, the beam flies back to the left-hand-side, moves down one pixel and then completes a row again (this is like the way a traditional typewriter works or an older cathode ray tube (CRT) television). After the flyback, the beam typically has a longer dwell time at the left-edge to allow for any hysteresis in the scan to damp out and the left-edge of the scan to be aligned at the same location for each row. This form of scanning is known to present difficulties with beam damage, particularly on the left-hand edge of the raster, and this has led to alternative spiral scanning approaches[1,2] to extract higher resolution images with less beam damage.

The beam size in the STEM is the same regardless of the magnification of the image, which can be as small as ~0.1nm for a Cs-corrected STEM[3,4]. In a low magnification image therefore, the area of the scan is large and the pixel size is correspondingly much larger than the size of the beam. For example, for the Cs-corrected STEM above, in a 1000 x 1000-pixel scan covering 1mm x 1mm, the pixel size is 1μm, i.e. 1000 x the size of the beam. To achieve atomic resolution in STEM, the magnification of the microscope is increased to the point where the pixel size approaches atomic separation, i.e. ~0.1-0.5nm. In the highest resolution images, the magnification is turned up to a level where the pixel size is actually much smaller than the probe size, leading to an oversampled image where beam damage is prevalent[5].

When a STEM is at low-magnification, beam damage is not a critical issue, as the distance between beam locations is very large, and the likelihood that the scan hits exactly the same location in successive sweeps is very small – damage can still occur, but it is below the scale of the image resolution. It is only when the beam and pixel size starts to converge that the damage becomes serious, and this is of course the condition for the highest spatial resolution images. If we think about the problem from the perspective of overlapping beam positions and their effect on the measurable damage, then it is clear that if we can increase the spacing of the beam positions at high magnification then we will be able to avoid/reduce the beam damage problem that plagues high resolution SEM/STEM (Figure 1). This reduction in beam damage effects is actually what has been seen in cases where expanding the time and space between measurements reduces damage[6-12]. Of course, the issue with this "sparse sampling" approach is that we would then need a means to reconstruct the full image from the sub-sampled acquisition. As the quality of the image then would depend on how many pixels were included, the optimal sampling would be defined as where beam damage is minimized while the reconstruction quality is maximized. We note here that the recent development of the Moiré STEM achieves atomic resolution images from a lower magnification image (i.e. sub-sampled scan) with reduced dose by utilising a geometrical interference effect between the STEM beam and the sample lattice[13-15]. This methodology represents a special case, i.e. where the sample geometry is known and an atomic resolution image is desired, of the general sub-sampling approaches and reconstructions described in the remainder of this manuscript.

2. Sparse Sampling and Reconstruction from a Single Image

As we can see from Figure 1, it is possible to obtain a sub-sampled scanned image by using both a set of "random" beam positions, and a "random walk" or "line-hop" scan. Practically, the line-hop approach is easier to implement on standard electron microscopes

as it avoids much of the hysteresis issue present in conventional scanning systems, permitting the system to run at the fastest possible speed[11] (more advanced electrostatic deflection systems can avoid this hysteresis[5] and use the full benefit of a truly random scan). The key challenge for all sub-sampling methods is to reconstruct the sub-sampled image. Compressive sensing[16,17] is a method of efficient signal acquisition and reconstruction via the solving of a set of undetermined linear equations. Like traditional image compression techniques, it relies upon the fact that given an appropriate coordinate system (or 'Dictionary'), complex high dimensional signals such as an image can be expressed within a margin of error by a potentially much smaller set of parameters, describing a linear combination of signal patterns with their respective scalar coefficients[18,19]. The goal for any image reconstruction is to form a complete signal (with the smallest error) from as few measurements as possible[20,21].

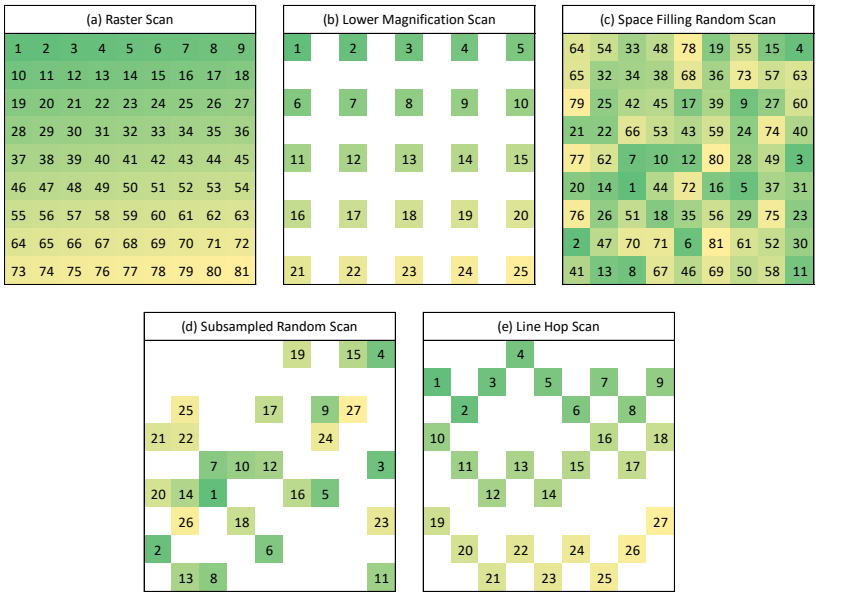


Figure 1: *Examples of various scanning patterns in a 9x9 grid. Number and colour indicate scanning order. (a) Raster scanning is the traditional method of scanning in STEM. (b) Down sampling akin to low magnification image acquisition. (c) Space filling random scanning has been shown to reduce beam damage in beam sensitive samples[10]. (d,e) Two scanning patterns possible using probe subsampling; subsampled random scanning and sub-sampled line hop (i.e. random walk) scanning at 33.3% sampling ratio.*

