

Making the link between ADF and 4D STEM: Resolution, transfer and coherence

Peter D. Nellist and Timothy J. Pennycook

Citation: Chin. Phys. B, 2024, 33 (**11**): 116803. DOI: 10.1088/1674-1056/ad8554
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Multidimensional images and aberrations in STEM

Eric R. Hoglund and Andrew R. Lupini
Chin. Phys. B, 2024, 33 (**9**): 096807. DOI: 10.1088/1674-1056/ad73b2

Multiphase cooperation for multilevel strain accommodation in a single-crystalline $BiFeO_3$ thin film

Wooseon Choi, Bumsu Park, Jaejin Hwang, Gyeongtak Han, Sang-Hyeok Yang, Hyeon Jun Lee, Sung Su Lee, Ji Young Jo, Albina Y. Borisevich, Hu Young Jeong, Sang Ho Oh, Jaekwang Lee, and Young-Min Kim
Chin. Phys. B, 2024, 33 (**9**): 096805. DOI: 10.1088/1674-1056/ad62e0

Symmetry quantification and segmentation in STEM imaging through Zernike moments

Jiadong Dan, Cheng Zhang, Xiaoxu Zhao(赵晓续), and N. Duane Loh
Chin. Phys. B, 2024, 33 (**8**): 086803. DOI: 10.1088/1674-1056/ad51f4

Cryogenic transmission electron microscopy on beam-sensitive materials and quantum science

Gang Wang(王刚) and Jun-Hao Lin(林君浩)
Chin. Phys. B, 2024, 33 (**8**): 086801. DOI: 10.1088/1674-1056/ad5af0

Cryo-EM combined with image deconvolution to determine ZIF-8 crystal structure

Kang Wu(吴抗), Baisong Yang(杨柏松), Wenhua Xue(薛文华), Dapeng Sun(孙大鹏), Binghui Ge(葛炳辉), and Yumei Wang(王玉梅)
Chin. Phys. B, 2024, 33 (**7**): 076802. DOI: 10.1088/1674-1056/ad48f8

SPECIAL TOPIC — Stephen J. Pennycook: A research life in atomic-resolution STEM and EELS

Making the link between ADF and 4D STEM: Resolution, transfer and coherence

Peter D. Nellist[1,†] and Timothy J. Pennycook[2]

[1] *Department of Materials, University of Oxford, Oxford, UK*
[2] *EMAT, University of Antwerp, Groenenborgerlaan* 171, 2020 *Antwerp, Belgium*

(Received 3 September 2024; revised manuscript received 5 October 2024; accepted manuscript online 10 October 2024)

Steve Pennycook is a pioneer in the application of high-resolution scanning transmission electron microscopy (STEM) and in particular the use of annular dark-field (ADF) imaging. Here we show how a general framework for 4D STEM allows clear links to be made between ADF imaging and the emerging methods for reconstructing images from 4D STEM data sets. We show that both ADF imaging and ptychographical reconstruction can be thought of in terms of integrating over the overlap regions of diffracted discs in the detector plane. This approach allows the similarities in parts of their transfer functions to be understood, though we note that the transfer functions for ptychographic imaging cannot be used as a measure of information transfer. We also show that conditions of partial spatial and temporal coherence affect ADF imaging and ptychography similarly, showing that achromatic interference can always contribute to the image in both cases, leading to a robustness to partial temporal coherence that has enabled high-resolution imaging.

Keywords: image forming and processing, phase retrieval, electron microscopy, scanning transmission electron microscope (STEM)

PACS: 68.37.Ma, 42.30.Va, 42.30.Rx, 87.64.Ee **DOI:** 10.1088/1674-1056/ad8554

1. Introduction

Prior to the early 1990s, atomic resolution electron microscope imaging was the preserve of the transmission electron microscope (TEM) configuration and was referred to as high-resolution TEM (HRTEM). That the situation is largely reversed in the current time, with the majority of atomic-resolution work now performed using the scanning transmission electron microscope (STEM), is largely a result of the pioneering work of Steve Pennycook and the development of the annular dark-field (ADF) STEM approach. The possibility of excluding diffraction effects using a sufficiently large ADF detector had previously been discussed by Howie,[1] but pioneers such as Pennycook pushed the approach to atomic resolution. In particular, Pennycook procured the first STEM instrument specifically designed for atomic resolution work — the VG Microscopes HB603U, a 300 kV instrument fitted with a high-resolution pole-piece. The instrument quickly produced a range of atomic-resolution results (see for example Refs. [2,3]). After that pioneering work, the technique of ADF STEM became the standard approach for atomic resolution work, alongside the spectroscopic techniques of electron energy-loss spectroscope (EELS) and energy-dispersive x-ray (EDX) spectroscopy.

Although, by the principle of reciprocity, phase contrast imaging can be performed in STEM using a small axial bright-field (BF) detector, it makes inefficient use of the transmitted electrons and is not competitive with HRTEM. A substantial development in STEM imaging has emerged over the past eight years or so — that of 4D STEM using fast pixelated detectors. This approach allows greatly improved access to phase information in STEM, closing the circle back to the phase contrast imaging that is behind the HRTEM approach. For high-resolution imaging, a number of 4D STEM imaging techniques have emerged, including ptychography, center of mass (CoM) (first moment),[4] and integrated center of mass (iCoM).[5] Such methods, being based on coherent interference effects in the scattered electrons, may seem quite distinct from ADF imaging which is regarded as being incoherent. Here we present a framework that has been developed in the context of ptychography to explain the main attributes of ptychography and ADF imaging. The methods are compared and contrasted, and in particular we explore the concepts of transfer functions and the effects of partial coherence on both methods.

2. The theoretical basis of STEM imaging

We start by considering a situation of perfect spatial coherence, under which conditions the STEM illuminating probe wavefunction $P(\boldsymbol{R})$ can be simply calculated as the inverse Fourier transform of the aperture function $A(\boldsymbol{K})$,

$$P(\boldsymbol{R}) = \int A(\boldsymbol{K}) \exp\left(\mathrm{i}2\pi \boldsymbol{K}\cdot\boldsymbol{R}\right) \mathrm{d}\boldsymbol{K}, \tag{1}$$

where $\boldsymbol{R}$ is a real-space vector in the sample plane and $\boldsymbol{K}$ is a reciprocal-space vector in the front focal plane. The

†Corresponding author. E-mail: peter.nellist@materials.ox.ac.uk

© 2024 Chinese Physical Society and IOP Publishing Ltd

value of $\boldsymbol{K}$ can be related to a convergence angle of a partial plane wave in the convergent, conical beam forming the probe through $\theta = \lambda |\boldsymbol{K}|$. The modulus of $A(\boldsymbol{K})$ is unity when $\boldsymbol{K}$ is inside the beam limiting aperture, often known as the condenser aperture but optically playing the role of an objective aperture, and zero outside the aperture. The phase of $A(\boldsymbol{K})$ is controlled by any lens aberrations present, being zero in a perfectly aberration-corrected system, and otherwise a function of $\boldsymbol{K}$ as given in many publications (see for example Ref. [6]).

