PHYSICAL REVIEW B **87**, 184108 (2013)

General framework for quantitative three-dimensional reconstruction from arbitrary detection geometries in TEM

Wouter Van den Broek* and Christoph T. Koch
Institute for Experimental Physics, Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany

(Received 20 February 2013; published 13 May 2013)

In Phys. Rev. Lett. **109**, 245502 (2012), a method for retrieving the object's three-dimensional potential distribution by inverting the dynamical scattering was presented and validated by a reconstruction from simulated atomic resolution transmission electron microscopy (TEM) images. In this paper, an extension to ptychography and scanning confocal electron microscopy is demonstrated and validated with simulations. Ultimately, this will make it possible to operate the microscope in the mode that yields the best reconstruction instead of accommodating the microscope settings to the linear approximation to the specimen-electron interaction used in most reconstruction algorithms. Furthermore, simultaneous estimation of the object and the unknown defocus from simulated atomic resolution TEM images is attained. This is an important step towards experimental reconstructions.

DOI: 10.1103/PhysRevB.87.184108 PACS number(s): 61.05.J−, 61.46.−w, 07.05.Mh, 02.30.Zz

I. INTRODUCTION

With the importance of nanomaterials on the rise, accurate and precise characterization becomes more important. Transmission electron microscopy (TEM) is the tool of choice if characterization of individual particles is needed. Imaging two-dimensional projections at atomic resolution has been common practice for years, but making the step to a full three-dimensional (3D) reconstruction remains challenging.

One approach to 3D characterization of the specimen is to match a certain specimen model to the measurements. A familiar example is the defocus-thickness tableaus generated by many image simulation packages, allowing one to retrieve the specimen thickness. In this manner, the specimen-electron interaction is treated correctly by using a multislice (MS) or Bloch wave calculation, and image effects like defocus, spherical aberration, and partial coherence are accounted for. This so-called quantitative approach has solved many problems over the years.

However, since this approach does not produce a true 3D reconstruction, deviations from the proposed model can easily go undetected. Inversion of the image formation process, i.e., direct retrieval of the three-dimensional object from the observations, is therefore desirable. Such techniques do exist in electron microscopy, examples include high angle annular dark field scanning TEM (HAADF STEM) tomography and scanning confocal electron microscopy (SCEM).

In HAADF STEM, the electron beam is condensed in a probe and scanned across the sample. For each probe position the electrons scattered to high angles are integrated, leading to an image with gray values approximately proportional to thickness[1] and to the atomic number raised to a power of about 1.7.[2] The object's 3D density distribution is then retrieved from projections at various object tilts through a tomographic reconstruction algorithm.[3–5] In SCEM[6] the electron beam is condensed in a probe and a second objective lens, placed behind the specimen and steered in concordance with the first, is focused on the beam crossover. A pinhole detector in the image plane selects the central part of the image, thereby yielding a vertical resolution of about 3.5 nm.[7,8]

However, a large part of the available information often remains unused since little knowledge of the state of the microscope can be included in the reconstruction algorithm. Furthermore, the electron-specimen interaction is approximated as a simple linear transformation. This linearity often comes at the cost of a decreased signal: For HAADF STEM and SCEM, for instance, linearity improves with a large detector angle and a narrow pinhole, respectively, both of which decrease the signal and increase the noise.

Even two seminal papers that recently attained atomic resolution tomographic reconstructions,[9,10] do so without making full use of the available prior knowledge and mainly rely on the increased point resolution of HAADF STEM brought forth by the recent improvements in hardware.[11,12] As a result, the samples are severely irradiated and mainly compounds very stable under the beam are investigated.

In this paper we expand the 3D reconstruction technique that recently has been presented in Ref. 13 and that is fully quantitative. The MS algorithm is recast as an artificial neural network (ANN), which allows for a numerically efficient least squares optimization. It takes into account the modulation transfer function (MTF) of the CCD, partial spatial and temporal coherence, inelastic scattering through an absorptive potential, and the complete state of the microscope, like for example the aberration coefficients. Since multiple scattering is accounted for exactly, the technique is dubbed IDES, inverse dynamical electron scattering.

IDES is versatile and modular and can treat any detection geometry in TEM. While in Ref. 13, IDES was validated with an object reconstruction from simulations of conventional high resolution TEM (HRTEM) images, in this paper an extension to coherent diffraction imaging,[14] and more in particular ptychography,[15] and SCEM is presented. Ultimately, this versatility will make it possible to operate the microscope in the mode that yields the best reconstruction instead of compromising on the reconstruction quality by accommodating the microscope settings to the linear approximation to the specimen-electron interaction used in most reconstruction algorithms. Furthermore, simultaneous estimation of the object and the unknown defocus from simulated atomic resolution

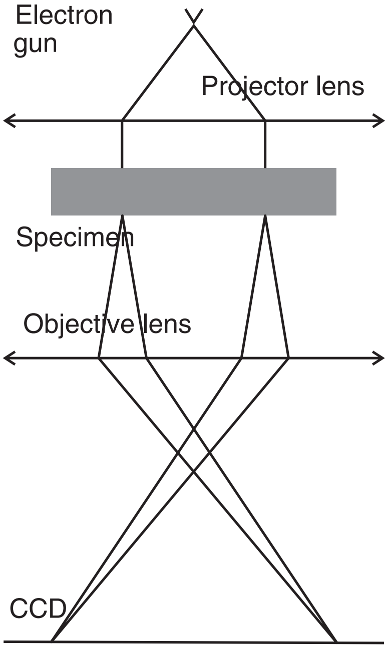

FIG. 1. Simplified ray diagram of the HRTEM geometry. The projector lens transforms the spherical wave emanating from the electron gun to a plane wave that illuminates the specimen. The objective lens produces an image on the CCD.

TEM images is demonstrated. This is an important step towards experimental reconstructions.

In Sec. II the basics of the IDES algorithm are explained for HRTEM, while in Secs. II E and II F more imaging effects are included and the extension to defocus estimation, ptychography, and SCEM is treated. In Sec. III simulations demonstrate the validity of IDES. The results and the technique are discussed in Sec. IV and the conclusions are drawn in Sec. V.

II. THE IDES ALGORITHM

At the core of the IDES algorithm is the MS algorithm that describes the propagation of the electrons through the specimen and thereby takes the dynamical scattering into account. The key realization is that the MS algorithm can be written as an ANN with the weights as a function of the object potential. Then, the backpropagation (BP) algorithm, which is a standard in the field of ANNs, can be adapted to calculate the derivatives of an error function with respect to the potential in each voxel of the object with just one extra pass through the ANN. Subsequently, the object is retrieved by gradient based optimization.