Traditional image Inpainting methods rely on a dictionary learning algorithm such as the Method of Optimal Directions (MOD)[22] or K-SVD[23] to form a dictionary of representative signal patterns from a fully sampled image, which via a sparse linear combination with corresponding scalar coefficients can closely represent any given patch (smaller segment) of the image. This dictionary, along with a now sub-sampled version of the same image may

then be passed to a sparse pursuit algorithm to solve the following system of equations (for each $\boldsymbol{i}$-th overlapping patch of the image):

$$\boldsymbol{v_i} = \boldsymbol{\Phi_i} \cdot (\boldsymbol{D\alpha_i} + \boldsymbol{\epsilon_i}), \qquad \text{(Eq. 1)}$$

$$\hat{\boldsymbol{x}}_{\boldsymbol{i}} = \boldsymbol{D\alpha_i}, \qquad \text{(Eq. 2)}$$

where $\boldsymbol{v_i} \in \mathbb{R}^n$ is the measured (sub-sampled) signal subject to noise $\boldsymbol{\epsilon_i} \in \mathbb{R}^n$, $\boldsymbol{\Phi_i} \in \{0,1\}^{\boldsymbol{n}\times\boldsymbol{n}}$ is the binary sensing matrix (or 'mask', determining the locations of missing pixels), $\hat{\boldsymbol{x}}_{\boldsymbol{i}} \in \mathbb{R}^n$ is the reconstructed (fully-sampled) signal represented using the given dictionary $\boldsymbol{D} \in \mathbb{R}^{n\times k}$ and corresponding 'weight' vector $\boldsymbol{\alpha_i} \in \mathbb{R}^k$.

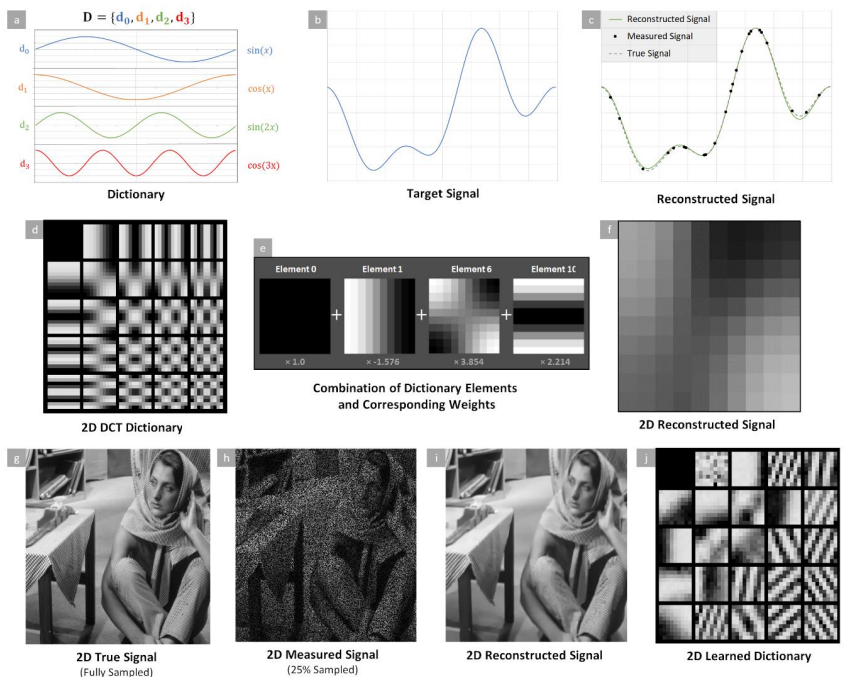


Figure 2: *(a) A series of 1D dictionary elements (Fourier components in this case) can be combined with defined scalar weightings to reproduce the true 1D signal (b). By reducing the sampling of our experiment, but still generate a good fit to the data (c). (d) A Dictionary of 2D elements (again Fourier components) can be combined with defined scalar weightings (e) to produce a reconstruction (f) which is indistinguishable from the true 2D signal. By reducing the number of points in the true signal (g) we can increase the speed and decrease the dose in the sampling of our experiment (h), but still generate a good fit to the data (i). For comparison (j) shows the dictionary learned directly from the sub-sampled image in (h).*

Examples of algorithms capable of solving this includes Orthogonal Matching Pursuit (OMP) and its many similar variants[24,25] or Basis Pursuit which involves the minimisation of

the l_1 norm[26]. More recently, a reformulation of the dictionary learning problem into the Bayesian regime has produced algorithms, such as Beta-Process Factor Analysis (BPFA)[27] which are capable of so-called "blind inpainting"[28], i.e. the formation of a dictionary and subsequent reconstruction of an image using *only* the subsampled image as its input (therefore, there is no need for a mask to be provided a priori or a fully-sampled version of the image to be acquired at any stage in the process). For this reason, BPFA represents an ideal starting algorithm for sub-sampled scanned images, which can be refined using dictionary seeding/transfer approaches (see later sections of this paper).

As an example of how this process can work, consider the case of a simple 1-dimensional (1D) signal, such as a wave shown in Figure 2. Here, a series of dictionary elements (in this case 1D Fourier components) can be used to re-construct a true signal (Figure 2b). But now what happens if we do not measure the complete signal? Figure 2c shows 25% subsampling of the true signal from Figure 2b. It is clear from Figure 2c that we can fit the dictionary elements to the sub-sampled observation, effectively "inpainting" the missing level of sampling in our experiment. As we reduce the level of sampling, the ability to "fit" to the data with a minimal error is reduced, until typically at ~2% sampling the error is unacceptably large[29]. However, given that the damage induced in the sample is a function of the speed of the scan, and the overlap of the beam positions, reducing the overall number of beam positions in the image by this factor of 50, can have a tremendous effect on the overall sample stability during the experiment. This approach is also easily extendable to higher dimension images, with the same approach as above shown for the reconstruction of the 2D image of "Barbara" (Figure 2d-i). In this case the reconstruction was obtained using BPFA to Inpaint the sub-sampled image[30]. In the use of the BPFA methodology there are a number of tuneable parameters that are used to increase the efficiency of the algorithms to reconstruct the images[11]. Here again, it is possible to reconstruct 2D images with high precision from a sampling of ~2%[31,32] (this also extends to non-rectangular scans[33], 3D tomography[34,35], 4D methods such as ptychography[36,37], and higher dimensional datasets).