We now assume that the interaction with the sample can be described through a multiplicative transmission function $\phi(\boldsymbol{R})$, which modifies the amplitude and phase of the transmitted electron wave function, with the modification in phase being the most important. The electron wave exiting the sample is then given by

$$\psi(\boldsymbol{R}, \boldsymbol{R}_\mathrm{p}) = P(\boldsymbol{R} - \boldsymbol{R}_\mathrm{p})\phi(\boldsymbol{R}), \tag{2}$$

where a new variable $\boldsymbol{R}_\mathrm{p}$ has been introduced to account for the ability to scan the probe. The wavefunction in the detector plane of the STEM, which can be regarded as being in the far-field, can then be written as a function of the probe position, $\boldsymbol{R}_\mathrm{p}$, and the position in the detector plane, given by the reciprocal space vector $\boldsymbol{K}_\mathrm{f}$, as

$$M(\boldsymbol{R}_\mathrm{p}, \boldsymbol{K}_\mathrm{f}) = \int \psi(\boldsymbol{R}, \boldsymbol{R}_\mathrm{p}) \exp(-\mathrm{i}2\pi \boldsymbol{K}_\mathrm{f} \cdot \boldsymbol{R})\, \mathrm{d}\boldsymbol{R}. \tag{3}$$

Given that, from Eq. (2), ψ is a product of P and ϕ, M now becomes a convolution in the detector plane between the Fourier transform of P, which we denote as A, and the Fourier transform of ϕ, which we denote by simply changing its argument, so that M can be written as

$$M(\boldsymbol{R}_\mathrm{p}, \boldsymbol{K}_\mathrm{f}) = \left[A(\boldsymbol{K}_\mathrm{f}) \exp(-\mathrm{i}2\pi \boldsymbol{K}_\mathrm{f} \cdot \boldsymbol{R}_\mathrm{p})\right] \otimes_{\boldsymbol{K}_\mathrm{f}} \phi(\boldsymbol{K}_\mathrm{f}) \tag{4}$$

and is illustrated in Fig. 1. The probe position is now encoded as a phase ramp across the aperture function through the Fourier shift theorem and each scattering vector from the sample, $\phi(\boldsymbol{K}_\mathrm{f})$, is broadening through the convolution with the aperture function. Different scattering vectors from the sample can now overlap and interfere at the same $\boldsymbol{K}_\mathrm{f}$ leading to interference. This is most easily seen for a crystalline sample with discrete Bragg beams, as shown in Fig. 1 where the beams broaden to form overlapping discs, but we emphasize that the analysis presented here is not limited to periodic objects. The interference between different scattered beams will be dependent on the relative phases of the beams, which are dependent on (i) the probe position as shown by Eq. (4), (ii) any lens aberrations present through the phase of A and (iii) the relative phase of the different scattered beams. Point (i) is at the heart of STEM imaging: As the probe is scanned, the interference will vary between constructive and destructive and an intensity detector in the overlap region will give contrast in a STEM image. The information about the sample structure is carried by the dependence explained in point (iii). Positional information is most strongly carried by the phase of the scattered beams so determining their phase allows images to be formed.

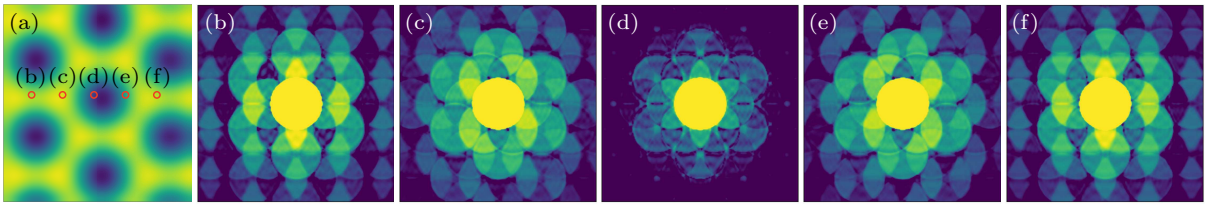


Fig. 1. Panel (a) shows an SSB image showing the positions for which $|M(\boldsymbol{R}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})|$ vs. $\boldsymbol{K}_\mathrm{f}$ is displayed in panels (b)–(f). Data simulated for graphene at 60 kV with a 20 mrad convergence angle. Notice how the intensity in the disk overlaps in $|M(\boldsymbol{R}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})|$ oscillates due to the phase ramp encoded in the aperture function in Eq. (4). $|M(\boldsymbol{R}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})|$ is displayed on a log scale with contrast boosted to show the differences between probe positions.

The measured 4D STEM data set is the modulus-squared of Eq. (3). Rodenburg and Bates[7] showed that the mathematics simplified beautifully if the expression was fully in reciprocal space by taking the Fourier transform of the 4D STEM data set with respect to the probe position,

$$G(\boldsymbol{Q}_\mathrm{p}, \boldsymbol{K}_\mathrm{f}) = \int \left|M(\boldsymbol{R}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})\right|^2 \exp(-\mathrm{i}2\pi \boldsymbol{R}_\mathrm{p} \cdot \boldsymbol{Q}_\mathrm{p})\, \mathrm{d}\boldsymbol{R}_\mathrm{p}, \tag{5}$$

which results in the following convolution:

$$\begin{aligned} &G(\boldsymbol{Q}_\mathrm{p}, \boldsymbol{K}_\mathrm{f}) \\ &= \left[A(\boldsymbol{K}_\mathrm{f}) A^*(\boldsymbol{K}_\mathrm{f} + \boldsymbol{Q}_\mathrm{p})\right] \otimes_{\boldsymbol{K}_\mathrm{f}} \left[\phi(\boldsymbol{K}_\mathrm{f}) \phi(\boldsymbol{K}_\mathrm{f} - \boldsymbol{Q}_\mathrm{p})\right]. \end{aligned} \tag{6}$$

This equation can be seen as being a master equation for STEM imaging. The left operand in the equation is the overlap function between two off-set aperture functions. It shows that information arising in a STEM image at a spatial frequency of $\boldsymbol{Q}_\mathrm{p}$ is arising from the interference between pairs of beams in the incident cone of illumination separated by $\boldsymbol{Q}_\mathrm{p}$, as shown in Fig. 2. This overlap feature is then convolved over all pairs of scattered beams that are separated by $\boldsymbol{Q}_\mathrm{p}$.