In this section the IDES algorithm is explained in detail: The MS algorithm, the translation of MS to an ANN, and the optimization procedure are covered. These concepts are developed and illustrated with HRTEM in mind, the detection geometry of which is illustrated with a ray diagram in Fig. 1. However, to ensure an accurate and quantitative description a wave optics approach is used in all mathematical derivations. In Secs. II E and II F the imaging model is extended to include defocus estimation, the MTF, incoherent aberrations, ptychography, and SCEM.

A. The MS algorithm

The MS algorithm[16–20] describes the propagation of fast electrons through a solid, neglecting backscattering. The specimen is divided in N slices normal to the beam direction, each supporting a transmission function t. Step j of the MS algorithm consists of multiplying the transmission function with the impinging electron wave function ψ and propagating the result to the next slice by means of Fresnel propagation:

$$\psi_{j+1} = p \otimes (\psi_j t_j), \quad \text{with} \quad t_j = \exp(\imath \sigma V_j), \tag{1}$$

where p is the Fresnel propagator, σ is the interaction constant, V_j is the projected electrostatic potential within slice j, and $\otimes$ denotes convolution. To simplify subsequent derivations, σ is absorbed in V_j for the remainder of this paper.

For HRTEM, the impinging electron wave ψ_1 is a plane wave and the exit wave ψ_{N+1} is transferred to the image plane by the objective lens:

$$\psi_{N+2} = \text{LF} \otimes \psi_{N+1}, \tag{2}$$

where LF is the lens function, i.e., the inverse Fourier transform of the coherent transfer function[20] characterizing the objective lens through the aberration coefficients. An explicit expression for LF is given in Appendix A. The final measurement is the intensity of wave function ψ_{N+2}:

$$I = |\psi_{N+2}|^2. \tag{3}$$

The problem has to be discretized in order to be represented in the computer. In the z direction, parallel to the optical axis, the projected potential is sampled in intervals of Δ_z equal to the distance between successive transmission functions, and indexed by the index j of Eq. (1). In the (x,y) planes normal to the optical axis, the sampling distance is Δ_{xy} in the x and y directions, for notational simplicity, these data points are indexed with a single index k. In the following, the value of the wave function in slice j and sample point k is denoted as ψ_{jk} and the notation $\psi_{j\cdot}$ is used to address the whole of slice j at once.

B. Translation to an ANN

An ANN[21] consists of nodes connected through edges that carry a certain weight, see Fig. 2. Each node represents a function that accepts as its input the output of other nodes, multiplied with the weight of the interconnecting edges. ANNs are most commonly used for classification tasks, e.g., assigning the correct letter of the alphabet to the image of a handwritten letter. The network is trained by feeding it with a large amount of objects of known class, updating the weights of the edges until the classification error is minimal.

The standard way to minimize the error E is through a derivative based optimization. However, numerically computing the derivatives of E with respect to the potential in each voxel would require a prohibitive amount of computation time because each derivative would necessitate a run through the entire network. Fortunately, the backpropagation algorithm of Sec. II C retrieves all derivatives with only one extra pass through the network by cleverly exploiting the chain rule.

In Fig. 2(a) the ANN for HRTEM is given, the description of other recording geometries and further imaging effects is deferred to Secs. II E and II F. The incoming wave ψ_1 is multiplied elementwise with the first transmission function

