

<https://doi.org/10.1038/s41524-025-01766-4>

BEACON—automated aberration correction for scanning transmission electron microscopy using Bayesian optimization



Alexander J. Pattison¹, Stephanie M. Ribet¹, Marcus M. Noack², Georgios Varnavides^{1,3}, Kunwoo Park⁴, Earl J. Kirkland⁵, Jungwon Park^{4,6,7,8}, Colin Ophus^{9,10} & Peter Ercius¹ ✉

Aberration correction is an important aspect of modern high-resolution scanning transmission electron microscopy. Most methods of aligning aberration correctors require specialized sample regions and are unsuitable for fine-tuning aberrations without interrupting on-going experiments. Here, we present an automated method of correcting first- and second-order aberrations called BEACON, which uses Bayesian optimization of the normalized image variance to efficiently determine the optimal corrector settings. We demonstrate its use on gold nanoparticles and a hafnium dioxide thin film showing its versatility in nano- and atomic-scale experiments. BEACON can correct all first- and second-order aberrations simultaneously to achieve an initial alignment and first- and second-order aberrations independently for fine alignment. Ptychographic reconstructions are used to demonstrate an improvement in probe shape and a reduction in the target aberration.

Aberration correction is an important component of high-resolution scanning transmission electron microscopy (STEM). Modern multipole aberration correctors effectively correct for the main spherical aberration (C_3 in round electron lenses) but add smaller parasitic aberrations that also need to be corrected^{1–5}. This is typically accomplished using an automated program based on the analysis of Ronchigrams (for example⁶) or a (CEOS) probe tableau to estimate aberration coefficients from a set of images at different defocii and beam tilts. The program then modifies multipole lenses to compensate for the measured aberrations^{7–10}. Recognizable features in Ronchigrams (such as streaks) can also be used to manually correct lower-order aberrations such as 2-fold astigmatism (A_1) and axial coma (B_2). Microscope and corrector lens settings drift on the minutes to hours time scales and fine-tuning during an experiment is necessary to achieve optimal performance¹¹. Methods based on the Ronchigram require a flat amorphous sample region (e.g. a thin carbon film), which can be difficult to find on monolithic crystalline films. For the probe tableau method, a special sample

is required and is typically only used to tune the corrector at the start of an experiment. A general method for tuning aberrations that can be used on a wide range of samples is therefore desirable. Ideally, such a method would be fully automated and apply the least amount of beam dose possible to minimize time, damage (especially to beam-sensitive samples), and carbon contamination of a region of interest.

Recently, there has been promising progress in the development of new methods for automating aberration correction. DeepFocus uses a convolutional neural network to calculate the defocus and astigmatism in SEM images from two images at different focus values¹². Convolutional neural networks have also been used to extract various metrics from Ronchigrams that can be used to correct first- and second-order aberrations in STEM, such as beam emittance growth^{13,14} and Strehl ratio¹⁵, as well as directly assess aberrations from the Ronchigrams themselves¹⁰. In ref. 1, Kirkland demonstrated via simulation that the normalized variance (σ^2/μ^2) of a high-angle annular dark field (HAADF-) STEM image increases when the

¹Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ²Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ³Miller Institute for Basic Research in Science, University of California, Berkeley, CA, USA. ⁴School of Chemical and Biological Engineering, Institute of Chemical Processes, Seoul National University, Seoul, Republic of Korea. ⁵School of Applied and Engineering Physics, Cornell University, Ithaca, NY, USA. ⁶Center for Nanocrystal Research, Institute for Basic Science (IBS), Seoul, Republic of Korea. ⁷Institute of Engineering Research, College of Engineering, Seoul National University, Seoul, Republic of Korea. ⁸Advanced Institute of Convergence Technology, Seoul National University, Suwon, Republic of Korea. ⁹Department of Materials Science and Engineering, Stanford University, Stanford, CA, USA. ¹⁰Precourt Institute for Energy, Stanford University, Stanford, CA, USA. ✉e-mail: percius@lbl.gov

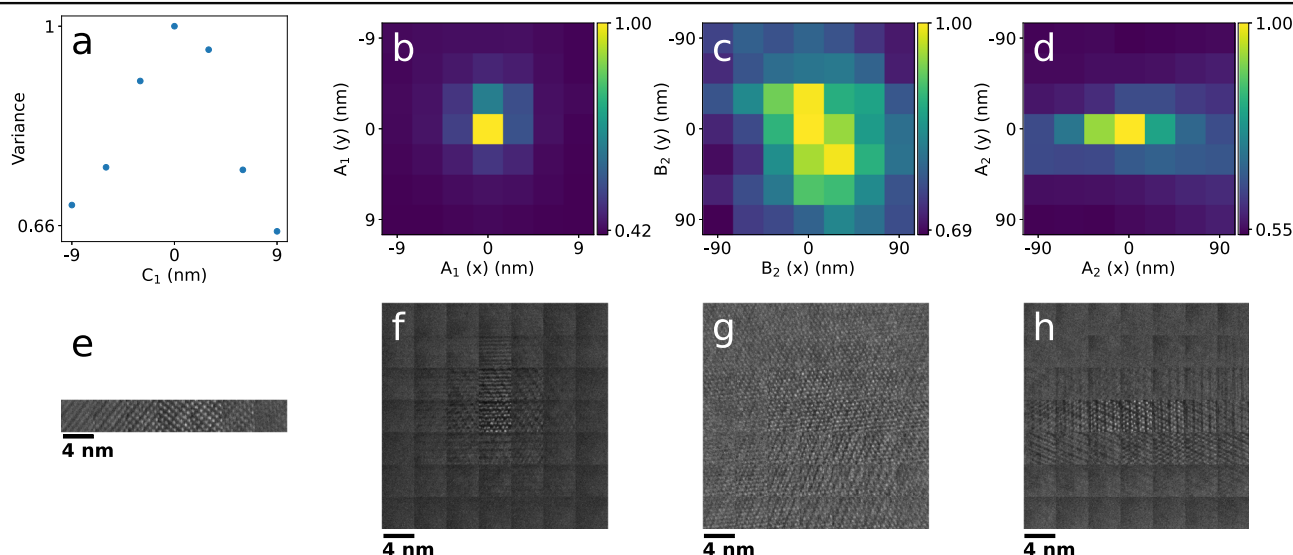


Fig. 1 | Grid searches of individual aberrations on the HfO₂ thin film. Grid searches of C_1 (a, e), A_1 (b, f), B_2 (c, g), and A_2 (d, h) for the HfO₂ thin film. a–d Image NV normalized to the maximum value. e–h Tiled cutouts (4.3 nm

field of view) of the corresponding full real space images. The zero aberration value was the initial well-corrected state as determined by the CEOS alignment software.

magnitude of first, second and third-order aberrations decreases. Ishikawa et al. experimentally demonstrated the same relationship using image standard deviation (σ) to correct first- and second-order aberrations¹⁶. Together, these works demonstrate that adjusting aberrations to maximize σ or σ^2/μ^2 can effectively correct aberrations without directly measuring them. Kirkland used a simplex method to iteratively determine the set of aberration coefficients that maximizes σ^2/μ^2 using STEM image simulations¹, while Ishikawa et al. developed a system that maximized σ for each aberration coefficient sequentially¹⁶. A similar concept is utilized by companies such as Thermo Fisher in their OptiSTEM+ software and has been shown to improve microscope alignment for atomic resolution imaging.

In this paper, we present BEACON—Bayesian-Enhanced Aberration Correction and Optimization Network—which uses Bayesian optimization to autonomously correct first- and second-order aberrations, both individually and collectively. Bayesian optimization can efficiently estimate unknown functions with a reduced set of measurements, making it useful in situations where each measurement is expensive in terms of time, cost and/or, in this case, beam dose. It is used for autonomous experiments in synchrotrons¹⁷, fMRI studies¹⁸ and scanning-probe microscopy¹⁹. We employ the metric of normalized variance (NV), proposed by Kirkland for HAADF-STEM and others for SEM, as a metric for image quality^{20–25}. We demonstrate the use of BEACON at atomic- and nanoscale resolution using a polycrystalline hafnium dioxide (HfO₂) thin film and a sample composed of 5 nm gold nanoparticles (Au NP) on a carbon support to show the method’s viability in different experimental scenarios. We demonstrate that BEACON efficiently compensates for aberrations, leading to higher quality and higher resolution images. Further, we show through ptychographic reconstruction of pre- and post-corrected data that BEACON indeed reduces the target aberrations. The simplicity and robustness of BEACON makes it useful for fine-tuning aberration corrector alignment in low- and high-resolution experiments without human intervention.