Examples of the amplitude of $G(\boldsymbol{Q}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})$ showing regions in the detector plane where the interference is strong are displayed in Fig. 3 for two spatial frequencies with 20 mrad and 35 mrad convergence angles, simulated for graphene at

60 kV, noise free using the *ab*TEM and pyPtychoSTEM packages. The disk overlaps that occur in the central BF disk region can be seen to be repeated throughout the entire range of $\boldsymbol{K}_\mathrm{f}$ for each spatial frequency due to the convolution in Eq. (6) and the nature of the diffraction pattern for a crystal with many beams being separated by $\boldsymbol{Q}_\mathrm{p}$ even at higher $\boldsymbol{K}_\mathrm{f}$ values.

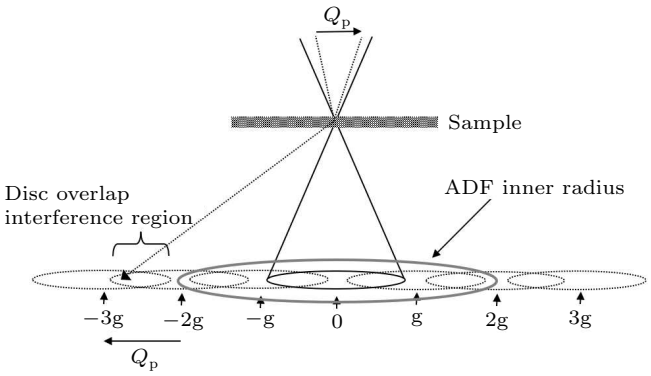


Fig. 2. A schematic showing how STEM images are formed from the interference between beams in the incident cone of illumination that are scattered into a final wavevector and can interfere. For a particular spatial frequency in an image, $\boldsymbol{Q}_\mathrm{p}$, all pairs of beams in the incident cone separated by $\boldsymbol{Q}_\mathrm{p}$ contribute.

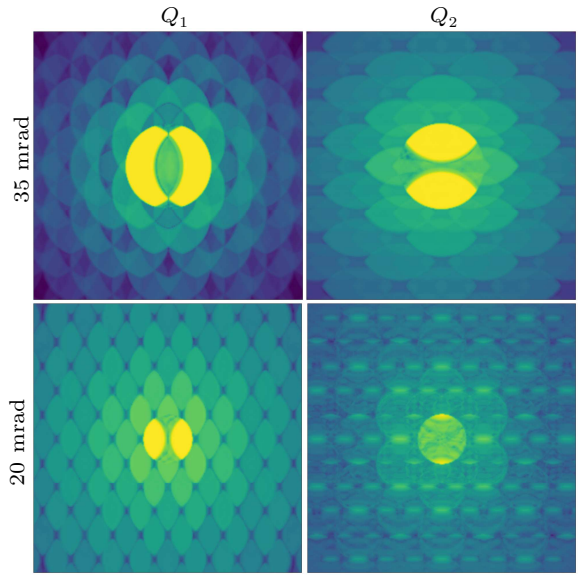


Fig. 3. A plot of $|G(\boldsymbol{Q}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})|$ as a function of $\boldsymbol{K}_\mathrm{f}$ for graphene displayed for two spatial frequencies with $Q_2 > Q_1$ with 20 mrad and 35 mrad convergence angles at 60 kV. Q_2 is on the very edge of the high end of the frequencies transferred by the 20 mrad convergence angle, while it is midrange for the 35 mrad convergence angle. Notice how the pattern inside the central disk of $G(\boldsymbol{Q}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})$ for a given $\boldsymbol{Q}_\mathrm{p}$ repeats throughout $\boldsymbol{K}_\mathrm{f}$ due to the convolution in Eq. (6).

The interferometry that occurs between beams scattered at different angles due to the range of angles in the coherent illumination is the source of phase information in ptychography.[8] In fact, all STEM imaging can be seen in terms of this interference, and indeed the question has been raised of whether the STEM should actually be referred to as a ptychograph (or should that be ptychoscope?)![9] Different STEM imaging methods can be regarded as being methods for extracting the sample information, $\phi(\boldsymbol{K}_\mathrm{f})$, from Eq. (6). Here we will use Eq. (6) to compare ADF imaging and ptychography, and also highlight some of the similarities of the impact of partial spatial coherence.

3. Incoherent imaging using an annular dark-field detector

For a single channel detector, such as an annular dark-field (ADF) detector, bright-field (BF) or indeed annular bright-field (ABF) detector, the image can be calculated quite simply from Eq. (6) by integrating over the relevant range of $\boldsymbol{K}_\mathrm{f}$. This is valid because $\boldsymbol{K}_\mathrm{f}$ and $\boldsymbol{Q}_\mathrm{p}$ are independent of each other. Of course, this approach yields the image as a function of $\boldsymbol{Q}_\mathrm{p}$ and can then be inverse Fourier transformed back to the real-space image. Defining the area of the ADF detector through a function $D_\mathrm{ADF}(\boldsymbol{K}_\mathrm{f})$, which has a value of unity in the detected region and zero outside, the Fourier transform of the ADF image can then be calculated by

$$I_\mathrm{ADF}(\boldsymbol{Q}_\mathrm{p}) = \int D_\mathrm{ADF}(\boldsymbol{K}_\mathrm{f}) G(\boldsymbol{Q}_\mathrm{p}, \boldsymbol{K}_\mathrm{f})\, \mathrm{d}\boldsymbol{K}_\mathrm{f}, \tag{7}$$

in which the convolution can be written explicitly using a dummy variable for the convolution, $\boldsymbol{K}'_\mathrm{f}$

$$\begin{aligned} I_\mathrm{ADF}(\boldsymbol{Q}_\mathrm{p}) = \iint & D_\mathrm{ADF}(\boldsymbol{K}_\mathrm{f}) A(\boldsymbol{K}'_\mathrm{f}) A^*(\boldsymbol{K}'_\mathrm{f} + \boldsymbol{Q}_\mathrm{p}) \\ & \times \phi(\boldsymbol{K}_\mathrm{f} - \boldsymbol{K}'_\mathrm{f}) \phi^*(\boldsymbol{K}_\mathrm{f} - \boldsymbol{K}'_\mathrm{f} - \boldsymbol{Q}_\mathrm{p})\, \mathrm{d}\boldsymbol{K}_\mathrm{f}\, \mathrm{d}\boldsymbol{K}'_\mathrm{f}. \end{aligned} \tag{8}$$