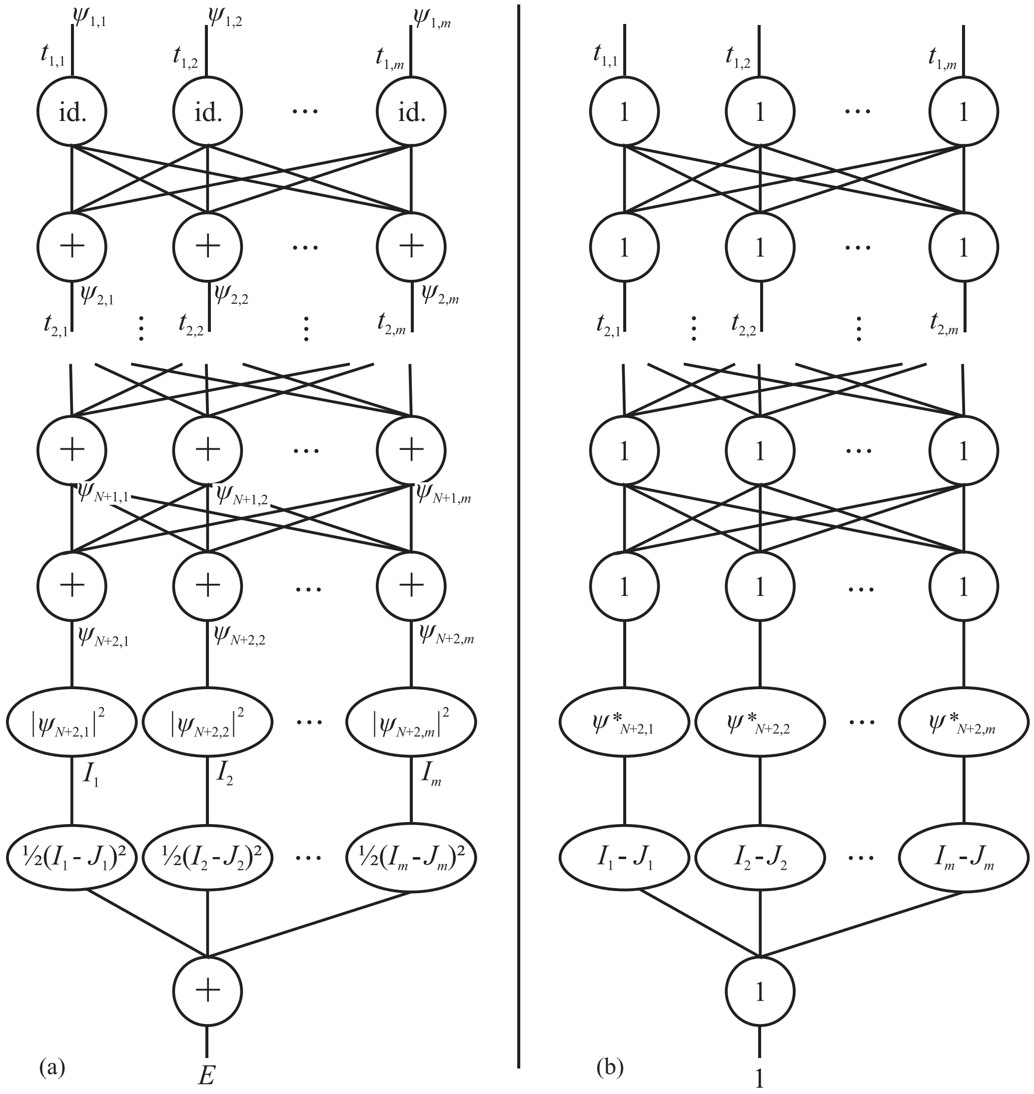

FIG. 2. The ANN and the BP algorithm. (a) The ANN that describes the image formation in HRTEM. The incoming wave ψ_1 is multiplied with the transmission function t_1. The next layer of edges and nodes encodes a real-space convolution with the Fresnel propagator p, resulting in ψ_2. This is repeated until the exit wave ψ_{N+1} is produced. Then, a real space convolution with LF produces ψ_{N+2}, and the intensity I is computed. The last layers compute E from the measurements J. (b) The BP algorithm. Each node contains the derivative with respect to the node input. E is replaced with 1 and the network is run backwards with the same weights, incoming information to each node is added and multiplied with the stored derivatives, resulting in the derivative of E with respect to that node's input.

t_1, therefore, the elements of t_1 act as the weights of the first layer of edges. The nodes in the second layer all represent the identity function, meaning that the output of this layer equals $\psi_1 t_1$. The next layer of edges and nodes encodes a real-space convolution with the Fresnel propagator p, resulting in ψ_2. This pattern is repeated until the exit plane of the specimen is reached and the exit wave ψ_{N+1} is produced. The next layer of edges and nodes realizes a convolution with the lens function and produces ψ_{N+2}. The following layer calculates the intensity $I = |\psi_{N+2}|^2$. The last layers compute the error, which is defined as

$$E = \sum_k \frac{1}{2}(I_k - J_k)^2, \tag{4}$$

with J the measured intensity.

C. Backpropagation algorithm

The backpropagation (BP) algorithm computes the derivatives of E with respect to the inputs of the nodes in only one extra pass through the entire ANN. In this paper only the basics of the BP algorithm are explained, for an in-depth discussion, the reader is referred to Ref. 21.

While passing forward through the ANN, the partial derivatives of the nodes' functions with respect to the inputs are evaluated at the input values and stored in the nodes, see Fig. 2(b). In the backpropagation step, the value of E is replaced with 1 and the network is run backwards with the same weights as before, incoming information to each node is added and the result is multiplied with the stored derivatives. The resulting value at each node is the derivative of E with respect to that node's input.

Although BP yields the derivatives with respect to the node inputs, i.e., $\partial E/\partial(\psi_{jk}t_{jk})$, for the purpose of object retrieval, one needs the derivative with respect to the projected potential:

$$\frac{\partial E}{\partial V_{jk}} = -2\,\mathrm{Im}\left(\psi_{jk}t_{jk}\frac{\partial E}{\partial \psi_{jk}t_{jk}}\right), \tag{5}$$

where $\mathrm{Im}(z)$ takes the imaginary part of z. This result is derived in Appendix B. In Ref. 22 a similar result is arrived upon through a different route.

D. Optimization procedure

In this work, a so-called on-line optimization is chosen in which E is evaluated and minimized serially for each recorded image. Preliminary results have shown that this yields convergence in fewer epochs[23] than an off-line optimization where E is evaluated and minimized over the complete set of recorded images simultaneously.

The error function E is minimized with a steepest descent algorithm. If the index i indicates the iterations of the minimization procedure, then

$$V_{jk}^{(i+1)} = V_{jk}^{(i)} - d^{(i)}\frac{\partial E}{\partial V_{jk}^{(i)}}. \tag{6}$$

The gradient indicates the direction of steepest ascent of E, and the scalar $d^{(i)}$ is the step size. In this paper, $d^{(i)}$ is set with a secant line search[24]: The gradient of E is evaluated in $V_{..}^{(i)}$ and in $V_{..}^{(i)} - d^{(i-1)}\partial E/\partial V_{..}^{(i)}$, and $d^{(i)}$ is set to the position of the minimum of the fitted parabola.

In the field of compressed sensing,[25,26] it is common practice to search for solutions that are maximally sparse in a certain basis set by minimizing the ℓ_1 norm of the coefficients of the object in that basis set, while the measurements act as constraints. The problem at hand is too large to be dealt with by the simplex algorithm, instead the approach of Ref. 27 is adopted, where the error function is ℓ_1 regularized, yielding the total error E_{tot}, and minimized with a steepest descent algorithm:

$$E_{\mathrm{tot}} = \sum_k \frac{1}{2}\left(I_k^{(i)} - J_k\right)^2 + \mu\frac{n_1n_2n_3}{m_1m_2m_3}\sum_{jk}\left|V_{jk}^{(i)}\right|, \tag{7}$$

where μ is a user defined constant, n_1 and n_2 are the number of pixels in the x and y direction of the measurements, n_3 is the number of measured images, m_1 and m_2 are the number of voxels in the x and y direction of the object, and m_3 is the number of slices in the specimen. As shown in Sec. II F, the projected potential is sparse in the voxel basis, especially when a generalized potential is used, and ℓ_1 regularization is therefore justified. The first term of E_{tot} can still be differentiated with BP, while the derivative of the second term can be calculated from

$$\frac{\partial|x|}{\partial x} = \frac{x}{|x|}, \quad \text{with } |x| = \sqrt{x^2+\epsilon^2}, \tag{8}$$

with ϵ^2 a small number of typically 10^{-8}. This redefinition of the absolute value prevents an undefinedness at $x = 0$, while having negligible influence for finite values of x.

Charge flipping (CF)[28] is an *ab initio* structure-determination algorithm that solves the phase problem in x-ray crystallography with much weaker assumptions than classical direct methods. CF iterates back and forth between real space, containing the electron density, and reciprocal space, containing the structure factors, by means of a Fourier transform. With each iteration, the moduli in reciprocal space are replaced with the observed moduli and in real space the sign of the electron density is flipped wherever it is below a certain positive threshold. In practice, CF has proven to successfully solve global, nonconvex optimization problems.