Results

Verifying image NV relationship

To verify the correlation between the image NV and aberration coefficient magnitude for both samples, we systematically varied corrector aberration values in small steps (called a “grid search”) and acquired HAADF-STEM images for each of the first- and second-order aberrations: defocus (C_1), two-fold astigmatism (A_1), axial coma (B_2), and three-fold astigmatism

(A_2). Figure 1 (HfO₂ thin film) and Fig. 2 (Au NPs) demonstrate that the image NV (top row) and visual image quality (bottom row) both vary inversely with the aberration coefficient magnitude^{16,26}.

Individual aberrations

We initially tested BEACON on each first- and second-order aberration individually. We started from a corrected state before applying a known aberration value and running BEACON for a total of 35 iterations. Unless otherwise specified, we always initialized the model with 5 randomly chosen points followed by a specified number of iterations using the upper confidence bound (UCB) algorithm (see Section “Bayesian Optimization”). Images were taken before and after the runs to verify the results. The middle row of Fig. 3 shows the final model overlaid with each experimental measurement (colored dots). BEACON finds the peak of the surrogate model (background image), while the top and bottom rows show pre- and post-correction HAADF-STEM images of the HfO₂ thin film, respectively, and demonstrate a visual improvement in the image quality. Supplementary Fig. 1 shows the same results for the Au NP sample.

To quantitatively assess the performance and reliability of BEACON, we did 5 runs on each sample starting from the same initial aberration settings. We evaluated the correction precision by calculating the standard error of all final aberration coefficients compared to the mean, as a measure of self-consistency (see Table 1). We also assessed the convergence of BEACON by measuring the difference (ΔMM) between the final maximum of the surrogate model and the maximum of the surrogate model after every iteration, then calculating the standard deviations of these differences over the 5 runs (see Fig. 4 and Supplementary Fig. 2). We set benchmarks for a good corrector alignment as ± 1 nm for first-order corrections (C_1 , A_1) and ± 20 nm for second-order corrections (B_2 , A_2). We defined the number of iterations needed for convergence as the iteration at which the standard deviation of the differences fell below these benchmarks (see Table 1).

Combined aberrations

In the previous section, we corrected each aberration separately with a maximum of two dimensions. However, Bayesian optimization can handle higher dimensional problems and each image NV provides information about multiple aberrations simultaneously.

We grouped first- (C_1 , A_1) and second-order (B_2 , A_2) aberrations because they affect image NV similarly, simplifying the procedure. Figure 5 shows the C_1 - A_1 optimization (three-dimensional) before (Fig. 5a) and after

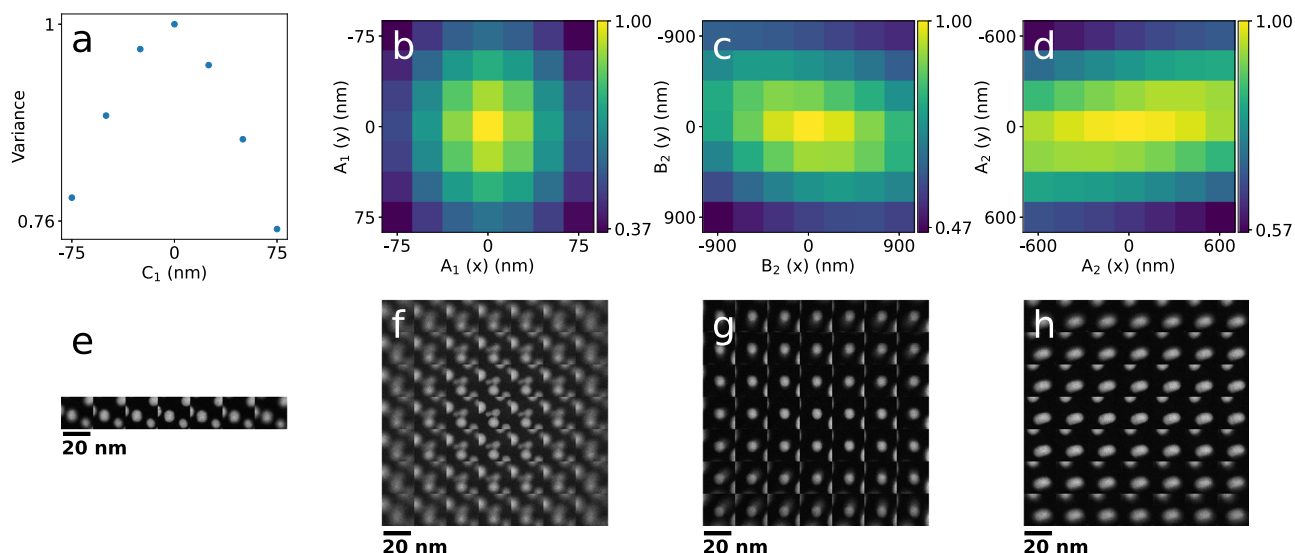


Fig. 2 | Grid searches of individual aberrations on the Au NP sample. Grid searches of C_1 (a, e), A_1 (b, f), B_2 (c, g), and A_2 (d, h) for the Au NP sample. a–d Image NV normalized to the maximum value. e–h Tiled cutouts (24.4 nm

field of view) of the corresponding full real space images. The zero aberration value was the initial well-corrected state as determined by the CEOS alignment software.