Note that the domain of the $\boldsymbol{K}'_\mathrm{f}$ is controlled by the overlap region shown in Figs. 1 and 2, whereas the domain of $\boldsymbol{K}_\mathrm{f}$ is controlled by the detector size. As discussed in more details in Ref. [10] building on work in Ref. [11], if we assume that the detector geometry is much larger than the aperture overlap region, it is reasonable to approximate Eq. (8) in the following form:

$$\begin{aligned} I_\mathrm{ADF}(\boldsymbol{Q}_\mathrm{p}) = & \int A(\boldsymbol{K}'_\mathrm{f}) A^*(\boldsymbol{K}'_\mathrm{f} + \boldsymbol{Q}_\mathrm{p})\, \mathrm{d}\boldsymbol{K}'_\mathrm{f} \\ & \times \int D_\mathrm{ADF}(\boldsymbol{K}_\mathrm{f}) \phi(\boldsymbol{K}_\mathrm{f}) \phi^*(\boldsymbol{K}_\mathrm{f} - \boldsymbol{Q}_\mathrm{p})\, \mathrm{d}\boldsymbol{K}_\mathrm{f}. \end{aligned} \tag{9}$$

Being a product in reciprocal space, taking the inverse Fourier transform of Eq. (9) to form a real space image results in a convolution. It can also be noted that the integral involving the aperture function is the autocorrelation function of A which, when inverse Fourier transformed, gives the intensity of the illuminating probe convolved with an object function which is a high-pass filtered version of the transmission function, thus

$$I_\mathrm{ADF}(\boldsymbol{R}_\mathrm{p}) = \left|P(\boldsymbol{R}_\mathrm{p})\right| \otimes O(\boldsymbol{R}_\mathrm{p}), \tag{10}$$

where $O(\boldsymbol{R}_\mathrm{p})$ is the modulus squared of a high-pass filtered version of the transmission function.[10] The high-pass filtering of the transmission function means that the higher angle parts of the scattering factors are retained, which arise predominantly from scattering by the nucleus and hence gives Z-contrast, where Z denotes the atomic number.

Although the image contrast predicted by Eq. (10) has arisen from coherent interference occurring between beams in the illumination, it is still regarded as demonstrating that ADF imaging is incoherent imaging. The reason for this is the analogy with conventional imaging of a self-luminous object which would involve the modulus-squared of the point-spread function. By reciprocity, the ADF detector can be thought of as being equivalent to incoherent illumination in a TEM over a wide range of illumination angles. It was initially suggested that the incoherent nature of ADF imaging arose from the high-angle detector collecting predominantly thermal diffuse scattering (TDS). Note that the formulation presented here does not require the scattering itself to be incoherent such as would occur with TDS.

A significant development by Steve Pennycook was realizing that the incoherent nature of the imaging simplified the imaging even under dynamical scattering conditions.[12] Developing this theme further, Nellist and Pennycook[13] showed that Eq. (8) could be written for a dynamical object using the Bloch wave formulation

$$\begin{aligned} G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f}) \;=\; & \sum_{\boldsymbol{g},j,k}\int \delta(\boldsymbol{K}_\mathrm{f}-\boldsymbol{K}-\boldsymbol{g})A(\boldsymbol{K})A^*(\boldsymbol{K}+\boldsymbol{Q}_\mathrm{p}) \\ & \times \Phi_0^{(j)*}(\boldsymbol{K})\Phi_{\boldsymbol{g}}^{(j)}(\boldsymbol{K})\Phi_{\boldsymbol{Q}_\mathrm{p}}^{(k)}(\boldsymbol{K})\Phi_{\boldsymbol{g}}^{(k)*}(\boldsymbol{K}) \\ & \times \exp\left\{-\mathrm{i}2\pi z\left[k_z^{(j)}(\boldsymbol{K})-k_z^{(k)}(\boldsymbol{K})\right]\right\}\mathrm{d}\boldsymbol{K}, \end{aligned} \tag{11}$$

where the summation is over reciprocal lattice vectors $\boldsymbol{g}$ and Bloch wave indices j and k. The Bloch waves are represented through the Fourier coefficients of the eigenfunctions, Φ, and the associated eigenvalues k_z, and z represents the crystal thickness. The off-set aperture overlap function is again visible here, and Nellist and Pennycook showed how the integral over both $\boldsymbol{K}$ in this expression and the integral over $\boldsymbol{K}_\mathrm{f}$ that occurs with an integral over the detector reduces the degree of interference between Bloch waves, reducing the impact of dynamical effects.

4. Coherent imaging using ptychography

We have seen that one way to separate the sample information from Eq. (6) is to integrate over a geometrically large detector, leading to incoherent Z-contrast imaging. The availability of fast, pixelated, direct electron detectors has made recording the full 4D STEM data set described in Eq. (4) routine, allowing the $G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f})$ function to be measured. The obvious approach now to separating the sample information is to deconvolve the offset aperture overlap function. There are a number of ways to achieve this. A direct deconvolution can be performed, such as the Wigner distribution deconvolution (WDD) approach, which has been used in a number of applications (see for example Ref. [14]). It is also possible to solve the deconvolution blind, which means simultaneously solving for the aperture function and the sample transmission function, and this is what the single slice iterative algorithms such as ePIE achieve.[15] For thicker samples, an iterative multislice approach can cope with multiple scattering and obtain three-dimensional information.[16]

For a phase object, an alternative route to access the phase information is to use the integrated center of mass approach.[5,17] In this approach, the first moment of the intensity is used to get a measure of the grad vector of the potential,

$$I_\mathrm{CoM}(\boldsymbol{Q}_\mathrm{p}) = \int \boldsymbol{K}_\mathrm{f} G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f})\,\mathrm{d}\boldsymbol{K}_\mathrm{f}. \tag{12}$$

Much like we saw for the ADF case in Section 2, the integral over all of $\boldsymbol{K}_\mathrm{f}$ allows for the separation of the offset aperture overlap function and the sample transmission function. Writing the convolution within G explicitly gives

$$\begin{aligned} I_\mathrm{CoM}(\boldsymbol{Q}_\mathrm{p}) \;=\; & \iint \boldsymbol{K}_\mathrm{f}A(\boldsymbol{K}_\mathrm{f}')A^*(\boldsymbol{K}_\mathrm{f}'+\boldsymbol{Q}_\mathrm{p})\phi(\boldsymbol{K}_\mathrm{f}-\boldsymbol{K}_\mathrm{f}') \\ & \times\phi^*(\boldsymbol{K}_\mathrm{f}-\boldsymbol{K}_\mathrm{f}'-\boldsymbol{Q}_\mathrm{p})\,\mathrm{d}\boldsymbol{K}_\mathrm{f}\,\mathrm{d}\boldsymbol{K}_\mathrm{f}'. \end{aligned} \tag{13}$$