In the reconstruction algorithm presented here, it is made possible to apply flipping to the sign of the reconstructed potential in each iteration, henceforth referred to as potential flipping (PF). The choice of the threshold is subject to debate, but in this work the recommendation in Ref. 28 is followed and it is taken proportional to the standard deviation of the potential in each iteration with a proportionality constant of 1.1.

E. Optimization of the focus

In real life measurements the microscope parameters are often known only approximately. In Refs. 29 and 30 it is demonstrated that ptychographic reconstructions improve drastically if the shape and position of the impinging probe is estimated along with the object.

In this article, the focus values C_1 of the individual images are optimized simultaneously with the object. The derivative of the error function with respect to the focus, i.e., $\partial E/\partial C_1$, is derived in detail in Appendix B and can be included directly in the optimization procedure explained in Sec. II D.

In Fig. 3 a typical error as function of the defocus displays a 6.5 nm wide valley around the minimum, suggesting that the starting value for the focus must be known with an accuracy of about 3 nm. In Sec. III convergence with starting values deviating up to 5 nm from the correct focus is attained.

To give priority to the reconstruction of the object during the first iterations, the values of $\partial E/\partial C_1$ are multiplied with

$$\gamma^{(i)} = 1 - 2^{-i/i_{\mathrm{H}}}, \tag{9}$$

where i denotes the iteration number and i_{H}, fixed to 128 in this paper, is the so-called half-time.

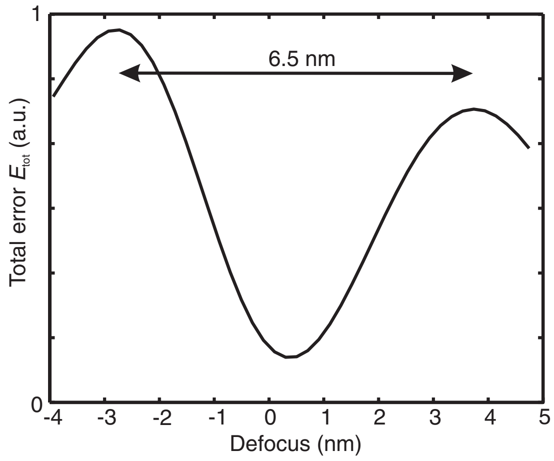

FIG. 3. Error as a function of defocus. The correct defocus resides in a valley of about 6.5 nm wide.

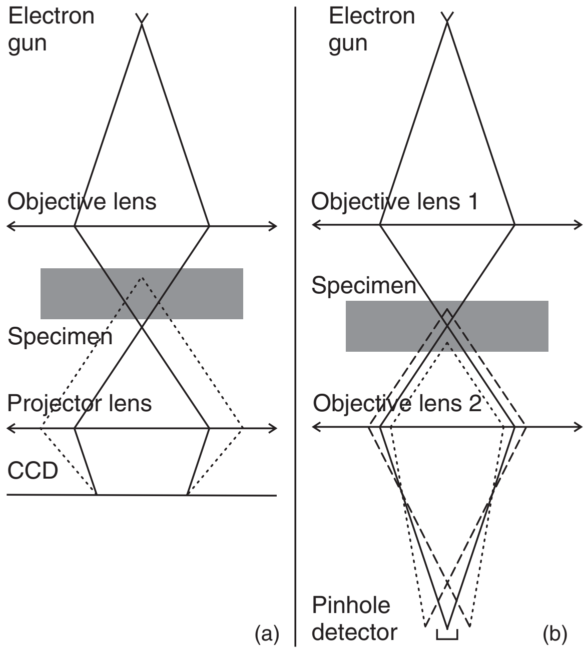

FIG. 4. Simplified ray diagrams. (a) Ptychography. The objective lens illuminates the specimen with a defocused electron probe and the post-specimen projector lens images a diffraction pattern on the CCD. (b) Conventional SCEM. The first objective lens illuminates the specimen with a condensed electron probe and the second objective lens images the beam crossover. The pinhole detector rejects most of the intensity from above and below the crossover, thus creating depth sensitivity.

F. Extensions of IDES

The scheme described so far is easily extended to different recording modes and to include further imaging effects.

In ptychography,[15] the object exit wave is retrieved from a set of diffraction patterns recorded with a defocused electron probe, placed such that the illuminated areas overlap, see Fig. 4(a). In this paper the ptychographic setup is combined with specimen tilt to retrieve the object in 3D. The incoming wave function ψ_1 is now taken as a converged beam. Since in diffraction mode the objective lens transfers ψ_{N+1} to the diffraction plane, the ANN layer in question is made to encode a Fourier transform.

In conventional SCEM[6] the electron beam is condensed in a probe and a second objective lens, placed behind the specimen and steered in concordance with the first, is focused on the beam crossover, see Fig. 4(b). While normally a pinhole detector in the image plane selects the central part of the image, the setup in this paper collects all the electrons in the image plane with a CCD. The incoming wave ψ_1 is taken as a converged beam produced by the objective lens in front of the specimen. The ANN layer following the exit wave ψ_{N+1} is made to encode the lens function of the objective lens behind the specimen.

TABLE I. The simulation parameters for defocus estimation. U is the acceleration voltage, C_1 is the focus value, C_3 is the spherical aberration constant, α_{ill} is the illumination semiangle, Δf is the focal spread, Δ_{pix} is the CCD pixel size, Δ_{xy} is the size of the horizontal dimensions of the voxels of the reconstruction, Δ_z is the slice thickness, and a, c, and d characterize the MTF through Eq. (12).

U	C_1 (nm)	C_3	α_{ill}	Δf	Δ_{pix}
40 kV	−10.15	14 μm	0.1 mrad	1 nm	25 pm
Δ_{xy}	Δ_z	a	c	d	
25 pm	210 pm	0.58	2.7 pix	3.9 pix	

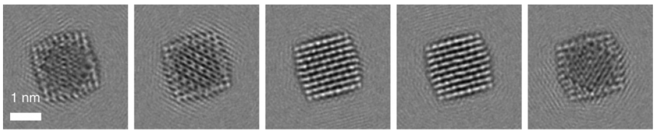

FIG. 5. Five typical simulated HRTEM images. From left to right, the α tilt is -10°, -5°, 0°, $+5^\circ$, and $+10^\circ$ and the β tilt equals 0°.