3. Dictionary Transfer and Seeding with Fast Simulations

From the results in the previous section, it is clear that we can use a dictionary for our single sub-sampled image with standard Fourier components or learn the dictionary for the reconstruction directly from the sub-sampled image using BPFA[11]. However, could we use a dictionary that we have learned from one image to reconstruct another? The reason we may want to consider this is that if one type of image had a better signal-to-noise ratio (SNR) then we may get a more accurate dictionary, improving the speed and fidelity of the reconstruction of subsequent images if we use the dictionary from that "best image"[30]. We may also be able to create a master dictionary that would allow us to sub-select the best dictionary to reconstruct any particular set of images. When using BPFA, which has the benefit of working directly on sub-sampled images, the most time-consuming part of the reconstruction process is the dictionary determination – we can speed up significantly by using an existing optimal dictionary. In order to test this dictionary transfer concept, we can simply try it for two images that show strikingly different contrast and see what happens. Figure 3 shows two images, their dictionaries learnt by K-SVD and the reconstructions of 25% sub-sampled images using the dictionary from the other image. The quality of the reconstructions in each case are shown in table 1. As can be clearly seen, it is possible to swap dictionaries and still achieve a

high-quality reconstruction. The limitation in this example is that it takes longer to achieve the required reconstruction quality.

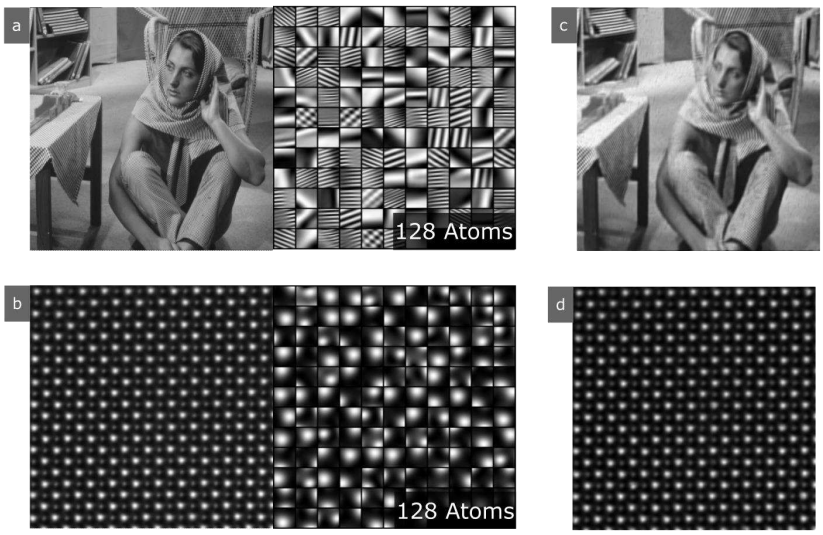


Figure 3: *(a) Image of Barbara and dictionary (b) atomic resolution STEM image of $SrTiO_3$ and dictionary (c) 25% Barbara reconstructed with STEM dictionary and (d) 25% STEM reconstructed with Barbara dictionary.*

		Input Image			
		Barbara		STEM	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM
Dictionary	Barbara	24.72	0.77	25.66	0.9
	STEM	23.81	0.63	31.46	0.98
	Simulation*	-	-	26.13	0.92

Table 1: *Comparison of reconstruction quality for the images shown in Figures 3 and 4 using standard Peak Signal to Noise Ratio (PNSR)[38] measured in decibels (dB) (the higher the number the better the reconstruction) and Structural Similarity Index Measure (SSIM)[39] metrics (a number closer to 1 is better). As a general rule of thumb, a PSNR value greater than 20 dB is an acceptable reconstruction, over 25 dB is a very good reconstruction, and over 30 dB is indistinguishable from the ground truth (also the same for an SSIM of 0.95 or above) *Note that dictionary generated from image simulation was applied to a cropped area of the full STEM image.*

A particularly useful application of the dictionary transfer process is in the use of simulations for image analysis. For many types of images in STEM, SEM and other microscopies, simulations are used to match expected contrast to actual contrast and quantify the structure giving rise to the image. Usually this is a computationally intensive

methodology as the potential unknowns in the image (composition, thickness, detection efficiency, defocus, etc) require multiple simulations to be performed to match to the experimental image[4]. In the case of the sub-sampling approach, however, we can use the dictionary from a simulated image to seed the experimental reconstruction, or we can use a dictionary learned from experimental image to match to the theory, or we can combine the two images together to make a combined dictionary that is optimal for both[34]. As we do not need the full dataset for the simulations (BPFA can reconstruct a partial simulation), we can also sub-sample the number of frozen phonon configurations, locations in real space and experimental conditions to increase the speed for the simulations, potentially allowing for each of them to be performed in real time during experiments[40]. As with the reconstructions described previously, by having a dictionary and reconstruction it means that there are quantifiable fits to the experiment, theory, and combination of the two. Rather than matching to experimental conditions, the theory and experiment are solved together and the differences can be quantified and identified from a single simulation. As a simple example of this approach, 10% subsampled simulations of $SrTiO_3$ obtained using MULTEM[41] are shown in Figure 4a, along with their dictionary and full reconstruction. Also shown in Figure 4 is the reconstruction of the 25% $SrTiO_3$ image shown in Figure 3, using the simulated dictionary in Figure 4c. A quantified comparison of the reconstruction of this image is shown in Table 1. A key factor here is that unlike the example of Figure 3, where the dictionary was poorly suited in the transfer, making the reconstruction become more complex, here we are simplifying the reconstruction by making the dictionary specifically for the experimental image we want to reconstruct