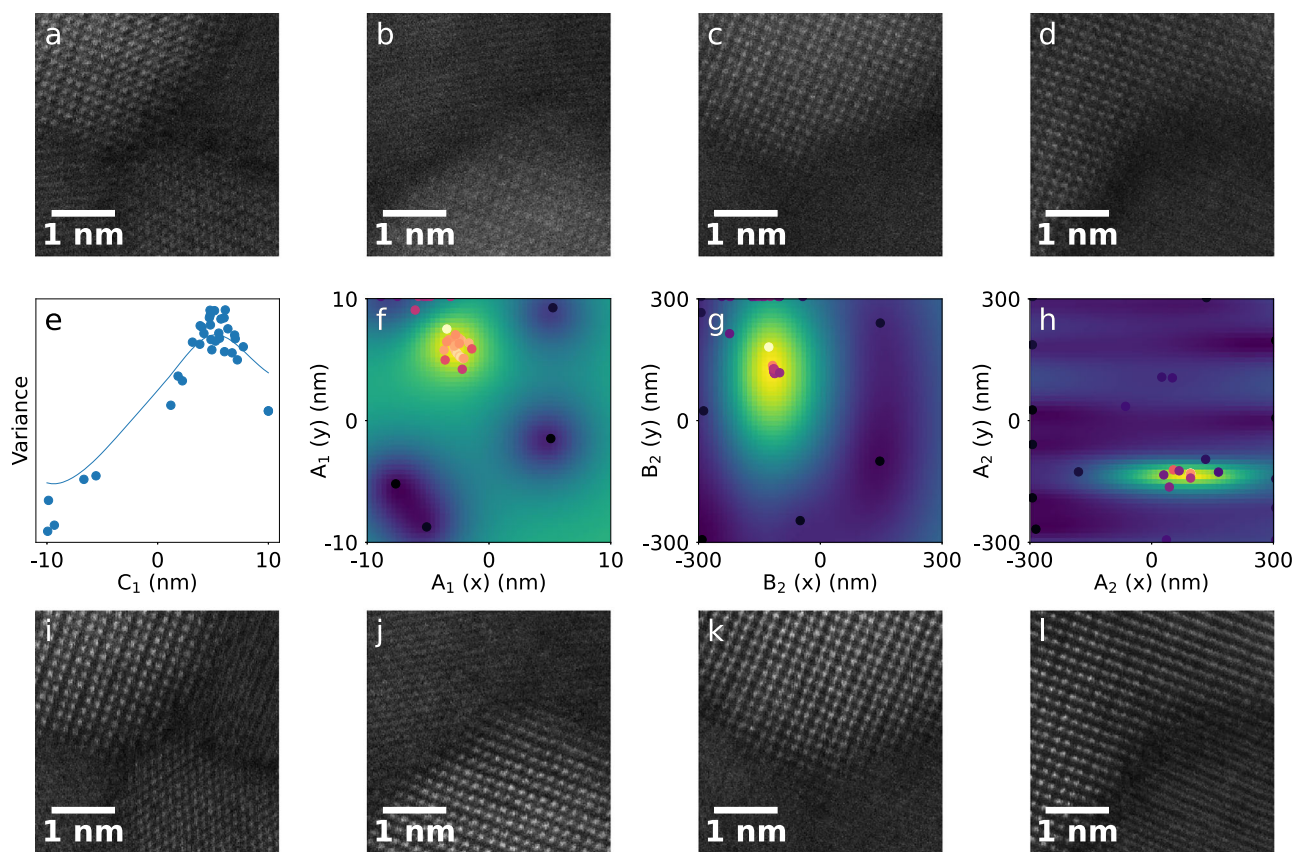


Fig. 3 | Correction of individual aberrations. BEACON corrections of C_1 (offset = 5 nm) (a, e, i), A_1 (x offset = -5 nm, y offset = 5 nm) (b, f, j), B_2 (x offset = -150 nm, y offset = 150 nm) (c, g, k), and A_2 (x offset = 100 nm, y

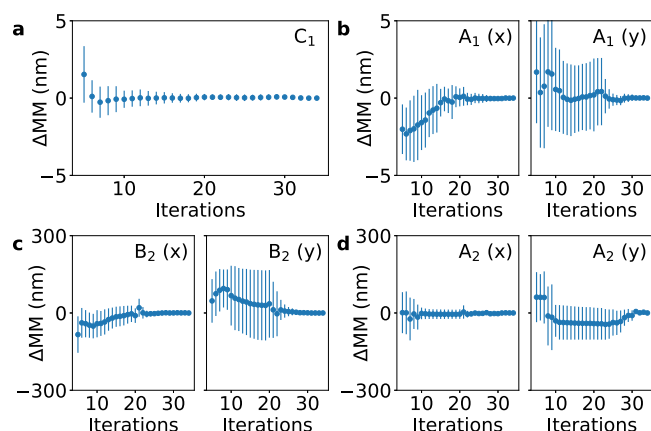
offset = -150 nm) (d, h, l) aberrations on HfO_2 sample. a–d Images before correction. e–h Surrogate models with sampled points marked by dots. i–l Images after correction.

(Fig. 5c) 65 iterations using the HfO_2 thin film sample. Figure 5b shows the three-dimensional volumetric plot of the surrogate model of aberration space (estimated NV for each aberration set). A second optimization run for combined B_2 - A_2 with 85 iterations is shown in Fig. 6. The aberration space plots in Fig. 6c–f) are three-dimensional plots obtained by fixing one of the

aberration coefficients to its optimum value, enabling the visualization of the four-dimensional model. Equivalent low-magnification results using the Au NP sample are shown in Supplementary Figs. 3, 4. The convergence results of multiple runs on both samples are presented in Table 2. Figure 7 shows the differences between the model maxima and final model maxima

Table 1 | Correction precision and number of iterations required for convergence for individually corrected aberrations

Corrected aberration	Precision of HfO ₂ measurement (nm)	Convergence iteration for HfO ₂	Precision of Au NP measurement (nm)	Convergence iteration for Au NP
C ₁	0.29	8	1.25	8
A ₁	1.18	24	0.17	13
B ₂	16.18	25	15.04	9
A ₂	24.27	31	40.14	9

**Fig. 4 | Convergence of individual aberration runs.** Plots of the average difference (ΔMM) between the model maximum and the final model maximum during BEACON runs on the HfO₂ thin film for (a) C₁, (b) A₁, (c) B₂ and (d) A₂.

(ΔMM). Equivalent low-magnification results using the Au NP sample are shown in Supplementary Fig. 5.

Finally, we optimized all first- and second-order aberrations simultaneously using 150 iterations initialized with 20 random points. Figure 8 shows the results of these corrections on the HfO₂ thin film sample. Equivalent low-magnification results using the Au NP sample are shown in Supplementary Fig. 6. Table 3 includes the precision of these corrections but does not list the number of iterations needed for convergence as none of the standard deviations of the differences from the final model maximum fell below our benchmarks. Figure 9 shows the differences between the model maxima and final model maxima (ΔMM). Equivalent low-magnification results using the Au NP sample are shown in Supplementary Fig. 7. These were only saved every 10 iterations due to the large computation incurred to calculate the full 7-dimensional surrogate model.

Discussion

The before and after images in Figs. 3, 5, 6, 8 show qualitative visual improvement in image quality for every optimization performed on the HfO₂ sample. Supplementary Figures 1, 3, 4, 6 show similar improvements for the Au NP sample.

Table 1 indicates that optimizations of individual aberrations reached below the aforementioned benchmarks within a reasonable number of iterations. The one exception is that the precision of the A₂ correction on the Au NP sample was roughly double the benchmark. In our experience, this still constitutes an acceptable correction of A₂ which is typically corrected to < 50 nm using the CEOS alignment software. With the exception of C₁, convergence was slower for the HfO₂ as compared to the Au NP sample. We attribute this to the different noise levels, which can be seen by comparing the scatter of NV measurements for different C₁ values in Fig. 3 and Supplementary Fig. 1. We also note that the two samples required different noise functions (see Section “BEACON operation”). However, the number of iterations required to converge is relatively small for both samples, and BEACON achieves a corrected state with fewer images as compared to the traditional tableau method. Table 4 shows the time per iteration and full run time including additional overhead (normalization, plotting, etc.) for each

type of optimization. For comparison, acquisition of a single tableau using the CEOS alignment software used for correcting higher-order aberrations requires 44 images. Each tableau takes 68 seconds and multiple passes are often required to correct second-order aberrations to within our benchmarks.

Table 2 shows that the precision of the C₁-A₁ corrections was similar to individual C₁ and A₁ corrections for both samples. The precision of B₂-A₂ corrections is better than the individual B₂ and A₂ corrections for the Au NP sample but worse for the HfO₂. In particular, the precision of the A₂ corrections is three times our benchmark. Additionally, the number of iterations required to converge is larger than the sum of the two individual aberrations in all cases except the C₁-A₁ correction on the Au NPs. Here, the number of iterations required for C₁ and A₁ individually (21) is similar to the number required for the combination (20).

Table 3 shows that the precision of C₁-A₁-B₂-A₂ corrections was comparable to individual optimizations for the HfO₂ thin film, with the exception of C₁. However, the precision of corrections using this method on the Au NP sample is worse than corrections of each aberration individually or grouped by order. Furthermore, the standard deviation of the model maxima between different runs does not converge to below 1 nm for C₁ and A₁ or 20 nm for B₂ and A₂ for either sample. It is therefore more efficient to either correct the aberrations individually or grouped by order.

Overall, the results indicate that optimizing aberrations individually or grouped by order is more efficient and gives more precise corrections than optimizing all first- and second-order aberrations simultaneously. Still, correcting all first- and second-order aberrations collectively may be useful in certain circumstances. For example, as shown in Fig. 9, the combined C₁-A₁-B₂-A₂ optimization successfully improves the alignment, but it fails to converge in a reasonable amount of time. Such a coarse alignment could be followed by fine correction of each aberration individually. Additionally, further hyperparameter optimization may improve the performance of the combined correction method.