Partial temporal coherence, characterized by the focal spread Δf, can be modeled approximately by a convolution with

$$\mathcal{F}^{-1}\{\exp[-2(\Delta f \nu^2/\lambda)^2]\} \tag{10}$$

following the lens function.[31] The variable ν is the radial distance in reciprocal space, expressed in radians, and λ is the electron wavelength.

If the spherical aberration is small—as is the case for the aberration corrected machines assumed in this paper—partial spatial coherence, characterized by the illumination convergence semiangle α_{ill}, can be accounted for by an extra convolution of the intensity[32] with

$$\mathcal{F}^{-1}\{\exp[-(\pi\alpha_{\text{ill}}C_1\nu/\lambda)^2]\}. \tag{11}$$

The MTF[33] of the CCD can be accounted for by a convolution of the intensity with the associated point spread function. In this paper, following Ref. 34, the MTF is modeled by the expression

$$a\exp(-c\kappa) + (1-a)\exp(-d^2\kappa^2), \tag{12}$$

with κ as the radial distance in reciprocal space in units of inverted pixel size (pix^{-1}).

As shown in Ref. 20, small object tilts can be approximated by shifting the Fresnel propagator between consecutive slices.

If the recorded intensities are energy filtered, inelastic scattering events can be approximated by expanding the

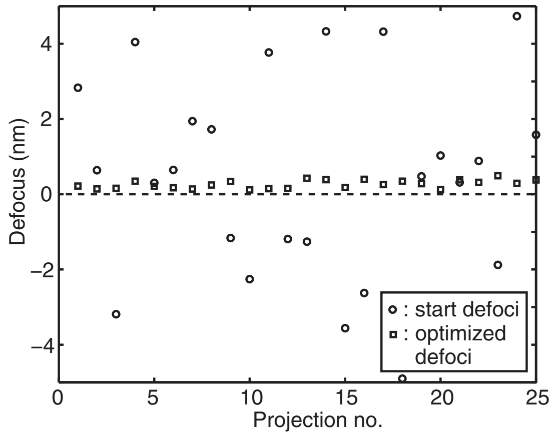

FIG. 6. The defoci at the start of the reconstruction process (*circles*) and at the end (*squares*).

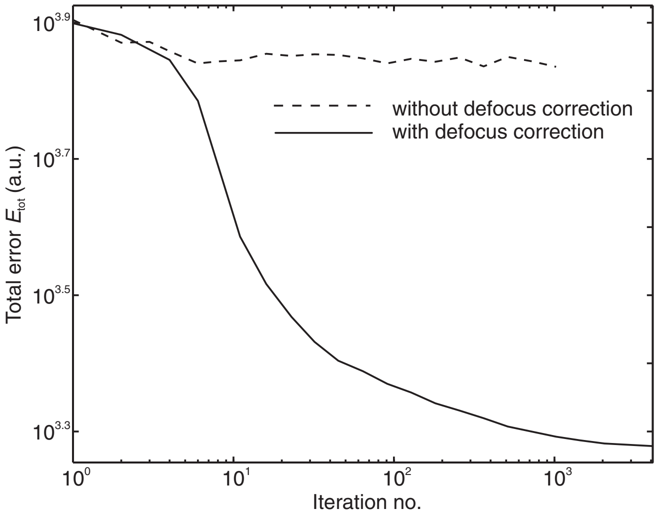

FIG. 7. Reconstruction error with (*full line*) and without (*dashed line*) defocus correction. After 4096 iterations, only little improvement is to be expected.

potential with an imaginary term $\imath W$.[35] This so-called absorptive potential causes a slight damping of the wave upon passing the transmission function. A straightforward adaptation of the derivation in Appendix B yields the derivative of E with respect to the elements of W:

$$\frac{\partial E}{\partial W_{jk}} = -2\,\mathrm{Re}\left(\psi_{jk} t_{jk} \frac{\partial E}{\partial \psi_{jk} t_{jk}}\right), \tag{13}$$

where $\mathrm{Re}(z)$ takes the real part of z.

In Refs. 36,37 it is shown to be highly advantageous to incorporate prior knowledge about the shape of the potential.

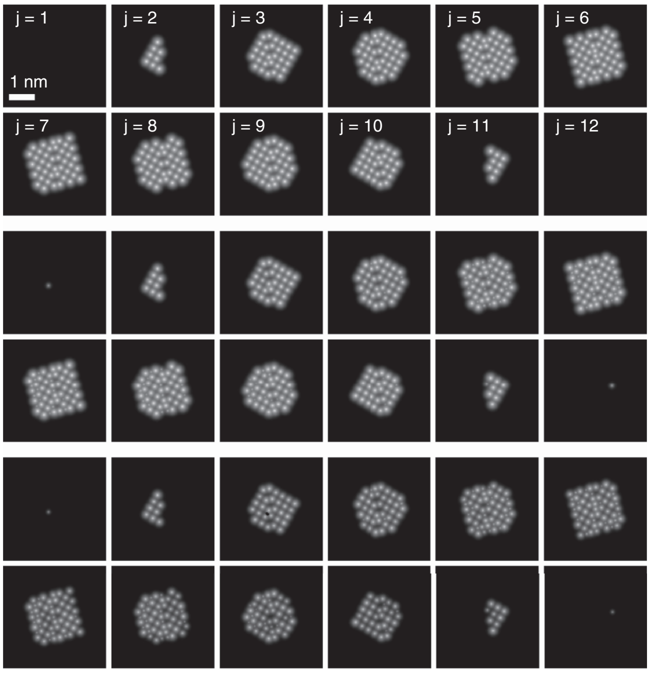

FIG. 8. Reconstruction with defocus correction. *Upper two rows*: Slices with the original potential V. *Middle two rows*: Reconstructed potential V. *Lower two rows*: Reconstructed absorptive potential W.

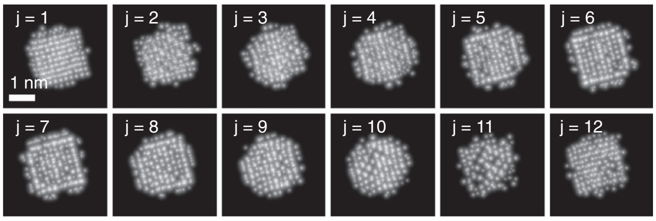

FIG. 9. Reconstruction without defocus correction, compare to the upper two rows of Fig. 8.