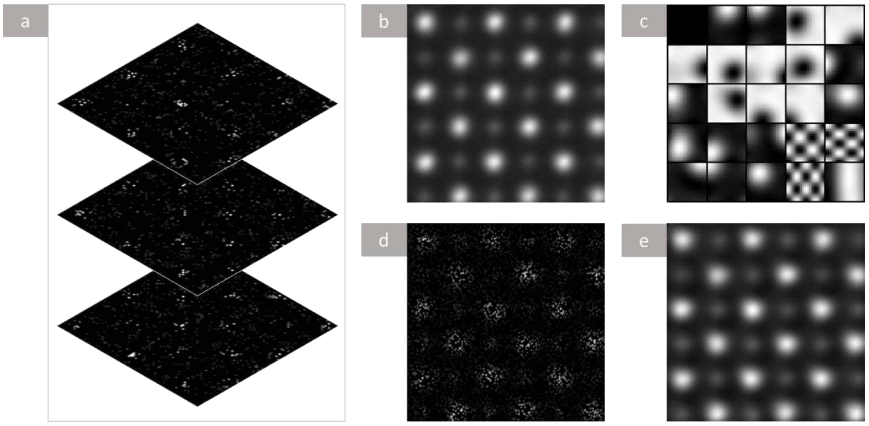


Figure 4: *(a) Three 10% sub-sampled simulations corresponding to different frozen phonon configurations of $SrTiO_3$, and the reconstructed simulation using BPFA (b) as well as the dictionary determined by BPFA (c). This dictionary is then used to reconstruct a 25% subsampled crop of the $SrTiO_3$ from Figure 3 (d), with the reconstruction (e) having a 92% similarity to the reference image.*

4. Learning for Compressed Hyperspectral Imaging

In the discussion so far, we have focused on an approach that uses the dictionary obtained from a single image (and our demonstration example has used atomic resolution images). However, if there are multiple images that are being acquired in an experiment (for example, with a movie and/or multiple different detector types), or we have multiple experiments performed on the same or similar samples, then there are opportunities to refine the reconstruction of sub-sampled images even further. This type of approach falls within the general topic of hyperspectral imaging (HSI)[42], and there are many powerful deep learning approaches that have been developed for this type of data structure[43,44]. In the case of their use in scanning (transmission) electron microscopy, the goal in the experiment is always to be ahead of the beam damage that is induced in the experiment and that means we will always be signal limited and looking for methods that will work under low SNR conditions. In the example we show in this section, we aim to reconstruct a general hyperspectral image datacube with 10 spectral channels and 1 BSE channel, and avoid the advantages of symmetry that can assist with the reconstruction of atomic resolution images.

Supervised methods[44] make use of a "ground truth" to train a deep neural network (DNN), capable of denoising images, by mapping the noisy image to a clean reference[46]. This can achieve state-of-the-art image performance with a large enough data set for training, but can struggle to adapt to unseen data with new models of noise. For the STEM case, a ground truth image may not be possible to acquire without damage. However, with unsupervised learning approaches it is possible for a DNN to instead learn the mapping between independently measured noisy images to predict a clean signal with no reference ground truth[47], achieving similar performance to the supervised method. *Noise2Self* [48] is one method that extended this idea to exploit the noise independence between pixels, relaxing the requirement for collecting two independent noisy images and providing theoretical performance guarantees of such an approach. A semi-supervised approach may also play an important role when there are a "ground truth" for part of the data.

Figure 5 shows a small section of a much larger area Energy Dispersive X-ray (EDS) spectrum map acquired with two different dwell times using a scanning electron microscope (SEM). Also shown are reconstructions of these two datasets using a 3D BPFA implementation, in which a high SNR fully sampled backscattered electron (BSE) image is included as one of the layers alongside the EDS data cube. Extended into 3D, rather than learning dictionary elements of size [b × b] as typical with 2D dictionary learning, dictionary elements of size [b × b × N] are used, where N is the number of layers to the data cube (N=11 for this example). In the case presented, each layer is a different elemental map produced, with the first layer being the BSE image. This dictionary, which considers features along the spectral axis, is designed to take advantage of the relationship that exists between both spatial and spectral information available in the data cube. In using the high signal BSE image, this method aims to take advantage of its high spatial resolution to aid in reconstruction. Two use cases are presented in Figure 5; a low dwell time acquisition (1µs) wherein the full spatial domain is sampled, and a high dwell time acquisition (100µs) in which only 1% of the spatial domain is sampled randomly. These two scenarios exhibit incompleteness in two domains; spectral and spatial, respectively. While the low dwell time dataset may sample every location, the spectral information collected at those locations is weak (low SNR). For the high dwell time dataset, the inverse is true; at each sampled location the full (high SNR) spectra is acquired but only a few datapoints are acquired. 3D BPFA can take both of these inputs and

produce detailed spectrum maps with significant quality increases over the original data. Both of these use cases, from an experimental point of view, would theoretically have the same incident electron dose rate and acquisition time. One of the main benefits of using a method such as this is that no training data or prior-learned model is required. All that is necessary is the input data, and a set of input parameters tailored to that input data.