Because both the $\boldsymbol{K}_\mathrm{f}$ and $\boldsymbol{K}_\mathrm{f}'$ integrals are over all space, the integrals can be separated to give

$$\begin{aligned} I_\mathrm{CoM}(\boldsymbol{Q}_\mathrm{p}) \;=\; & \int A(\boldsymbol{K}_\mathrm{f}')A^*(\boldsymbol{K}_\mathrm{f}'+\boldsymbol{Q}_\mathrm{p})\,\mathrm{d}\boldsymbol{K}_\mathrm{f}' \\ & \times \int \boldsymbol{K}_\mathrm{f}\phi(\boldsymbol{K}_\mathrm{f})\phi^*(\boldsymbol{K}_\mathrm{f}-\boldsymbol{Q}_\mathrm{p})\,\mathrm{d}\boldsymbol{K}_\mathrm{f}. \end{aligned} \tag{14}$$

Just like the ADF case, the inverse Fourier transform leads to a convolution between the probe intensity and, in this case, the grad vector of the sample phase which can be related to the projected electric field. The phase of the transmission function can then be retrieved by an integral of the field vector. Since the iCoM approach is a route to access the phase information that arises from the interferometry between different scattering vectors, it can be regarded as another form of ptychography.

5. Transfer functions for ADF imaging and ptychography

A transfer function describes how strongly a particular spatial frequency in the sample is transferred to the image. It is inevitable that in any imaging system, there will be an effective Fourier filter applied, often decreasing in transfer at high spatial frequencies. In the case of transfer of information for a weak-phase object in a conventional high-resolution TEM, the so-called phase contrast transfer function (PCTF) can be quite complex with negative regions indicating reversals in contrast at certain ranges of spatial frequency.

In the case of ADF imaging, Eq. (10) shows that the point-spread function is the probe intensity and therefore the transfer function can be regarded as being the Fourier transform of the probe intensity, which is the autocorrelation that forms the $\mathrm{d}\boldsymbol{K}_\mathrm{f}'$ integral in Eq. (6). This is known in light optics as the

optical transfer function (OTF) (see for example Ref. [18]). Because, in the absence of lens aberrations, the OTF is simply the integral over the overlap region of two discs separated by $\boldsymbol{Q}_{\mathrm{p}}$, this can be written analytically as

$$\mathrm{OTF}(\omega)=\frac{2}{\pi}\left[\cos^{-1}\left(\frac{\omega}{2}\right)-\frac{\omega}{2}\sqrt{1-\left(\frac{\omega}{2}\right)^2}\right]. \tag{15}$$

Here it is useful to work with a dimensionless quantity for spatial frequency, defined as $\omega=|\boldsymbol{Q}_{\mathrm{p}}|/K_{\max}$, where $K_{\max}$ is the maximum transverse momentum vector in the convergent illuminating beam, or equivalently the value of $\boldsymbol{K}$ at which the aperture function, A, falls to zero. It can be seen that the OTF falls to zero at $\omega=2$ and that therefore the resolution limit is given by the aperture diameter. It should be noted that in defining this OTF, we have neglected the filtering of the transmission function that occurs through the detector function in Eq. (8). The ADF image can be regarded as a map of the ability of the object to scatter to high angles. With the transfer defined through a simple OTF in this way, it is tempting to attempt to deconvolve the transfer from an ADF image. As shown in Ref. [19], this tends to lead to increased artefacts in the image. The gradual decay of the transfer smoothly to zero is close to optimal for minimizing artefacts. A deconvolution essentially involves dividing by the OTF to give a flat transfer, but the transfer must fall to zero at the information limit with the OTF falling to zero at $\omega=2$, resulting in a sharp cutoff in transfer and stronger artefacts. This can be overcome through the use of prior information in a Bayesian approach with early work being demonstrated in Pennycook's group.[20]

It is interesting to note that Eq. (14) suggests that the iCoM method has a similar transfer function, and transfer functions that are identical to the ADF OTF can be seen in Ref. [5]. The OTF has a maximum at zero spatial frequency, but there is a contradiction here because any in-line phase determination method, such as those applied for 4D STEM phase retrieval methods, cannot measure zero spatial frequency phase, which is the average phase over the field of view. Consider the addition of a constant phase to the electron wave exiting the sample. Such an addition will give rise to a uniform phase change of the wave in the far field, but will not lead to changes in amplitude. Since it is the intensity of the wave that is measured, the uniform phase addition cannot change the measurements in the far-field plane, and so will not be measurable. Only relative phase changes between the nonzero spatial frequencies are measurable, as only these will modify the far-field plane. The transfer for phase retrieval methods such as iCoM and ptychography must therefore fall to zero at zero spatial frequency in a way that is not captured by Eq. (14). Indeed, Ref. [5] notes the similarity between the iCoM CTF and the ADF OTF, but is careful to note that the iCoM method cannot retrieve the zero spatial frequency and the CTFs in that paper are not plotted to zero spatial frequency.

Indeed, if we look at the specific case of the weak-phase object approximation (WPOA)[21] one can see that the sample information is only transferred in a region corresponding to overlap between two diffracted discs, the triple overlap region not showing interference in the absence of aberrations. It is possible, therefore, to define a transfer function for 4D STEM phase retrieval within the WPOA. As can be seen in Fig. 4, for values of ω above one, this is the overlap that is identical to the ADF OTF. For $\omega<1$, however, the 4D STEM PCTF must start to fall again as the area of the double overlap region decreases. For $\omega<1$, it has been shown[22] that the PCTF has the form

$$\begin{aligned}\mathrm{PCTF}(\omega)=\frac{2}{\pi}\Bigg[&\cos^{-1}\left(\frac{\omega}{2}\right)-\cos^{-1}\omega+\omega\sqrt{1-\omega^2}\\&-\frac{\omega}{2}\sqrt{1-\left(\frac{\omega}{2}\right)^2}\Bigg],\end{aligned} \tag{16}$$

whereas for $\omega>1$, it follows the same form as the ADF OTF in Eq. (15). This transfer function is the one that is implicit to the single-side band (SSB) method that simply uses the information in the double overlap region to determine the amplitude and phase of a specific spatial frequency. Although both SSB and iCoM can form phase shift images in the WPOA from the same 4D STEM data, we can now see that different approaches to extracting the phase information have given rise to different transfer functions despite the information being expressed from the same interfering discs in the 4D STEM data set. We can conclude that there is no single fundamental transfer function for phase images reconstructed from 4D STEM, unlike ADF STEM. A direct imaging method, such as ADF STEM, can have a well-defined transfer function that gives a measure of information transfer for different spatial frequencies, whereas for 4D STEM phase retrieval methods such as ptychography and iCoM, the image is extracted from the 4D STEM data using an algorithm and in such cases a CTF is no longer a useful measure of information content.