As an independent verification of the BEACON method, we used ptychography to estimate aberrations from datasets with intentionally applied aberrations. Electron ptychography is a phase retrieval technique^{27–33} that attempts to solve for both the object and the probe from a set of 2D diffraction patterns acquired at a set of probe positions, leading to the common term of 4D-STEM³⁴. Ptychographic reconstructions rely on careful control of the electron beam shape in order to acquire redundant information from the sample. Most often defocus is added to the incident probe to facilitate real space overlap between adjacent probe positions. It is also possible to correct for other aberrations based on the retrieved phase of the probe³⁵.

In this experiment, we added A₁ or B₂ to a well-corrected state before acquiring a 4D-STEM scan. We subsequently used BEACON to correct the aberrations and acquired a second dataset. We reconstructed both datasets using electron ptychography. Although it is possible to solve for an unknown probe with iterative ptychography, reconstructions converge more quickly with a good initial guess, which includes aberrations. Using phase retrieval methods implemented in py4DSTEM, we typically estimate the probe aberrations with either a tilt-corrected BF-STEM (or parallax imaging) reconstruction^{33,36,37}, or using an independent Bayesian optimization algorithm³³. The Bayesian optimization algorithm optimizes over self-consistency error of the ptychographic reconstruction, which in this case was used to estimate the initial aberrations in the probe. These fit aberration guesses provide

Fig. 5 | Correction of combined C_1 - A_1 aberrations. BEACON correction of combined C_1 (offset = -5 nm) and A_1 (x offset = 5 nm, y offset = -5 nm) on the HfO_2 thin film. Magnified portion of the full image **a** before correction and **c** after correction. **b** Volumetric plot of the surrogate model.

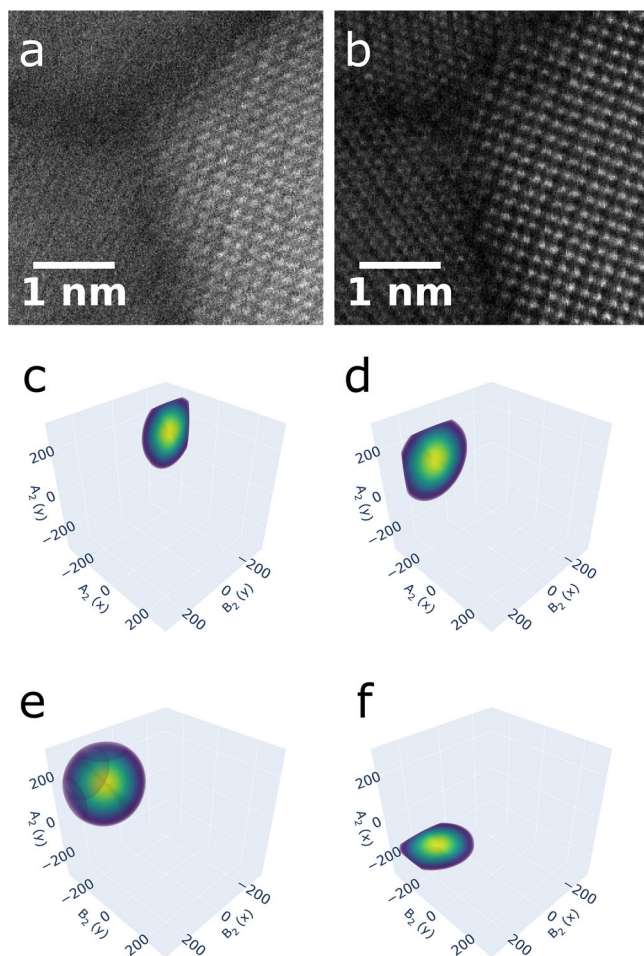
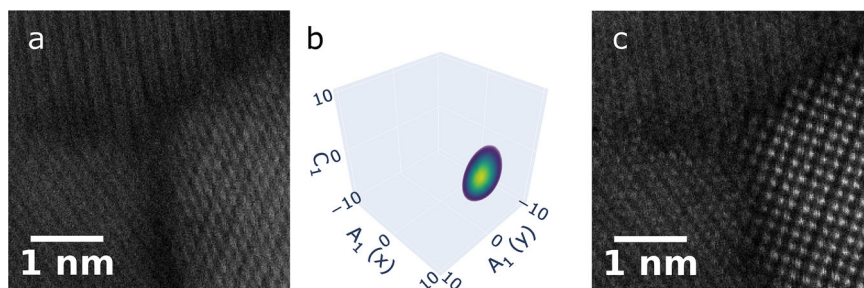


Fig. 6 | Correction of combined B_2 - A_2 aberrations. BEACON correction of combined B_2 (x offset = 150 nm, y offset = -150 nm) and A_2 (x offset = -150 nm, y offset = 150 nm) aberrations on the HfO_2 thin film sample. Magnified portion of the full image **a** before and **b** after correction. **c-f** Three-dimensional volumetric plots of the four-dimensional surrogate model. Each plot was generated using the optimum value for **c** $B_{2,x}$, **d** $B_{2,y}$, **e** $A_{2,x}$, **f** $A_{2,y}$.

an independent assessment of BEACON's ability to correct aberrations beyond image NV and visual image quality.

Figure 10 shows the results of our reconstructions from ptychography experiments with Au NPs. Our initial probe for the first particle had large C_1 and A_1 , as illustrated in Fig. 10a. The second column shows our second scan of the same particle after A_1 aberration correction by BEACON. Our estimate is that A_1 is reduced from 12 nm to 2 nm, as estimated from Bayesian optimization of probe parameters. The last two columns in Fig. 10 show a similar experiment with B_2 instead of A_1 . In this case, the probe begins with approximately 554 nm of B_2 , which is decreased to 224 nm. Real space

probes are shown in Supplementary Fig. 8. We note that aberration coefficients are defined differently by different groups and the CEOS coefficients used by BEACON are not expected to match the conventions used by py4DSTEM. While large aberration values would interfere with conventional imaging, the ptychographic approach yields high-quality reconstructions with clearly resolved atomic features and an estimate of the probe shape for all scans. Some small differences between subsequent images of the sample particle can be due to beam damage and carbon contamination. Comparing the fast Fourier transforms (FFT) between these two scans (Fig. 10i-j, k-l), we see similar spatial resolution but increased Thon ring contrast with the BEACON-corrected probes. An estimate for the improvement in Thon ring uniformity is shown in Supplementary Table 1.

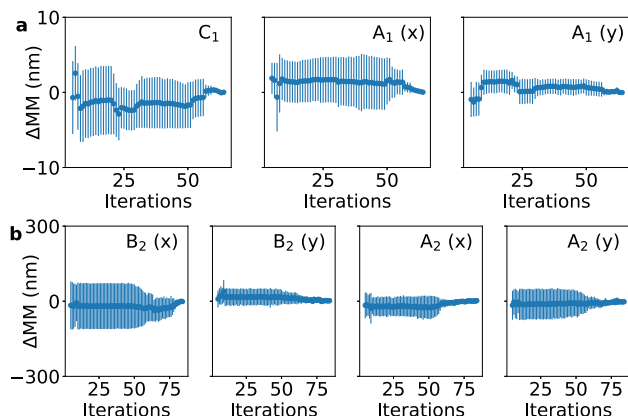
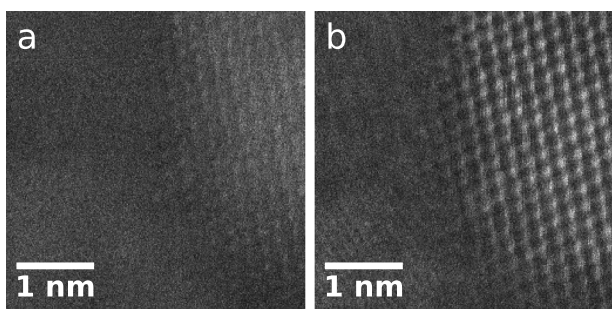
It has been suggested that increasing the phase diversity of the probe, through physical devices or aberration modulation, can improve information in ptychography reconstructions^{35,38-40}. However, it is often preferred to remove (especially unknown) higher-order aberrations from the incident beam. In this experiment, there was minimal impact on information transfer in the ptychographic reconstruction from the aberrated probe, but a clear reduction in the target aberration as evidenced by the final probe shape. Our ptychographic reconstructions highlight that BEACON lends itself especially well to the tuning of aberrations for many types of experiments, including conventional dark-field imaging and phase contrast experiments.