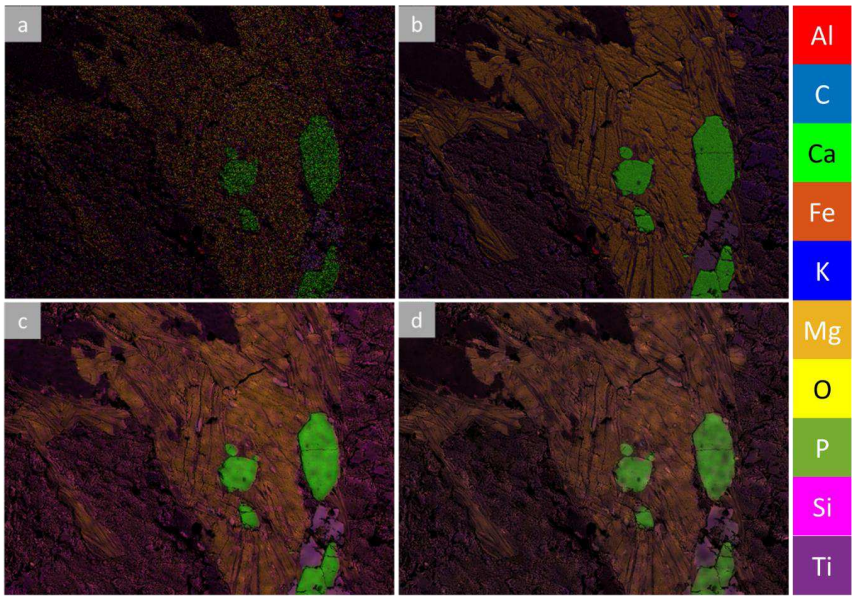


Figure 5: *(a) Untreated EDS spectrum image with 1 μs dwell time at 100% sampling from a sample of granite, containing the following minerals; muscovite (purple), illite (orange), quartz (dark purple), apatite (green) and rutile (lilac). (b) Untreated EDS spectrum image with 100 μs dwell time at 100% sampling. (c) 3D BPFA treated version of (a), 1 μs dwell time at 100% sampling. (d) 3D BPFA treated version of (b), with 100 μs dwell time but sampled randomly at 1%. Both (c) and (d) have the same theoretical electron dose exposure and acquisition time. Changes in image contrast between (a) and (b) are due to increased signal-to-noise ratio when scanning with higher dwell times. Changes in image contrast between (c) and (d), which are formed from the same electron dose rate, are due to the efficiency of the algorithm to construct the complete data set. All images contain the BSE image as a layer in the data cube. Image colour denotes chemical composition, as indicated by the legend. Image dimensions are 1.11 x 0.835 mm. We would like to thank Dr Louise Hughes and Dr Matthew Hiscock for supplying this data and help with analysis.*

5. Conclusions and Outlook

The methodology described here can allow scanned images to be acquired at least 10-100x faster and coupled with a dose fractionation that maximises the beam spacing in space and time, thereby minimising electron beam damage. Although the full potential of these

techniques, such as BPFA and DNN are yet to be fully exploited and evaluated, such an approach increases the range of beam sensitive samples that can be studied by advanced STEM (and other scanned) methods. In addition, this methodology is particularly important for the rapid throughput of scanned experiments, as technique developments over the last few have moved to more and more pixels over an ever-increasing analysis area. For example, ptychography in STEM[36], serial block face SEM[49] and FIB-SEM slice and view[50] are all pushing large scan dimensions even up to the 100 million pixel range, making the time of image acquisition the main limitation that is key to all future applications of these methods. In some cases, such as Z-contrast STEM[4], by reducing the number of pixels in an image, dynamic in-situ phenomena can be directly observed where the beam does not significantly change the observation[51]. Techniques that were once only thought of as only applicable to high resolution imaging can therefore now become in-situ methods. The use of DNNs to learn an inpainting model from a set of samples means that we can also in principle teach a microscope how to image a particular type of sample more efficiently.

Another aspect of scanned imaging that this form of sub-sampling and Inpainting reconstruction facilitates is the potential for dynamic control of the microscope. If we take a 1 mega pixel example with a 1 μs dwell time, then on a standard scan we would acquire ~1 frame per second (not accounting for the flyback time). At 10% sampling this becomes 10 frames/second, which means that we effectively can perform sub-original-frame analysis for key things like defocus, stigmation, drift, tilt etc or measure damage and damage rate independently to determine the best acquisition time for high resolution images. While the ability to reconstruct the image and perform the analytics to change the microscope alignment parameters would have to be performed in real time, there is increasing evidence that this should be routinely possible soon. The codes that perform the Inpainting work on a patch-by-patch basis across the whole image and each refinement can be performed in parallel, meaning that the speed can be increased by scaling efficient codes on a GPU platform. The difference in time between the Inpainting of a 5% sampled image and a 10% sampled image is negligible, meaning that the final speed of the imaging process is limited only by the acquisition time at these sampling levels. For things like defocus and drift it is also important to understand that we also do not need to reconstruct across the whole image to accomplish control, a few patches at specific locations would be all that would be needed. Another key part of the speed discussion is that there is a difference between the image that is needed to control the microscope and the image that is used for the final analysis. A poorer quality, i.e. faster reconstruction, could be enough to control the microscope and ensure the best experimental imaging parameters were used during the experiment. After the experiment is completed all the images could be transferred and analysed with a range of dictionaries, algorithms and approaches to determine the best result after the optimal reconstruction methods were used. Of course, at this point, the best overall result would then be used to modify the dictionary/algorithm in a positive way, allowing all previous images to be re-analysed to improve existing/prior results.

It is worth noting that all of the above discussions for use of the data are with images that are ~1-10% of the size of the original fully sampled image. At the time of acquisition, whatever the best approach to reconstruction exists would set the limit of sub-sampling and ensure that there is no data loss ever. Improvements in reconstruction and analysis that may come from a better algorithm in the future are not inherent in the storage architecture, meaning that the smallest amount of data is always kept for the best algorithm available – a

better algorithm may lead to all existing images being reduced even further. Another key part in the discussion of scanned images, is that often it is not the image that is important but the analysis of what the image shows; for example, the number and size distribution of nanoparticles in a particular area or part of the structure. As there is enough information in the sub-sampled image to reconstruct a full image, it means there is enough information to work with advanced image analytics. So in the example above, the image analytics for edge detection and size determination would be able to run directly on the sub-sampled data, but as it is sub-sampled then there is less data to analyse and a faster result[52,53]. It may be possible that complete data analysis could be performed during a microscope experiment and users of the instruments leave with their data rather than their results to be analysed at a later time. As with the final resolution of the images, even if there are sacrifices in having large error bars in data to get it quickly, the quality of data will be enough to determine the course of the experiment and the errors can always be improved off-line using the best supporting data/algorithms.

In summary, whether sub-sampling and Inpainting is used to improve image speed, control dose fractionation and reduce damage, provide real time adaptive scanning and autonomous instrument controls, learn the optimal reconstruction and de-noising approach to all (sub- and fully- sampled) images, or simply to reduce size of images for storage and/or transmission, these applications can all fundamentally change the way that we acquire and use images from STEM and other scanned microscopes in the future.