The reason for this is because 4D STEM phase image reconstruction methods such as ptychography and iCoM also do not transfer the noise uniformly, whereas the noise in an ADF imaging is Poissonian, and therefore uniformly distributed in spatial frequency. The iCoM method makes use of a mathematical step that involves dividing by the spatial frequency, and therefore will amplify the noise alongside the signal at low spatial frequencies resulting in no gain in information content, as noted in the signal to noise analysis presented in Ref. [5] which does fall to low values at low spatial frequencies. To evaluate the performance of 4D STEM phase retrieval methods, a more sophisticated measure is needed otherwise the high apparent transfer of iCoM at low frequencies might lead

one to imagine that there is strong information transfer there. Suggested ways to capture the real information content may be based on the noise-normalized transfer function suggested by Ref. [23] or the use of Fisher information indices.[24] We conclude this section by noting that ptychography can exceed the diffraction limit for objects with strong enough diffraction outside the bright-field disc by using the phase difference measured between adjacent scattered beams to then solve for the phase of all measurable coherent diffraction.[25,26]

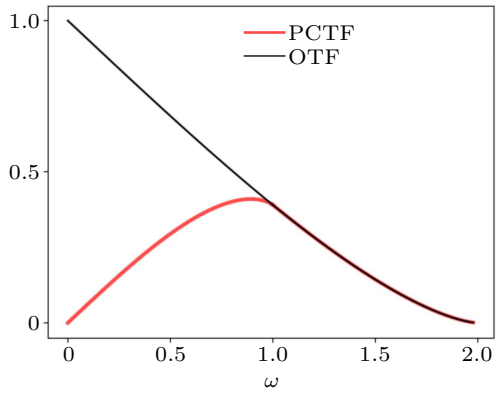


Fig. 4. A comparison of the OTF for incoherent imaging with the intrinsic PCTF for 4D STEM in the WPOA at zero defocus. Note that they are identical above $\omega = 1$ because in both cases the overlap between two diffracted discs is interfering.

6. Partial coherence

6.1. Spatial coherence

Finally we consider the impact of partial coherence on both ADF and ptychography. In STEM, partial spatial coherence refers to the finite extent of the electron source. Ignoring diffraction and aberration effects, the illuminating STEM probe can be regarded as being a demagnified image of the electron source, which will have an effective incoherent width. For both ADF and 4D STEM phase retrieval, the approach of incorporating this is similar, the intensity of any measurement is convolved over a range of probe positions controlled by the effective size of the demagnified electron source.

For ADF, the signal is a convolution in real-space, therefore there is an attenuating envelope in reciprocal space, thus

$$I_{\mathrm{src}}(\boldsymbol{Q}_{\mathrm{p}}) = I(\boldsymbol{Q}_{\mathrm{p}})S(\boldsymbol{Q}_{\mathrm{p}}), \tag{17}$$

where $I_{\mathrm{src}}(\boldsymbol{Q}_{\mathrm{p}})$ and $I(\boldsymbol{Q}_{\mathrm{p}})$ are the intensity of the ADF image with and without the effect of source size partial coherence included. The envelope, $S(\boldsymbol{Q}_{\mathrm{p}})$, is the Fourier transform of the effective source size at the sample after demagnification. The greater the degree of demagnification, usually achieved by exciting the first condenser lens harder, the less the envelope attenuates higher spatial frequencies, but at the expense of lower beam current.

A similar argument applies to 4D STEM where the partially coherent version of G relates to the perfect coherence version via

$$G_{\mathrm{src}}(\boldsymbol{Q}_{\mathrm{p}}, \boldsymbol{K}_{\mathrm{f}}) = G(\boldsymbol{Q}_{\mathrm{p}}, \boldsymbol{K}_{\mathrm{f}})S(\boldsymbol{Q}_{\mathrm{p}}). \tag{18}$$

The impact of this is that G is attenuated in the $\boldsymbol{Q}_{\mathrm{p}}$ direction but not in the $\boldsymbol{K}_{\mathrm{f}}$ direction. For methods that only use the BF disc, such as the SSB method or methods that use segmented detectors such as a quadrant on the BF disc, the reconstructed complex image is simply convolved with the effective source size. Note that this convolution is on the complex image and not just on the phase component.

Methods that can achieve resolution beyond the diffraction limit by using larger values of $\boldsymbol{K}_{\mathrm{f}}$ are not limited by spatial coherence. Rather counterintuitively, resolutions higher than the spatial coherence or indeed the stability of the microscope can be achieved.[25,26]

6.2. Temporal coherence

Partial temporal coherence refers to the defocus spread induced by chromatic aberrations in combination with the spread of the temporally varying energy of the electrons from the source, instabilities in the accelerating voltage and fluctuations in the lens current of the objective lens. In HRTEM, inelastic scattering of the specimen can also contribute significantly to the energy spread of the beam entering the objective lens. In HRTEM, this chromatic defocus spread imposes a damping envelope on the transfer function

$$E_{\mathrm{chr}}(\boldsymbol{Q}_{\mathrm{p}}) = \exp\left(-\frac{1}{2}\pi^2\lambda^2\Delta^2|\boldsymbol{Q}_{\mathrm{p}}|^4\right), \tag{19}$$

where Δ is the defocus spread, assuming a Gaussian defocus spread. In STEM, the probe forming optics are before the specimen, so the energy spread created by inelastic scattering with the sample is not a significant concern in contrast to the situation in HRTEM. This means that the energy spread reducing temporal coherence can be much lower in STEM, particularly for thick specimens. However there is always a defocus spread imposed by the microscope itself, typically principally through the energy spread of the electron source, which can be reduced by the use of a monochromator.