Note that

$$V_{jk} + \imath W_{jk} \simeq [V_0 \otimes v_{j\cdot} + \imath W_0 \otimes w_{j\cdot}]_k, \tag{14}$$

where V_0 and W_0 are the generalized regular and absorptive potentials of individual atoms and v and w are approximate arrays of Dirac δ functions centered on the atom positions. If the specimen consists of only one element, V_0 and W_0 are taken as the projected potential of the element and the semiequality becomes exact. In the case of multiple elements a weighted average based on the stoichiometry can be used.

The error metric can then be optimized with respect to v and w instead of V and W, by using

$$\frac{\partial E}{\partial v_{jk}} = \left[\frac{\partial E}{\partial V_{j\cdot}} \otimes V_{0,180^\circ}\right]_k, \tag{15}$$

$$\frac{\partial E}{\partial w_{jk}} = \left[\frac{\partial E}{\partial W_{j\cdot}} \otimes W_{0,180^\circ}\right]_k, \tag{16}$$

where $V_{0,180^\circ}$ and $W_{0,180^\circ}$ equal V_0 and W_0 rotated over 180°. Since v and w are sparse by construction, this constraint can be imposed by ℓ_1 regularization as explained in Sec. II D.

III. SIMULATIONS

The reconstruction scheme has been validated previously in Ref. 13 on simulations of HRTEM images of a Au nanoparticle. In this section the algorithm's versatility is shown with reconstructions from a set of simulated HRTEM images, diffraction patterns, i.e., ptychography, and SCEM images.

In all cases, the test object is an Au cubeoctahedron composed of 309 atoms arranged in an fcc lattice. Dynamical scattering is attained by setting the acceleration voltage to 40 kV.[13] The absorptive potential W is taken as one tenth of the regular potential V.[38] The projected potential for Au, as parametrized by Ref. 20, is chosen as the generalized potentials V_0 and W_0. At the start of the iterative optimization, the potential is set to zero. Neither the fact that $W = V/10$, nor that the particle has an fcc lattice, has been used as prior knowledge in the reconstructions.

TABLE II. The simulation parameters for ptychography. θ_0 is the condenser aperture semiangle, the other quantities are explained in Table I.

U	C_1	C_3	θ_0	Δ_{pix}
40 kV	26 nm	14 μm	30 mrad	0.84 mrad
Δ_{xy}	Δ_z	a	c	d
21 pm	21 pm	0.58	3.2 pix	4.7 pix

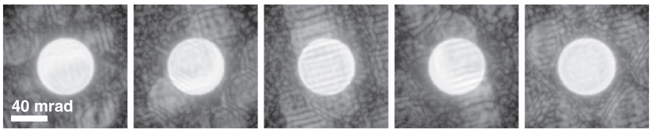

FIG. 10. Five typical diffraction patterns for ptychography, displayed on a logarithmic gray scale. From left to right, the α tilt is -10°, -5°, 0°, $+5^\circ$, and $+10^\circ$ and the β tilt is 0°. Note the noise in the higher frequencies.

The potential V of the original object is compared to V and W of the reconstruction by displaying all slices alongside each other on a logarithmic gray scale. The reconstructed V and W are obtained by convolution of v and w with V_0 and W_0 after intensities below the PF threshold have been set to zero.

A. Focus estimation

A series of 25 HRTEM images is simulated with tilt angles α and β taking the values -10°, -5°, 0°, $+5^\circ$, and $+10^\circ$. In each image, the defocus was set to 0 with respect to the Scherzer focus. The microscope and imaging parameters are summarized in Table I.

The object consists of $320 \times 320 \times 12$ voxels and to avoid wrap-around artifacts, only the central part of the images, measuring 160×160 pixels, is used for reconstruction. Poisson noise, equivalent with a dose of 100 electrons per pixel, has been applied to the intensity right before the MTF. A set of five typical measurements is presented in Fig. 5.

The reconstruction uses potential flipping and the weight μ for the ℓ_1 norm is equal to 1. The 25 starting values for the defoci were chosen randomly between -5 and $+5$ nm, with a mean absolute value of 2.2 nm. After 4096 iterations, the mean absolute defocus is reduced to 0.27 nm, see Fig. 6, and only little improvement in the error is to be expected, see Fig. 7.

The central part (160×160 pixels) of the reconstructions of V and W are displayed in Fig. 8. Both potentials are reconstructed faithfully and all atoms are retrieved. A finite vertical resolution manifests itself as the intensity of the atoms extending into adjacent layers. The residual defocus of 0.27 nm on average does not result in a translation of the reconstruction, as is demonstrated by the reconstruction of the empty upper and lower slices.

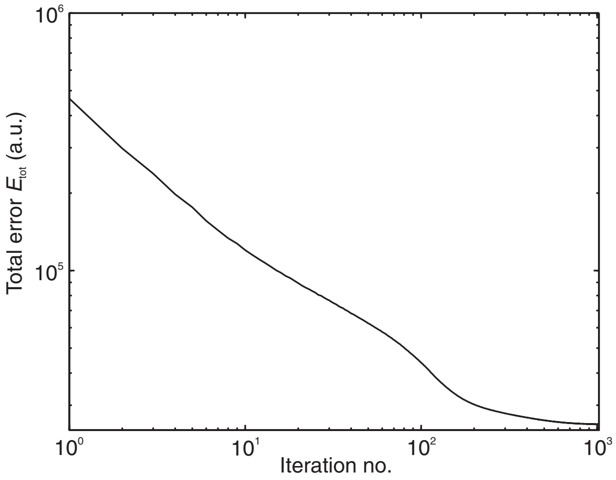

FIG. 11. The error of the ptychographic reconstruction as a function of iteration number. After 1024 iterations, the error has leveled off.