6. References

1. X. Sang, A. R. Lupini, R. R. Unocic, M. Chi, A. Y. Borisevich, S. V. Kalinin, E. Endeve, R. K. Archibald, S. Jesse, *Advanced Structural and Chemical Imaging* **2**, 6 (2016)
2. X. Sang, A. R. Lupini, J. Ding, S. V. Kalinin, S. Jesse, R. R. Unocic, *Scientific Reports* **7**, 43585 (2017)
3. S. J. Pennycook and P. D. Nellist, *Scanning transmission electron microscopy: imaging and analysis*. (Springer Science & Business Media, 2011).
4. P. E. Batson, N. Dellby, and O. L. Krivanek, Nature **418** (6898), 617 (2002).
5. R. F. Egerton, Ultramicroscopy **127**, 100 (2013)
6. A. Stevens, H. Yang, L. Carin, I. Arslan, and N. D. Browning, Microscopy **63** (1), 41 (2014).
7. L. Kovarik, A. Stevens, A. Liyu, and N. D. Browning, Applied Physics Letters **109** (16), 164102 (2016).
8. A. Stevens, L. Luzi, H. Yang, L. Kovarik, B. L. Mehdi, A. Liyu, M. E. Gehm, and N. D. Browning, Applied Physics Letters **112** (4), 043104 (2018).
9. D. Nicholls, J. Lee, H. Amari, A. J. Stevens, B. L. Mehdi, and N. D. Browning, Nanoscale **12** (41), 21248 (2020).
10. D. Nicholls, J. Wells, A. Stevens, Y. Zheng, J. Castagna, and N. D. Browning, Ultramicroscopy **233**, 113451 (2022).
11. A. Velazco, A. Béché, D. Jannis, and J. Verbeeck, Ultramicroscopy **232**, 113398 (2022).
12. D. Jannis, A. Velazco, A. Béché, and J. Verbeeck, Ultramicroscopy, 113568 (2022).
13. A. Ishizuka, M. Hytch, K. Ishizuka, *Microscopy* **66**, 217-21 (2017).
14. K. Reidy, G. Varnavides, J. D. Thomsen, A. Kumar, T. Pham, A. M. Blackburn, P. Anikeeva, P. Narang, J. LeBeau, F. M. Ross, *Nature Communications* **12**, 1290 (2021).
15. X. Ke, M. Zhang, K. Zhao, D. Su, *Small Methods* **6**, 2101040 (2022)

16 D. L. Donoho, IEEE Transactions on information theory **52** (4), 1289 (2006).
17 E. J. Candès, J. Romberg, and T. Tao, IEEE Transactions on information theory **52** (2), 489 (2006).
18 K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T. W. Lee, and T. J. Sejnowski, Neural computation **15** (2), 349 (2003).
19 R. G. Baraniuk, IEEE signal processing magazine **24** (4), 118 (2007).
20 M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, presented at the Proceedings of the 27th annual conference on Computer graphics and interactive techniques, 2000 (unpublished).
21 M. Fornasier and H. Rauhut, Handbook of mathematical methods in imaging **1**, 187 (2015).
22 K. Engan, S. O. Aase, and J. H. Husoy, presented at the 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258), 1999 (unpublished).
23 M. Aharon, M. Elad, and A. Bruckstein, IEEE Transactions on signal processing **54** (11), 4311 (2006).
24 D. Needell and R. Vershynin, Foundations of Computational Mathematics **9** (3), 317 (2009).
25 D. Needell and J. A. Tropp, Applied and computational harmonic analysis **26** (3), 301 (2009).
26 S. Chen and D. Donoho, presented at the Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers, 1994 (unpublished).
27 J. Paisley and L. Carin, presented at the Proceedings of the 26th annual international conference on machine learning, 2009 (unpublished).
28 S. Sertoglu and J. Paisley, presented at the 2015 23rd European Signal Processing Conference (EUSIPCO), 2015 (unpublished).
29 D. Nicholls, A. Robinson, J. Wells, A. Moshtaghpour, M. Bahri, A. Kirkland, and N. Browning, presented at the ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022 (unpublished).
30 J. Wells, D. Nicholls, A. Robinson, Y. Zheng, J. Castagna, and N. D. Browning, In Preparation.
31 A. Zobelli, S. Y. Woo, A. Tararan, L. H. G. Tizei, N. Brun, X. Li, O. Stephan, M. Kociak, M. Tence, *Ultramicroscopy* **212**, 112912 (2020)
32 J. M. Ede and R. Beanland, *Scientific Reports* **10**, 8332 (2020)
33 X. Li, O. Dyck, S, V. Kalinin, S. Jesse, *Microscopy & Microanalysis* **24**, 623-633 (2018)
34 Z. Saghi, B. Winter, R. Leary, E. Spiecker, C. Ducati, P. Midgley, *Ultramicroscopy* **160**, 230-238 (2016)
35 R. Leary, Z. Saghi, P. A. Midgley, and D. J. Holland, Ultramicroscopy **131**, 70 (2013).
36 J. M. Rodenburg, Advances in Imaging and Electron Physics 150, 87-184 (2008)
37 A. Stevens, H. Yang, W. Hao, L. Jones, C. Ophus, P. D. Nellist, and N. D. Browning, Applied Physics Letters **113** (3), 033104 (2018).
38 A. Horé, and D. Ziou, *IET Image Processing* **7**, 12-24 (2013).
39 Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, *IEEE transactions on image processing* **13**, 600-612 (2004).
40 A. W. Robinson, D. Nicholls, J. Wells, A. Moshtaghpour, A. Kirkland, and N. D. Browning, Ultramicroscopy **242**, 113625 (2022).
41 I. Lobato and D. Van Dyck, Ultramicroscopy **156**, 9 (2015).

[42] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **5** (2), 354 (2012).