In STEM, partial temporal coherence must be taken into account by integrating over the defocus spread that enters the STEM master equation, Eq. (6), via the multiplicative pairs of aperture functions in the left bracket. For a given $\boldsymbol{Q}_{\mathrm{p}}$, in the absence of aberrations other than a defocus δ,

$$\begin{aligned} A(\boldsymbol{K}_{\mathrm{f}})A^*(\boldsymbol{K}_{\mathrm{f}}+\boldsymbol{Q}_{\mathrm{p}}) = 2a(\boldsymbol{K}_{\mathrm{f}})a(\boldsymbol{K}_{\mathrm{f}}+\boldsymbol{Q}_{\mathrm{p}}) \\ \times \exp\left[-\mathrm{i}\pi\lambda\delta\left(2\boldsymbol{K}_{\mathrm{f}}\cdot\boldsymbol{Q}_{\mathrm{p}}+|\boldsymbol{Q}_{\mathrm{p}}|^2\right)\right], \end{aligned} \tag{20}$$

where $a(\boldsymbol{K}_{\mathrm{f}})$ is the real top hat amplitude function defining the extent of the aperture. Thus within each double disk overlap there exists a line along the perpendicular bisector of $\boldsymbol{K}_{\mathrm{f}}$ and $\boldsymbol{Q}_{\mathrm{p}}$, $\boldsymbol{K}_{\mathrm{f}}\cdot\boldsymbol{Q}_{\mathrm{p}} = |\boldsymbol{Q}_{\mathrm{p}}|^2/2$, for which the defocus has no effect. These lines are known as achromatic lines because a chromatic

defocus spread will have no effect along them, and occur when the interfering beams shown in Fig. 2 are equidistant from the optic axis and therefore suffer the same phase shift due to defocus. The achromatic lines exist for every $\boldsymbol{Q}_\mathrm{p}$ due to the cancellation of the spherically symmetric aberrations along the line equidistant from the centers of the $A(\boldsymbol{K}_\mathrm{f})$ and $A^*(\boldsymbol{K}_\mathrm{f}+\boldsymbol{Q}_\mathrm{p})$ terms because of the complex conjugation of the second aperture function in Eq. (20).

As the temporal defocus spread is a fluctuation, it can be informative to consider the summation of two equal and opposite defocus values σ. From Eq. (20) we obtain via trigonometric identity

$$A(\boldsymbol{K}_\mathrm{f})A^*(\boldsymbol{K}_\mathrm{f}+\boldsymbol{Q}_\mathrm{p}) = 2a(\boldsymbol{K}_\mathrm{f})a(\boldsymbol{K}_\mathrm{f}+\boldsymbol{Q}_\mathrm{p}) \times 2\cos\left[\pi\lambda\sigma\left(2\boldsymbol{K}_\mathrm{f}\cdot\boldsymbol{Q}_\mathrm{p}+|\boldsymbol{Q}_\mathrm{p}|^2\right)\right]. \quad (21)$$

The period of this cosine is $2/(\lambda\sigma)$, so for small values of defocus the interference between the two defocused beams causes a slow falloff in the amplitude in the double disk overlap regions. For large values of defocus, an oscillation across the overlap region results. If one increases the number of defoci pairs in the spread, then these different cosine variations in the amplitude interfere, but always retaining an unaffected achromatic line. It has been shown[27,28] that integrating Eq. (20) weighted with physically reasonable distribution of defocus spread gives, for overlaps including the bright-field disc,

$$G_\mathrm{chr}(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f}) = G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f}) \times \exp\left[-\frac{1}{2}\pi^2\lambda^2\Delta^2\left(2\boldsymbol{K}_\mathrm{f}\cdot\boldsymbol{Q}_\mathrm{p}+|\boldsymbol{Q}_\mathrm{p}|^2\right)\right], \quad (22)$$

and overlap regions at higher $\boldsymbol{K}_\mathrm{f}$ values are attenuated by the same envelope if they contribute at the same $\boldsymbol{Q}_\mathrm{p}$ value, as can be observed in Fig. 6. Notice that there is no dependence on the specific defocus value now, but only the chromatic defocus spread itself.

Thus the effect of partial temporal coherence is to modulate $G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f})$ by a frequency dependent 4D envelope function, but it is not a simple envelope that damps higher frequencies to zero exponentially. Examples of the achromatic lines appearing in the amplitude of $G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f})$ are shown in Fig. 5. Moving off an achromatic line perpendicular to it in $\boldsymbol{K}_\mathrm{f}$ there is a Gaussian attenuation of the amplitude depending on the size of the defocus spread and following a $1/\boldsymbol{Q}_\mathrm{p}$ dependence on spatial frequency. The extent of significant amplitude within the double overlap regions can either be limited by extent of the double overlap itself or by the falloff due to the Gaussian attenuation due to partial temporal coherence.

This means if the defocus spread is large enough to be limiting for a set of frequencies, the transfer function itself will exhibit a $1/\boldsymbol{Q}_\mathrm{p}$ dependence. As the size of the double disk overlaps themselves drops off faster than this towards the $\omega = 2$ resolution limit, partial temporal coherence actually minimally affects the transfer of the highest spatial frequencies, significantly less than the midrange frequencies.[27] Beams interfering near to the resolution limit will have passed through the objective lens almost diametrically opposite and therefore experienced the same phase change due to defocus variations and are achromatic.

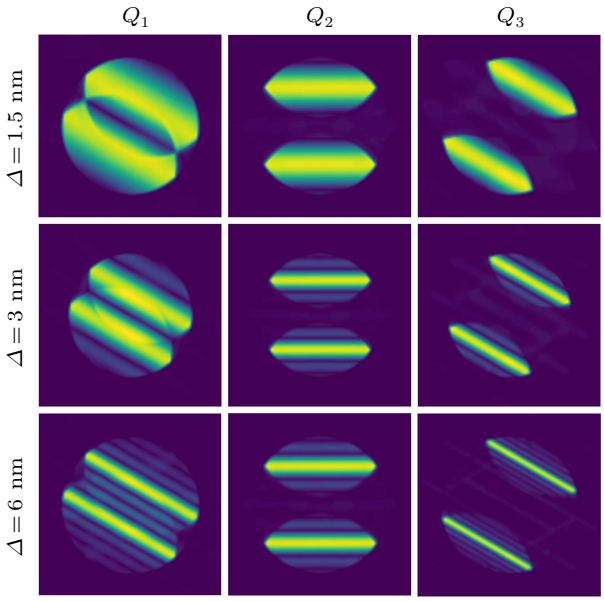


Fig. 5. Achromatic lines appearing in the amplitude of $G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f})$ vs. $\boldsymbol{K}_\mathrm{f}$ displayed for three frequencies with $Q_3 > Q_2 > Q_1$ at three levels of defocus spread as simulated for graphene. The falloff in amplitude moving perpendicular to the achromatic lines can be seen to increase with both increasing defocus spread and increasing spatial frequency.