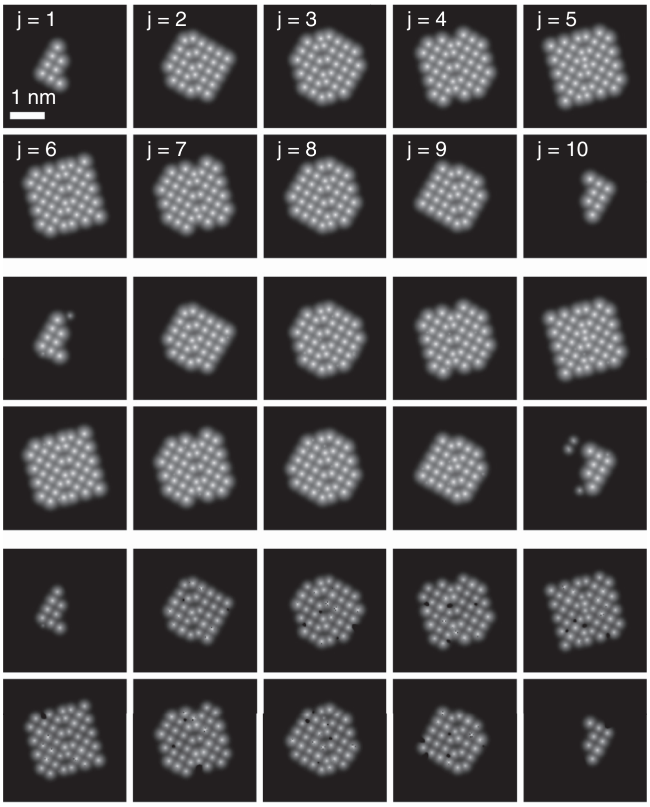

FIG. 12. Ptychographic reconstruction. *Upper two rows*: Slices with the original potential V. *Middle two rows*: Reconstructed potential V. *Lower two rows*: Reconstructed absorptive potential W.

In Fig. 9 a reconstruction with identical parameters, but without defocus optimization, is shown. The error function had leveled off after 1024 iterations, see Fig. 7, and the reconstruction clearly is not physical.

B. Ptychography

The α tilt takes values -10°, -5°, 0°, $+5^\circ$, and $+10^\circ$, while the β tilt equals 0°, and, vice versa, the β tilt takes values -10°, -5°, $+5^\circ$, and $+10^\circ$, while the α tilt equals 0°. At each tilt, 14 diffraction patterns are recorded with a defocused beam that is placed on random but known positions, resulting in a

TABLE III. The simulation parameters for SCEM. C_3^{hi}, C_3^{lo} and θ_0^{hi}, θ_0^{lo} are the spherical aberration constants and the aperture semiangles of the lenses above and below the specimen. The other quantities are explained in Table I.

U	C_3^{hi}	C_3^{lo}	θ_0^{hi}	θ_0^{lo}	
40 kV	0.0 μm	0.0 μm	50 mrad	50 mrad	
Δ_{pix}	Δ_{xy}	Δ_z	a	c	d
30 pm	30 pm	210 pm	0.58	2.2 pix	3.8 pix

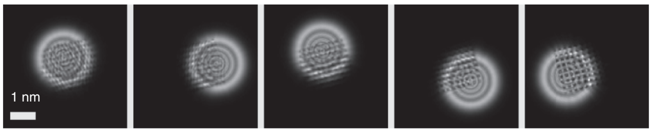

FIG. 13. Five typical SCEM measurements, displayed on a linear gray scale. From left to right, the β tilt equals -10°, -5°, 0°, $+5^\circ$, and $+10^\circ$ and the α tilt equals 0°.

total of 126 diffraction patterns. The microscope and imaging parameters are summarized in Table II.

The object consists of $342 \times 342 \times 10$ voxels and to avoid aliasing artifacts, only the central part of the diffraction patterns, measuring 160×160 pixels, is used for reconstruction. Poisson noise, equivalent with a dose of 225 electrons per pixel in the central disk, is applied to the intensity right before the MTF. In Fig. 10 five typical diffraction patterns are shown.

Diffraction patterns have a large dynamical range and the region outside the central disk will therefore make a negligible contribution to the first term of the error function E_{tot} in Eq. (7). This term has therefore been changed to

$$\sum_k \frac{1}{2J_k}\left(I_k^{(i)} - J_k\right)^2. \tag{17}$$

No potential flipping was needed for this reconstruction and the weight μ for the ℓ_1 norm was set equal to 1. In Fig. 11 it can be seen that after 1024 iterations the error has leveled off. The central part (160×160 pixels) of the reconstructions of V and W are displayed in Fig. 12. Both potentials are reconstructed faithfully and all atoms are retrieved. A finite vertical resolution can be noticed as the intensity of the atoms extends into adjacent layers.

C. Scanning confocal electron microscopy

The upper lens illuminates the specimen with a beam defocused by 25 nm with respect to the particle, and the lower lens is kept focused on the particle center. The specimen is tilted: The α tilt is set to -10°, -5°, 0°, 5°, and 10°, with the β tilt equal to 0°. At each tilt the probe is positioned at 25 random but known positions and an image is recorded with a CCD.

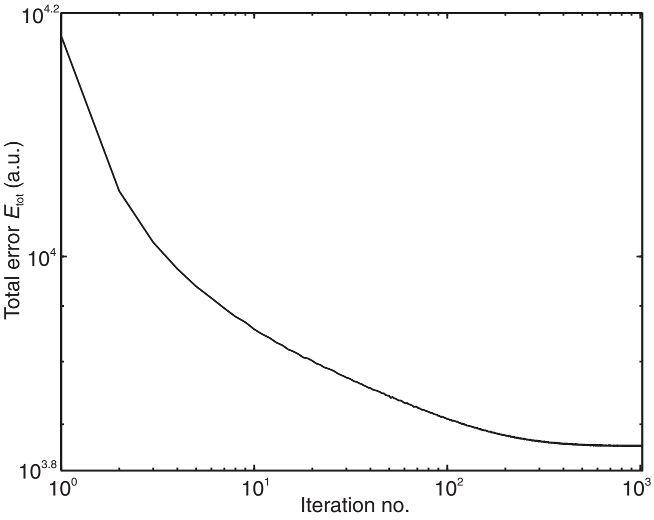

FIG. 14. The error of the SCEM reconstruction as a function of iteration number. After 1024 iterations, the error has leveled off.