Although this work demonstrates a reliable automated aberration correction method that works on different samples, several factors affect its behavior and performance. Since BEACON relies solely on image NV, it is difficult to use in situations where image NV changes independent of aberration values. We found that imaging different sample regions (i.e. due to sample drift or beam shift) reduced the method's reliability. BEACON incorporates a method based on cross-correlation to ensure the same region is imaged in every iteration. Further, electron beam-induced carbon contamination also tends to reduce the sensitivity of the method. Reducing the applied dose by using a lower magnification or lower beam current can mitigate this, but a clean sample is always preferred.

Additionally, the sample geometry and specific region chosen can affect the behavior and performance. The Au NP sample has many sharp edges and relatively large image contrast. The image NV varies above the noise level over a broad range of aberration magnitudes at low magnifications. Other nanoparticle samples should behave similarly. At atomic resolution, small changes in the imaged region unexpectedly resulted in relatively large changes in NV. Normalization (see equation (1)) is designed to compensate for this effect, but we found that using the same imaged region (identified by cross correlation) was an important feature to include in BEACON. Conversely, thin films such as the HfO_2 sample are less sensitive to this effect, because the intensity is relatively uniform across the entire sample. However, for this sample geometry, the NV varies significantly only when atom columns or fringes are resolvable. Once atomic or lattice resolution is lost, NV is insensitive to aberration magnitude. Therefore, when the parameter bounds of the optimization are large and/or the optimization was performed far from the optimum condition, BEACON struggled to converge. In these cases, the NV peak is very narrow (approximating a Dirac delta function)

Table 2 | Precision and number of iterations required for convergence for combined first- and combined second-order aberration correction

Corrected aberration	Precision of HfO ₂ measurement (nm)	Convergence iteration for HfO ₂	Precision of Au NP measurement (nm)	Convergence iteration for Au NP
C ₁ -A ₁	C ₁ : 0.64, A ₁ : 0.51	C ₁ : 57, A ₁ : 57	C ₁ : 1.29, A ₁ : 0.31	C ₁ : 20, A ₁ : 20
B ₂ -A ₂	B ₂ : 25.58, A ₂ : 60.06	B ₂ : 69, A ₂ : 49	B ₂ : 11.21, A ₂ : 12.54	B ₂ : 36, A ₂ : 71

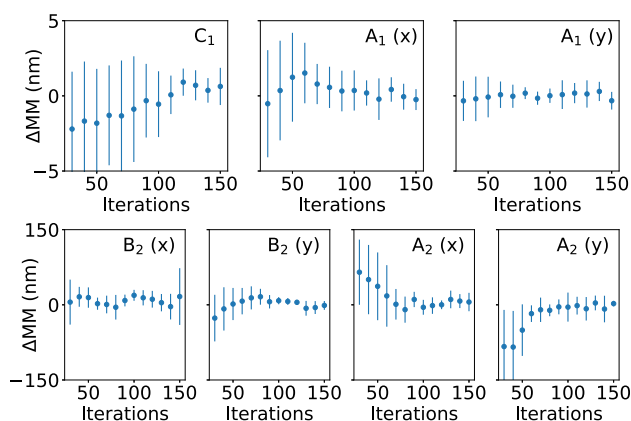
**Fig. 7 | Convergence of combined aberration runs.** Plots of the average difference (ΔMM) between the model maximum and the final model maximum during optimization runs on HfO₂ thin film for **a** C₁-A₁ correction and **b** B₂-A₂ correction.**Fig. 8 | Correction of combined C₁-A₁-B₂-A₂ aberrations.** BEACON correction of combined C₁ (offset = -5 nm), A₁ (x offset = -5 nm, y offset = 5 nm), B₂ (x offset = -150 nm, y offset = 150 nm) and A₂ (x offset = -150 nm, y offset = 150 nm) aberrations on HfO₂ thin film sample. Zoomed portion of the full image **a** before and **b** after correction.

compared to the parameter bounds. Tailoring the parameter bounds to approximately match the shape of the NV peak (which is sample-dependent) is essential to convergence. Further, samples with low intensity contrast in the field of view such as blank carbon films or amorphous thin-film materials cannot be used for this method. Lastly, single crystal samples which are tilted off-axis might also produce an erroneous result, especially for B₂.

Finally, the choice of noise estimation function was an important aspect of this work. Corrections performed on HfO₂ were relatively noisier when using the noise estimator designed for the Au NP sample. This was because the SNR of the HfO₂ NV measurements was intrinsically lower and the noise was amplified by the normalization procedure. Thus, runs on the HfO₂ sample used Eq. (5) and runs on the Au NP sample used Eq. (6). We find that Eq. (6) is an effective noise function for samples where the SNR of NV measurements is high, which typically occurs when there is large contrast within the field of view. Equation (5) is more effective for samples where the SNR of NV measurements is low, like for thin films.

Table 3 | Precision of C₁-A₁-B₂-A₂ aberration correction

Corrected aberration	Precision for HfO ₂ (nm)	Precision for Au NPs (nm)
C ₁ -A ₁ -B ₂ -A ₂	C ₁ : 4.44, A ₁ : 1.17, B ₂ : 20.66, A ₂ : 19.97	C ₁ : 5.10, A ₁ : 9.36, B ₂ : 107.10, A ₂ : 147.98

**Fig. 9 | Convergence of combined C₁-A₁-B₂-A₂ runs.** Plots of average difference (ΔMM) between model maximum and final model maximum during optimization runs on HfO₂ thin film for C₁-A₁-B₂-A₂.

To summarize these limitations, BEACON is most successful on clean samples with large contrast variations within the field of view, such as nanoparticles on a carbon film or well-resolved atomic lattices. Conditions where BEACON would fail include samples where carbon contamination builds up between images, beam-sensitive materials where the sample morphology changes under the beam, amorphous materials with low contrast variation and high-index zone axes with small spacings between atoms where the spacings are commensurate with the probe size. BEACON will also fail if the field of view drifts significantly during a single run, such as at extreme magnifications or cryo-cooled samples. Finally, success will also depend on the choice of the noise function used, which may depend on the specific sample characteristics and/or microscope properties.

Kirkland's simulations demonstrated that image NV can be used to optimize third-order aberrations: spherical aberration (C₃), four-fold astigmatism (A₃) and axial star (S₃)¹. Experimental verification proved difficult in our setup. Our software is unable to change C₃, but we note that C₃ and A₁ are coupled. A combined C₃-A₁ optimization could potentially correct both simultaneously essentially "learning" the relationship between the two aberrations. Also, changes to A₃ induced a large shift of the field-of-view which proved difficult to compensate for even with our cross-correlation method. The large image shifts necessary to recenter the field of view induce higher off-axis aberrations making the process self-defeating. The effect of S₃ on image NV is shown in Supplementary Fig. 9, but the effect was so subtle that BEACON was less effective than the CEOS alignment software. Third-order aberrations tend to drift much less than first- and second-order aberrations during experiments, and we thus concentrated on the lower orders in this manuscript.