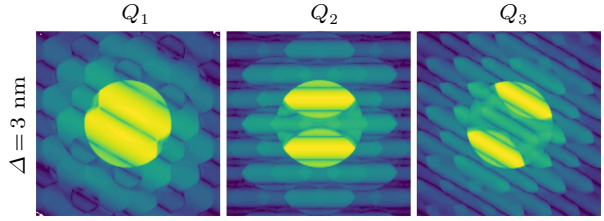


Fig. 6. The same $\Delta = 3$ nm amplitudes of $G(\boldsymbol{Q}_\mathrm{p},\boldsymbol{K}_\mathrm{f})$ Fig. 5 displayed on a log scale and showing a wider range of $\boldsymbol{K}_\mathrm{f}$, allowing one to see how similar attenuation occurs for the overlaps occurring at higher angles.

This is how temporal incoherence reduces the signal in STEM, for both ADF imaging and 4D STEM methods such as ptychography. However presence of the achromatic lines means these methods can be more robust than HRTEM for an equal amount of energy spread, as has been discussed previously.[27–29]

7. Conclusion

Here a common framework has been presented for STEM imaging which allows the links between high-resolution ADF STEM imaging, pioneered by Steve Pennycook, and the emerging field of 4D STEM and in particular ptychography to be made. For example, it is shown that because both ADF

imaging and the SSB method of ptychographical image reconstruction involve integrating in the detector plane over disc overlaps they exhibit similar transfer functions at higher spatial frequencies, though not at lower spatial frequencies due to SSB only being applicable in the WPOA. Similarly, because the iCoM technique integrates over the whole detector, its CTF is identical to the ADF OTF. However, whereas ADF is a direct imaging technique with Poisson distributed noise, and so its OTF is a meaningful quantity describing information transfer, we note that CTFs for 4D STEM phase retrieval techniques are not because the noise is not transferred uniformly with respect to spatial frequency. We also note that the impacts of partial spatial and temporal coherence on ADF and ptychographical images are very similar, and the contribution of achromatic interference in both cases leads to a strong robustness to partial temporal coherence.

It may seem surprising that such clear links can be made between ADF imaging, regarded as an incoherent mode, and ptychography which is regarded as a coherent mode of imaging. However, we have shown that ADF imaging relies on coherent interference between the converging beams forming the illuminating probe and the incoherence of ADF does not rely on intrinsically incoherent scattering. The apparent contradiction can be resolved by using the principle of reciprocity to reverse the beam paths for ADF. Such a configuration would then involve large, incoherent, annular illumination with the image forming lenses after the sample. Such incoherent illumination of objects is the most common in microscopy and in most other forms of imaging.

In the framework presented, important intrinsically incoherent scattering such as TDS has been neglected. An important attribute of ADF imaging is that such intrinsically incoherent scattering can also be collected and used in the image forming process and the formulation. Such scattering is incoherent in part because of the range of scattering momenta that it can generate, and so the similarity to incoherence generated through the integration over final wavevectors performed here can be made.

Finally, we note that we have illustrated some of the limitations of using transfer functions to compare techniques, especially with the emergence of 4D STEM phase retrieval techniques, and that new frameworks will be needed to enable comparisons to be made.

Acknowledgments

We acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme via Grant Agreement No. 802123-HDEM (TJP) and from the UK Engineering and Physical Sciences Research Council (EPSRC) via grant EP/M010708/1 (PDN).

References

[1] Howie A 1979 *Journal of Microscopy* **117** 11
[2] McGibbon A J, Pennycook S J and Angelo J E 1995 *Science* **269** 519
[3] Nellist P D and Pennycook S J 1996 *Science* **274** 413
[4] Müller K, Krause F F, Béché A, Schowalter M, Galioit V, Löffler S, Verbeeck J, Zweck J, Schattschneider P and Rosenauer A 2014 *Nat. Commun.* **5** 5653
[5] Lazic I, Bosch E G T and Lazar S 2016 *Ultramicroscopy* **160** 265
[6] Krivanek O L, Dellby N and Lupini A R 1999 *Ultramicroscopy* **78** 1
[7] Rodenburg J M and Bates R H T 1992 *Phil. Trans. Roy. Soc. Lond. Ser. A: Phys. Eng. Sci.* **339** 521
[8] Hoppe W 1969 *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* **25** 495
[9] Hawkes P W 1982 *Ultramicroscopy* **9** 27
[10] Nellist P D and Pennycook S J 2000 *Advances in Imaging and Electron Physics* **113** 147
[11] Loane R F, Xu P and Silcox J 1992 *Ultramicroscopy* **40** 121
[12] Pennycook S J and Jesson D E 1990 *Physical Review Letters* **64** 938
[13] Nellist P D and Pennycook S J 1999 *Ultramicroscopy* **78** 111
[14] Song W, Perez-Osorio M A, Marie J J, Liberti E, Luo X, O'Leary C, House R A, Bruce P G and Nellist P D 2022 *Joule* **6** 1049
[15] Maiden A M and Rodenburg J M 2009 *Ultramicroscopy* **109** 1256
[16] Maiden A M, Humphry M J and Rodenburg J M 2012 *J. Opt. Soc. Am. A* **29** 1606
[17] Müller-Caspary K, Krause F F, Grieb T, Löffler S, Schowalter M, Beché A, Galioit V, Marquardt D, Zweck J, Schattschneider P, Verbeeck J and Rosenauer A 2017 *Ultramicroscopy* **178** 62
[18] Black G and Linfoot E H 1957 *Math. Phys. Eng. Sci.* **239** 522
[19] Nellist P D and Pennycook S J 1998 *Journal of Microscopy* **190** 159
[20] McGibbon A J, Pennycook S J and Jesson D E 1999 *Journal of Microscopy* **195** 44
[21] Pennycook T J, Lupini A R, Yang H, Murfitt M F, Jones L and Nellist P D 2015 *Ultramicroscopy* **151** 160
[22] Yang H, Rutte R N, Jones L, Simson M, Sagawa R, Ryll H, Huth M, Pennycook T J, Green M L H, Soltau H, Kondo Y, Davis B G and Nellist P D 2016 *Nat. Commun.* **7** 12532
[23] Seki T, Ikuhara Y and Shibata N 2018 *Ultramicroscopy* **193** 118
[24] Dwyer C and Paganin D M 2024 *Phys. Rev. B* **110** 024110
[25] Chen Z, Jiang Y, Shao Y T, Holtz M E, Odstrčil M, Guizar-Sicairos M, Hanke I, Ganschow S, Schlom D G and Muller D A 2021 *Science* **372** 826
[26] Nellist P D, McCallum B C and Rodenburg J M 1995 *Nature* **374** 630
[27] Nellist P D and Pennycook S J 1998 *Phys. Rev. Lett.* **81** 4156
[28] Nellist P D and Rodenburg J M 1994 *Ultramicroscopy* **54** 61
[29] Pennycook T J, Martinez G T, Nellist P D and Meyer J C 2019 *Ultramicroscopy* **196** 131