[43] O. Sidorov and J. Yngve Hardeberg, presented at the Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019 (unpublished).

[44] R. Wong, Z. Zhang, Y. Wang, F. Chen, and D. Zeng, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **13**, 4369 (2020).

[45] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C. W. Lin, Neural Networks **131**, 251 (2020).

[46] X. Mao, C. Shen, and Y. B. Yang, Advances in neural information processing systems **29** (2016).

[47] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, arXiv preprint arXiv:1803.04189 (2018).

[48] J. Batson and L. Royer, presented at the International Conference on Machine Learning, 2019 (unpublished).

[49] W. Denk and H. Horstmann, PLoS biology **2** (11), e329 (2004).

[50] L. Holzer, F. Indutnyi, P. H. Gasser, B. Münch, and M. Wegmann, Journal of microscopy **216** (1), 84 (2004).

[51] B. L. Mehdi, A. Stevens, L. Kovarik, N. Jiang, H. Mehta, A. Liyu, S. Reehl, B. Stanfill, L. Luzi, and W. Hao, Applied Physics Letters **115** (6), 063102 (2019).

[52] S. Reehl, B. Stanfill, M. Johnson, D. Ries, N. D. Browning, B. L. Mehdi, and L. Bramer, Quality Engineering **32** (2), 244 (2020).

[53] B. A. Stanfill, S. M. Reehl, M. C. Johnson, N. D. Browning, B L.. Mehdi, P. C. Caragea, and L. M. Bramer, ChemCatChem **10** (14), 3115 (2018).

a Raster Scan

1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18
19	20	21	22	23	24	25	26	27
28	29	30	31	32	33	34	35	36
37	38	39	40	41	42	43	44	45
46	47	48	49	50	51	52	53	54
55	56	57	58	59	60	61	62	63
64	65	66	67	68	69	70	71	72
73	74	75	76	77	78	79	80	81

b Lower Magnification Scan

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

c Space Filling Random Scan

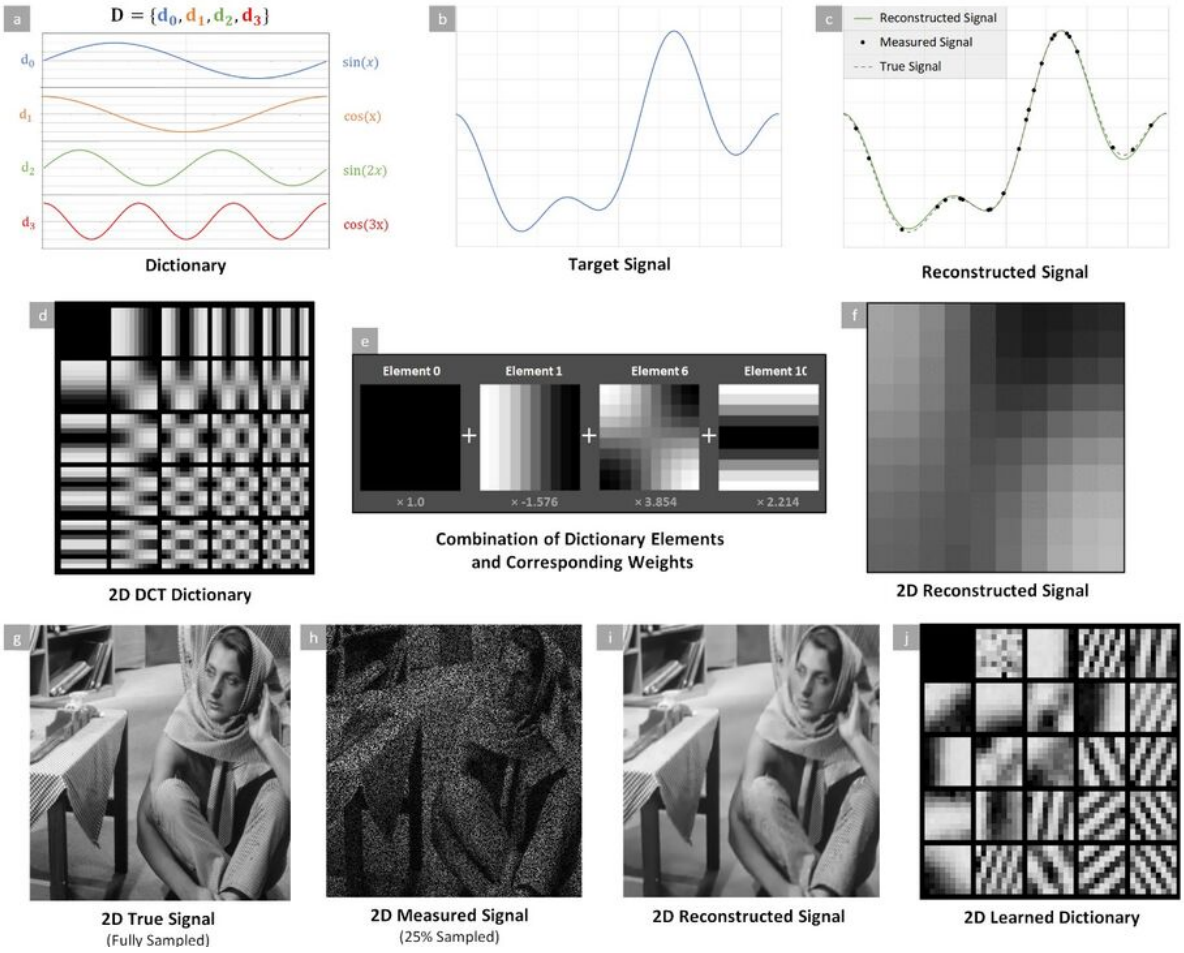
64	54	33	48	78	19	55	15	4
65	32	34	38	68	36	73	57	63
79	25	42	45	17	39	9	27	60
21	22	66	53	43	59	24	74	40
77	62	7	10	12	80	28	49	3
20	14	1	44	72	16	5	37	31
76	26	51	18	35	56	29	75	23
2	47	70	71	6	81	61	52	30
41	13	8	67	46	69	50	58	11

d Subsampled Random Scan

					19		15	4
	25			17		9	27	
21	22					24		
		7	10	12				3
20	14	1			16	5		
	26		18					23
2				6				
	13	8						11

e Line Hop Scan

			4					
1		3		5		7		9
	2				6		8	
10						16		18
	11		13		15		17	
		12		14				
19								27
	20		22		24		26	
		21		23		25		