Methods

Normalized variance (NV)

Normalized variance (NV) is expressed as:

$$f(C_a) = \frac{\sigma(C_a)^2}{\mu(C_a)^2} \quad (1)$$

where $C_a = [C_1, A_1, B_2, A_2, \dots]$ is the vector or set of aberration coefficients, possibly scaled to the tolerance for each order. σ^2 is the variance of the pixel intensities of an image, and μ is the mean pixel intensity, such that

$$\mu = \frac{1}{N} \sum_{ij} z_{ij}(C_a) \quad (2)$$

$$\sigma^2 = \frac{1}{N} \sum_{ij} (z_{ij}(C_a) - \mu(C_a))^2, \quad (3)$$

Table 4 | Iteration times and full run times for each category of optimization

Aberration	Time per iteration (s)	Full run time (s)
C1	1.802 ± 0.031	69.4
A1	1.863 ± 0.040	71.2
B2	2.183 ± 0.079	83.2
A2	2.027 ± 0.048	77.4
C1-A1	3.096 ± 0.110	210.0
B2-A2	3.361 ± 0.414	293.3
C1-A1-B2-A2	3.189 ± 1.729	473.7

where z_{ij} is the pixel intensity at position i, j and N is the total number of pixels¹. Normalization is designed to reduce the sensitivity to specimen drift and changes in beam current. We note that the level of noise affects the image NV but does not vary with changes in aberrations. A greater concern is the build up of carbon during STEM scanning, which decreases the image NV. This can be mitigated by reducing dose by using shorter dwell times, fewer scan points, and minimizing the number of optimization iterations (i.e. the number of images acquired). SEM and HAADF-STEM are both essentially incoherent images and the NV behaves similarly. Traditional BF-CTEM is primarily coherent phase contrast, and NV has not been found to behave as described below. However, incoherent BF-CTEM most likely behaves similarly to HAADF-STEM.

Bayesian optimization

Bayesian optimization is a quick and efficient method of estimating expensive-to-compute unknown functions by minimizing uncertainty⁴¹. It generates a surrogate model as a proxy for the unknown function, usually using Gaussian processes (GP)⁴¹. If we interpret data as a set of noisy function evaluations of a data-generating, inaccessible, ground-truth latent function $f(C_a)$, a GP assumes that a prior normal distribution can be placed over every finite subset of those function evaluations. This prior normal is fully defined by a prior mean and a covariance or kernel function⁴². This prior distribution can be optimized (trained) by maximizing the log marginal likelihood of the observed data. Once training is completed, conditioning on observations yields a posterior probability distribution for the latent function values across the domain. This posterior distribution provides a posterior mean (often considered the surrogate model) and posterior variances (often considered the uncertainties), both of which are used by the acquisition function such that maxima represent optimal new points for data collection. The choice of acquisition function defines the behavior of the optimization routine⁴¹. One commonly used acquisition function for

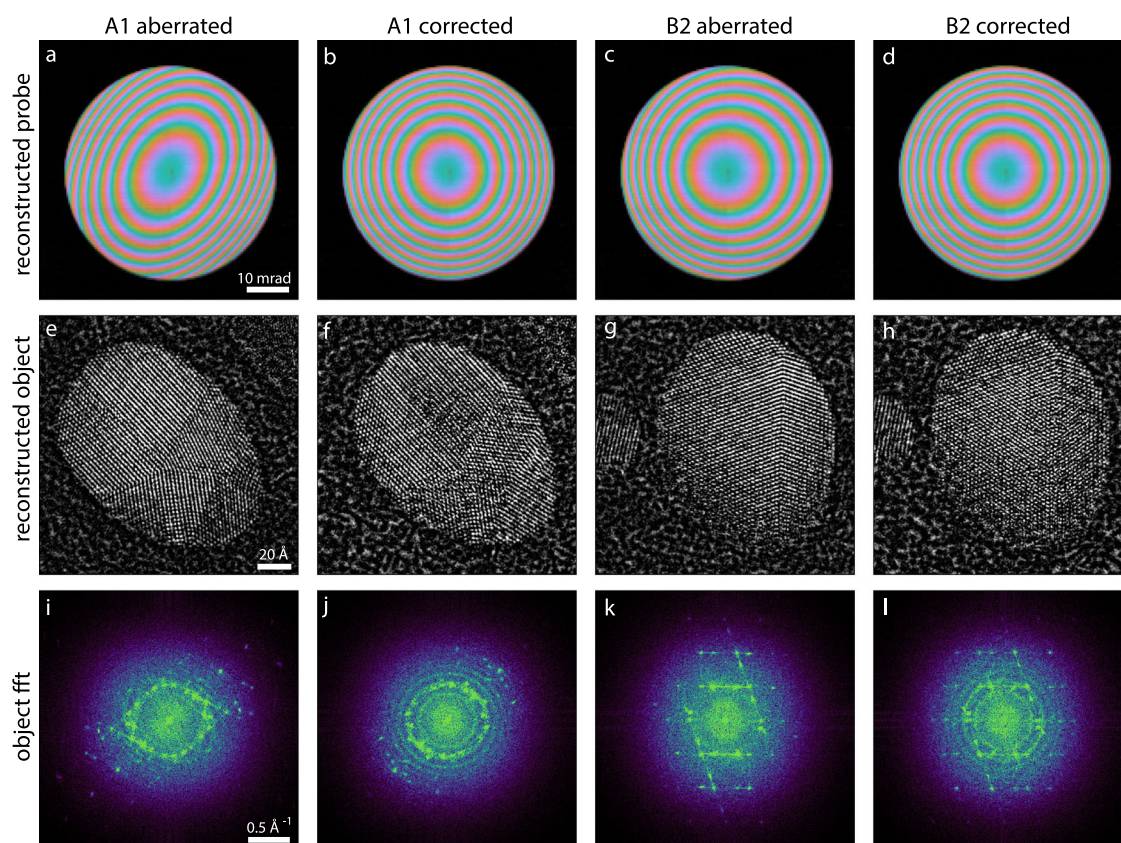


Fig. 10 | Verification by ptychography. Ptychographic reconstruction of two sets of particles: **a–d** probe in reciprocal space, **e–h** reconstructed object, and **i–l** object FFTs. Comparing **a** to **b** and **c** to **d** we observe reduction in aberrations after BEACON correction.