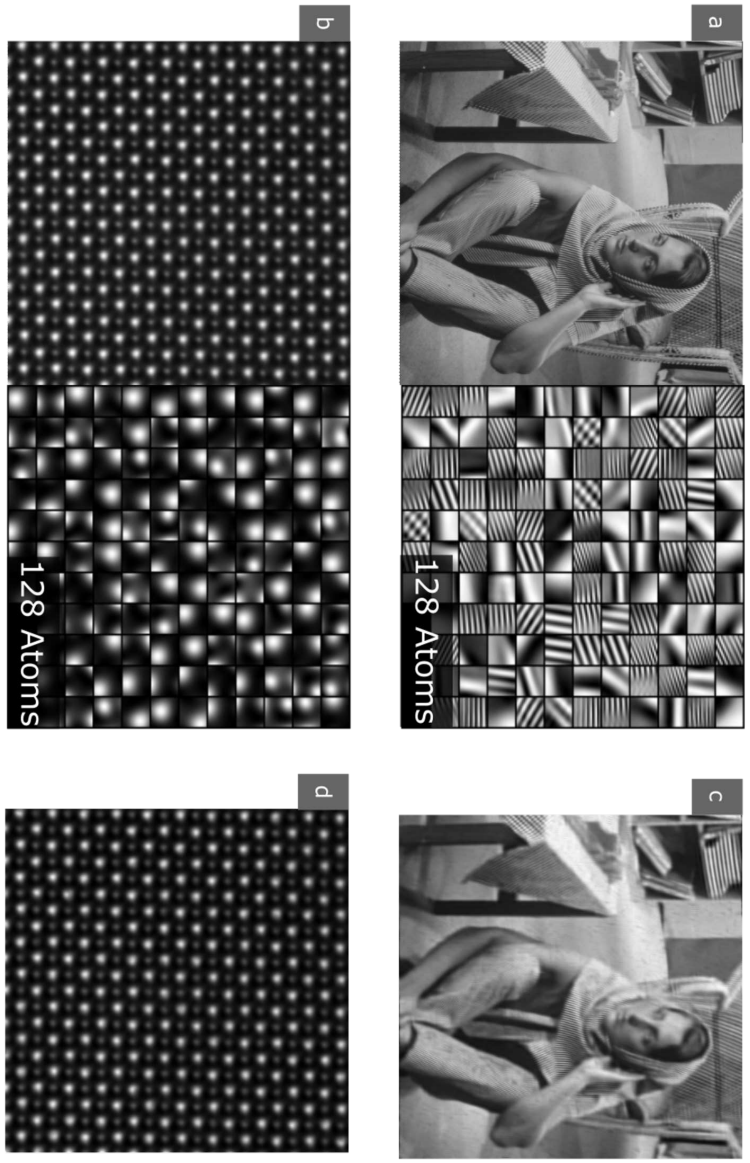


AIP Publishing

Applied Physics Letters

ACCEPTED MANUSCRIPT

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0135245

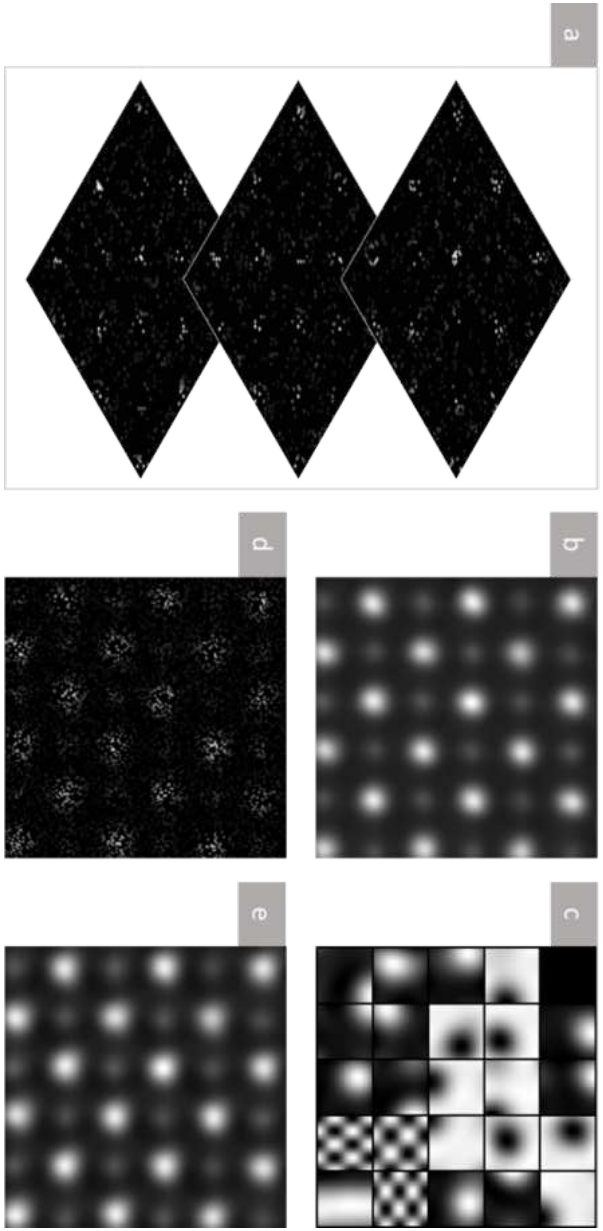


AIP Publishing

Applied Physics Letters

ACCEPTED MANUSCRIPT

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0135245



AIP Publishing

Applied Physics Letters

ACCEPTED MANUSCRIPT

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0135245