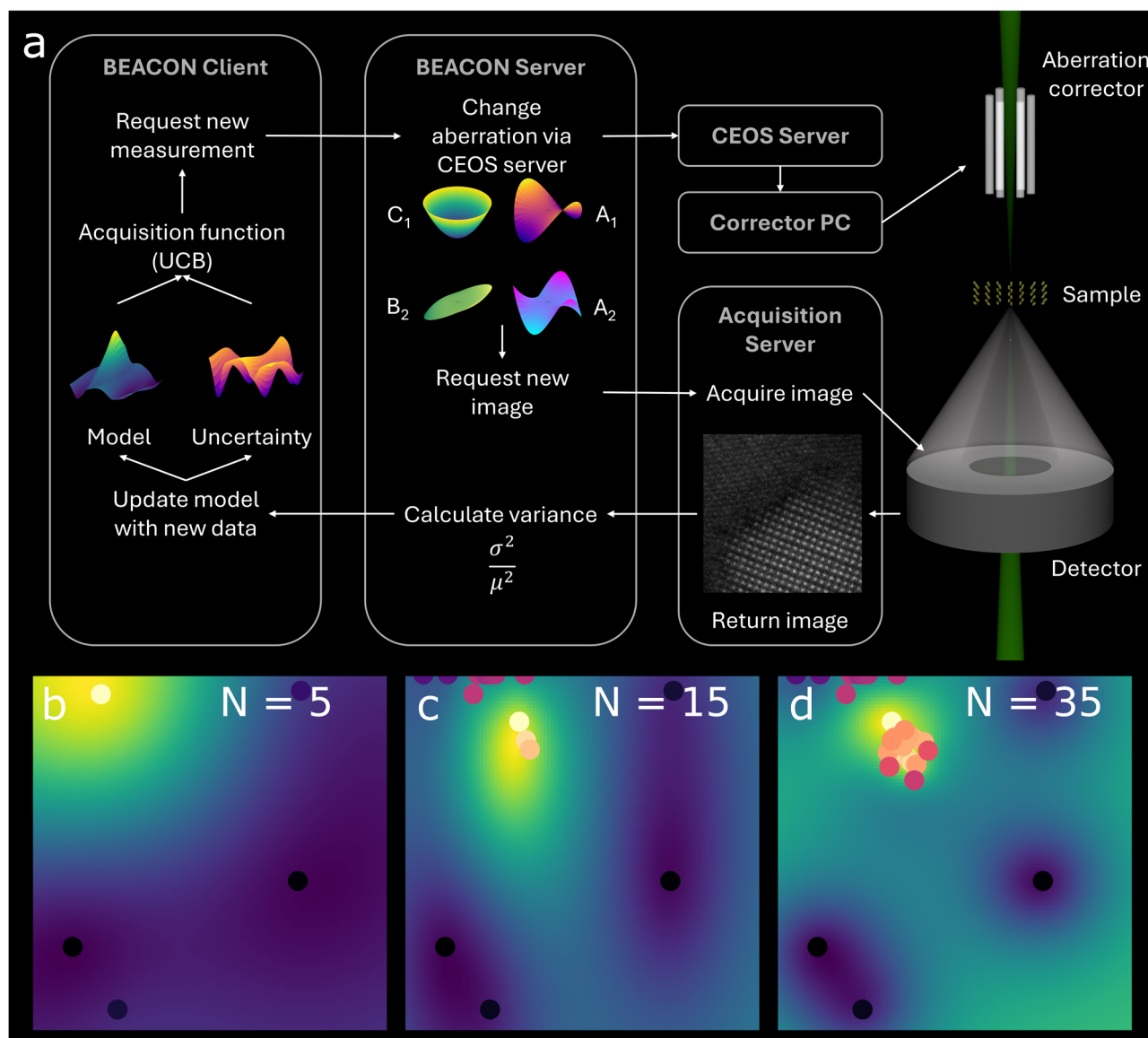


Fig. 11 | Schematic of BEACON workflow. **a** Software schematic of aberration correction workflow. Evolution of an A_1 surrogate model after **b** 5 iterations, **c** 15 iterations and **d** 35 iterations.

locating the maximum of an unknown function is the upper confidence bound (UCB) method, where the next measurement point is chosen by the maximum of the UCB. This is mathematically represented by:

$$C_a^* = \arg \max_{C_a} [m_{GP}(C_a) + c\sigma_{GP}(C_a)], \quad (4)$$

where C_a^* is the next sample point in aberration space, $m_{GP}(C_a)$ is the GP posterior mean, which is the predicted image quality (i.e. the image quality at aberration setting C_a), $\sigma_{GP}(C_a)$ is the posterior uncertainty of the posterior image quality at aberration setting C_a , and c is a coefficient that determines the size of the confidence bound⁴¹. Increasing c favors exploration (searching new regions for a higher maximum) while decreasing c favors exploitation (searching closer to the current maximum for a better estimate)⁴³. We used $c = 2$ throughout this work. Other acquisition functions for Bayesian optimization, such as predicted improvement and expected improvement⁴⁴, can also be used, but we have found UCB to be most effective. Here, we use Bayesian optimization to determine the optimum aberration corrector alignment. The latent function for the GP framework is the image NV, which we maximize to correct aberrations.

GP calculations in this work were handled by `gpCAM` v 8.1.4⁴⁵. We used an anisotropic Matern kernel of first order differentiability ($\nu = 1.5$).

BEACON operation

As shown in Fig. 11a, BEACON uses a client-server model to separate the Bayesian optimization process from the microscope control computer. First, the client requests a new measurement with a specific set of aberration coefficients. The server changes the state of the aberration corrector and acquires a new HAADF-STEM image. The server then calculates the image NV and returns this value to the client. The acquired image can also be returned to the client. The client updates the surrogate model of image quality for all sets of aberration coefficients as well as the model uncertainty. The UCB acquisition function then uses the surrogate model and the model uncertainty to choose the next set of aberration coefficients to measure. Figure 11b–d shows how the surrogate model evolves as more data points are added. The region of parameter space in which the peak resides is located with just 5 data points, as shown in Fig. 11b. Subsequent sampling efforts are focused in this region, and the approximate location of the peak is located with 15 data points, as shown in Fig. 11c. BEACON continues to search around this peak to accurately

identify the peak, as shown in Fig. 11d. This shows how BEACON can efficiently optimize the aberrations without needing to sample large, extraneous portions of parameter space.

We acquired 512 by 512-pixel images with a dwell time of 3 μ s and a current of 70 pA. Large adjustments to aberrations, particularly B_2 and A_2 , result in beam shifts that change the field of view. We use cross-correlation to ensure the same 256 by 256 pixel image region is used throughout an optimization run which also compensates for sample drift. The field of view was 7.8 nm for the HfO_2 thin film and 62.4 nm for the Au NP sample. BEACON incorporates a graphical user interface (GUI) as shown in Supplementary Fig. 10 that enables users to quickly tune the optimization parameters for their sample. This GUI is included in the BEACON software, which is generally available (see Code Availability for details).

Bayesian optimization can incorporate noise variances into the calculation of surrogate models. Initial testing showed that proper noise estimation was vital to both the speed and precision of optimization. Measurements from the HfO_2 sample were noisier than those from the Au NP sample, requiring the use of different noise estimation functions. For the HfO_2 sample, the noise function S was:

$$S = 0.0001 \times (y_{\max} - y_{\min}). \quad (5)$$

For the Au NP sample, we found that the default noise estimation function used by gpCAM,

$$S = 0.01 \times |E9|, \quad (6)$$

was effective.

Ptychographic reconstructions

For the ptychography experiments, scans were collected at 300 kV using a probe defined with a 25 mrad convergence angle. Approximately 400 Å of defocus was added, and the step size of the acquisition was 1.4 Å. A dose of approximately $2 \times 10^4 \text{ e}^-/\text{Å}^2$ was used. We used the 4D Camera to acquire scans⁴⁶. We reconstructed both datasets using a gradient descent algorithm with batched updates implemented in the open-source toolkit py4DSTEM v 0.14.8^{33,47}. A batch size of 1024 measurements was used (1/16 the total number of probe positions), and 100 iterations were run to converge the mean squared error. The probe amplitude was fixed using a vacuum probe measurement taken during the same session. Position refinement and slight low pass filtering was applied during each update. The initial defocus was estimated with a parallax reconstruction. The A_1 or B_2 values were subsequently estimated using Bayesian optimization, minimizing the mean squared error with 25 initial, uniformly spaced points followed by 25 more optimized calls. More details about the reconstruction formalism can be found elsewhere³³.

Data availability

The data used to generate figures 1 and 2 is available from the Zenodo repository (<https://doi.org/10.5281/zenodo.13891904>). Other data is available upon request.

Code availability

The BEACON code is available from our Github repository (<https://github.com/AlexanderJPattison/BEACON>).

Received: 18 October 2024; Accepted: 4 August 2025;

Published online: 24 August 2025

References

- Kirkland, E. J. Fine tuning an aberration corrected ADF-STEM. *Ultramicroscopy* **186**, 62–65 (2018).
- Biskupek, J., Hartel, P., Haider, M. & Kaiser, U. Effects of residual aberrations explored on single-walled carbon nanotubes. *Ultramicroscopy* **116**, 1–7 (2012).
- Haider, M., Uhlemann, S. & Zach, J. Upper limits for the residual aberrations of a high-resolution aberration-corrected STEM. *Ultramicroscopy* **81**, 163–175 (2000).
- Martis, J. et al. Imaging the electron charge density in monolayer MoS_2 at the ångström scale. *Nat. Commun.* **14**, 4363 (2023).
- Li, C., Zhang, Q. & Mayoral, A. Ten years of aberration corrected electron microscopy for ordered nanoporous materials. *ChemCatChem* **12**, 1248–1269 (2020).
- Lupini, A. R. The Electron Ronchigram. In Pennycook, S. J. & Nellist, P. D. (eds.) *Scanning Transmission Electron Microscopy, Imaging and Analysis*, 117–161 (Springer, New York, 2011).
- Krivanek, O., Dellby, N. & Lupini, A. Towards sub-Å electron beams. *Ultramicroscopy* **78**, 1–11 (1999).
- Haider, M. et al. Electron microscopy image enhanced. *Nature* **392**, 768–769 (1998).
- Zemlin, F., Weiss, K., Schiske, P., Kunath, W. & Herrmann, K.-H. Coma-free alignment of high resolution electron microscopes with the aid of optical diffractograms. *Ultramicroscopy* **3**, 49–60 (1978).
- Bertoni, G. et al. Near-real-time diagnosis of electron optical phase aberrations in scanning transmission electron microscopy using an artificial neural network. *Ultramicroscopy* **245**, 113663 (2023).
- Barthel, J. & Thust, A. On the optical stability of high-resolution transmission electron microscopes. *Ultramicroscopy* **134**, 6–17 (2013).
- Schubert, P. J., Saxena, R. & Kornfeld, J. Deepfocus: Fast focus and astigmatism correction for electron microscopy. *Nat. Commun.* **15**, 1–10 (2024).
- Ma, D. et al. Emittance minimization for aberration correction i: Aberration correction of an electron microscope without knowing the aberration coefficients. *Ultramicroscopy* **273**, 114137 (2025).
- Ma, D. et al. Emittance minimization for aberration correction ii: Physics-informed bayesian optimization of an electron microscope. *Ultramicroscopy* **273**, 114138 (2025).
- Schnitzer, N., Sung, S. H. & Hovden, R. Optimal stem convergence angle selection using a convolutional neural network and the strehl ratio. *Microsc. Microanal.* **26**, 921–928 (2020).
- Ishikawa, R. et al. Automated geometric aberration correction for large-angle illumination STEM. *Ultramicroscopy* **222**, 113215 (2021).
- Noack, M. M. et al. Gaussian processes for autonomous data acquisition at large-scale synchrotron and neutron facilities. *Nat. Rev. Phys.* **3**, 685–697 (2021).
- Lorenz, R., Hampshire, A. & Leech, R. Neuroadaptive Bayesian optimization and hypothesis testing. *Trends Cogn. Sci.* **21**, 155–167 (2017).
- Thomas, J. C. et al. Autonomous scanning probe microscopy investigations over WS_2 and Au {111}. *npj Comput. Mater.* **8**, 99 (2022).
- Erasmus, S. & Smith, K. An automatic focusing and astigmatism correction system for the SEM and CTEM. *J. Microsc.* **127**, 185–199 (1982).
- Santos, A. et al. Evaluation of autofocus functions in molecular cytogenetic analysis. *J. Microsc.* **188**, 264–272 (1997).
- Rudnaya, M., Mattheij, R. & Maubach, J. Evaluating sharpness functions for automated scanning electron microscopy. *J. Microsc.* **240**, 38–49 (2010).
- Dembele, S., Lehmann, O., Medjaher, K., Marturi, N. & Piat, N. Combining gradient ascent search and support vector machines for effective autofocus of a field emission–scanning electron microscope. *J. Microsc.* **264**, 79–87 (2016).
- Batten, C. F. Autofocusing and astigmatism correction in the scanning electron microscope. *Mphil thesis, University of Cambridge* (2000).
- Saxton, W., Smith, D. J. & Erasmus, S. Procedures for focusing, stigmating and alignment in high resolution electron microscopy. *J. Microsc.* **130**, 187–201 (1983).
- Kirkland, E. J. An image and spectrum acquisition system for a VG HB501 STEM using a color graphics workstation. *Ultramicroscopy* **32**, 349–364 (1990).

27. Hoppe, W. Beugung im inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen. *Acta Crystallogr. Sect. A* **25**, 495–501 (1969).
28. Hoppe, W. & Strube, G. Beugung in inhomogenen Primärstrahlwellenfeld. II. Lichtoptische Analogieversuche zur Phasenmessung von Gitterinterferenzen. *Acta Crystallogr. Sect. A* **25**, 502–507 (1969).
29. Hoppe, W. Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei unperiodischen Objekten. *Acta Crystallogr. Sect. A* **25**, 508–514 (1969).
30. Rodenburg, J. & Bates, R. The theory of super-resolution electron microscopy via wigner-distribution deconvolution. *Philos. Trans. R. Soc. Lond. Ser. A: Phys. Eng. Sci.* **339**, 521–553 (1992).
31. Nellist, P. D., McCallum, B. C. & Rodenburg, J. M. Resolution beyond the ‘information limit’ in transmission electron microscopy. *Nature* **374**, 630–632 (1995).
32. Chen, Z. et al. Mixed-state electron ptychography enables sub-angstrom resolution imaging with picometer precision at low dose. *Nat. Commun.* **11**, 2994 (2020).
33. Vamavides, G. et al. Iterative phase retrieval algorithms for scanning transmission electron microscopy. *arXiv preprint arXiv:2309.05250* (2023).
34. Ophus, C. Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microsc. microanalysis* **25**, 563–582 (2019).
35. Nguyen, K. X. et al. Achieving sub-0.5-angstrom-resolution ptychography in an uncorrected electron microscope. *Science* **383**, 865–870 (2024).
36. Lupini, A., Chi, M. & Jesse, S. Rapid aberration measurement with pixelated detectors. *J. Microsc.* **263**, 43–50 (2016).
37. Yu, Y. et al. Dose-efficient cryo-electron microscopy for thick samples using tilt-corrected scanning transmission electron microscopy, demonstrated on cells and single particles. *bioRxiv* <https://www.biorxiv.org/content/early/2024/08/15/2024.04.22.590491> (2024).
38. Yang, H., Ercius, P., Nellist, P. D. & Ophus, C. Enhanced phase contrast transfer using ptychography combined with a pre-specimen phase plate in a scanning transmission electron microscope. *Ultramicroscopy* **171**, 117–125 (2016).
39. Pelz, P. M., Qiu, W. X., Bückner, R., Kassier, G. & Miller, R. D. Low-dose cryo electron ptychography via non-convex Bayesian optimization. *Sci. Rep.* **7**, 9883 (2017).
40. Ribet, S. M. et al. Design of electrostatic aberration correctors for scanning transmission electron microscopy. *Microsc. Microanal.* **29**, 1950–1960 (2023).
41. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **104**, 148–175 (2016).
42. Noack, M. M. & Sethian, J. A. Advanced stationary and nonstationary kernel designs for domain-aware gaussian processes. *Commun. Appl. Math. Comput. Sci.* **17**, 131–156 (2022).
43. Berk, J., Gupta, S., Rana, S. & Venkatesh, S. Randomised Gaussian process upper confidence bound for Bayesian optimisation. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI’20* (2021).
44. Gan, W., Ji, Z. & Liang, Y. Acquisition functions in bayesian optimization. In *2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)*, 129–135 (2021).
45. 8.1.4 — gpcam.lbl.gov. <https://gpcam.lbl.gov/>. [Accessed 06-08-2024].
46. Ercius, P. et al. The 4D camera: An 87 kHz direct electron detector for scanning/transmission electron microscopy. *Microsc. Microanal.* **30**, 903–912 (2024).
47. Savitzky, B. H. et al. py4DSTEM: A software package for four-dimensional scanning transmission electron microscopy data analysis. *Microsc. Microanal.* **27**, 712–743 (2021).

Acknowledgements

This work was primarily funded by the US Department of Energy in the program "Electron Distillery 2.0: Massive Electron Microscopy Data to Useful Information with AI/ML." Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a Department of Energy Office of Science User Facility using NERSC award BES-ERCAP0028035. S.M.R. and C.O. acknowledge support from the U.S. Department of Energy Early Career Research Award program. J.P. acknowledges financial support from the National Research Foundation of Korea (NRF) grant, funded by the Korea government (MSIT) (Grant No. RS-2023-00283902, and RS-202400408823). We gratefully acknowledge CEOS, GmbH for providing the server enabling communication with the corrector and I. Massmann for technical assistance.

Author contributions

A.J.P. developed the BEACON software and assessed its performance. S.M.R. verified BEACON performance using ptychographic reconstructions. M.M.N. is the developer of gpCAM and assisted with the development of BEACON. A.J.P., S.M.R., M.M.N., and P.E. wrote the draft manuscript. S.M.R., G.V. and C.O. are developers of py4DSTEM. K.P. and J.P. created the HfO₂ thin film sample. E.J.K. is the developer of the image NV theory and provided helpful discussions. P.E. supervised the project and assisted with conceptualization. All authors have read and approved the manuscript.

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/s41524-025-01766-4>.

Correspondence and requests for materials should be addressed to Peter Ercius.

Reprints and permissions information is available at <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) 2025