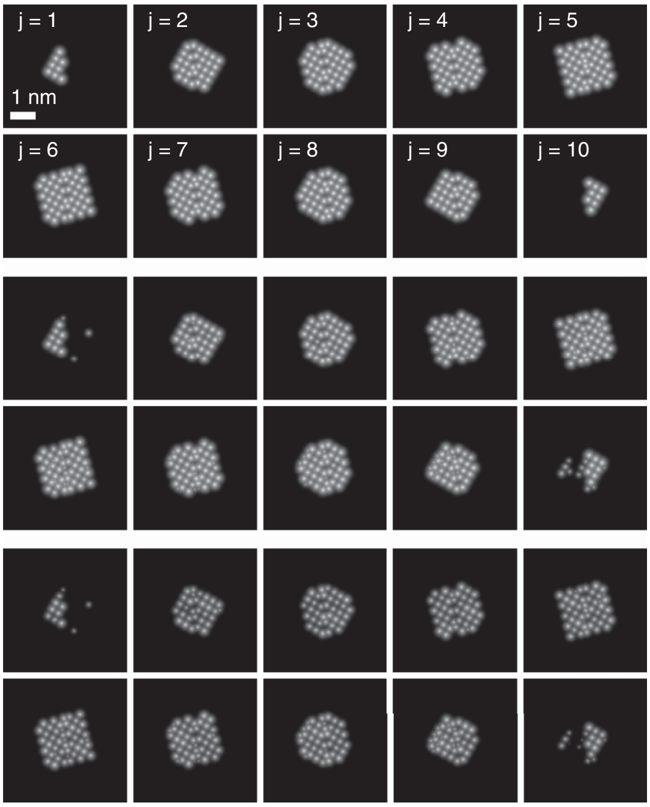

FIG. 15. SCEM reconstruction. *Upper two rows*: Slices with the original potential V. *Middle two rows*: Reconstructed potential V. *Lower two rows*: Reconstructed absorptive potential W.

The object consists of $320 \times 320 \times 10$ voxels and to avoid wrap-around artifacts, only the central part of the images, measuring 160×160 pixels, is used for reconstruction. The microscope and imaging parameters are listed in Table III. A set of five typical measurements is presented in Fig. 13 and no noise has been added.

The weight μ of the ℓ_1 norm is set to 1 and potential flipping is applied. In Fig. 14 it can be seen that after 1024 iterations the error has leveled off. The central part (160×160 pixels) of the reconstructions of V and W are displayed in Fig. 15. Both potentials are reconstructed faithfully and all atoms are retrieved. A finite vertical resolution is noticeable as the intensity of the atoms extends into adjacent layers.

The authors could not find a setup that yielded a SCEM reconstruction without allowing specimen tilt; however, it is possible that this is evidence of an inherent ill-definedness of the SCEM geometry, rather than of a shortcoming of the IDES algorithm.

IV. DISCUSSION

Since the tilt range can be kept relatively low (±10°) it could become possible to realize the tilt electronically through a beam tilt instead of physically through a specimen tilt. This has the advantage of increasing the speed and accuracy of the tilts, but the disadvantage is that this would cause larger aberrations.

However, if these larger aberrations can be quantified, they could be plugged in the expression for the lens function.

The number of knowns in the reconstruction problems in Secs. III A and III C is 52% and 63% of their respective number of unknowns. Therefore, unique solutions do not exist. However, due to the object's atomicity, the physically correct solution must be sparse, and the ℓ_1 regularization does just that: Restricting the solution space to sparse objects. In Ref. 25 an experimental reconstruction where the number of knowns is only 39% of the number of unknowns is demonstrated. Another successful reconstruction of a sparse object from a limited amount of data can be found in Ref. 39.

The optimization procedure makes use of two user-defined constants: the weight μ for the ℓ_1 norm and the threshold for PF. In Eq. (7) the former is normalized such that for μ a value of the order of 1 is expected. And indeed, this value yields good results in all three simulations. The threshold for PF was set to 1.1 times the standard deviation of the iterative solution, as recommended in Ref. 28. Since this value readily proofed to work well, no attempt was made to optimize it.

V. CONCLUSIONS

In this paper the workings of the IDES algorithm have been detailed. It has been shown through simulations that the defocus of each image in an HRTEM series can be estimated simultaneously with the object. This is an important step towards experimental reconstructions since in practice many microscope parameters are known only approximately.

Arbitrary detection geometries can be modeled. Ptychography simulations yielded a faithful three-dimensional reconstruction at the atomic level, even with noise in the diffraction patterns. An atomic resolution reconstruction was found from SCEM simulations as well. This will make it possible to operate the microscope in the mode that yields the best reconstruction instead of accommodating the microscope settings to the linear approximation to the specimen-electron interaction used in most reconstruction algorithms.

In the reconstruction, prior knowledge about the sparseness and the positivity of the object are used by means of ℓ_1 regularization and potential flipping. This puts a heavy, yet physically justified, constraint on the solution, allowing a reconstruction from a limited tilt range and a limited number of unknowns. No prior knowledge about the crystal structure of the particle has been used; therefore, phenomena like point defects, dislocations, or relaxation of the lattice parameters can be detected.

ACKNOWLEDGMENTS

The authors acknowledge the Carl Zeiss Foundation as well as the German Research Foundation (DFG, Grant No. KO 2911/7-1).

APPENDIX A: LENS FUNCTION

In this Appendix an explicit expression is given for the lens function LF used in Eq. (2).

The aberration function χ is expressed as

$$\chi = \tfrac{1}{2}C_1\nu^2 + \tfrac{1}{4}C_3\nu^4, \tag{A1}$$

where ν is the radial distance in reciprocal space, in radians, and C_1 and C_3 are the focus and the spherical aberration coefficient. In Ref. 40 a complete expansion, including noncircular symmetric terms, is given.

The coherent transfer function is defined as

$$\mathrm{CTF} = \exp(\imath 2\pi\chi/\lambda), \tag{A2}$$

with λ as the wavelength of the electrons. Finally, the lens function is

$$\mathrm{LF} = \mathcal{F}^{-1}(\mathrm{CTF}). \tag{A3}$$

APPENDIX B: DERIVATIVE OF THE ERROR FUNCTION

In this Appendix the result of Eq. (5), the derivative of E with respect to object potential, and the derivative with respect to the focus, is derived.

1. Derivative with respect to the object potential

Originally, the BP algorithm is formulated for real valued networks and it has to be adapted to complex numbers. In complex calculus, the chain rule reads as

$$\frac{\partial E}{\partial V_{jk}} = \frac{\partial E}{\partial \psi_{jk}t_{jk}}\frac{\partial \psi_{jk}t_{jk}}{\partial V_{jk}} + \frac{\partial E}{\partial (\psi_{jk}t_{jk})^*}\frac{\partial (\psi_{jk}t_{jk})^*}{\partial V_{jk}}, \tag{B1}$$

where the asterisk denotes the complex conjugate.

Denote two artificial neural networks as ANN and ANN*, both have the same structure and describe the same image formation process, but the incoming wave and the weights in ANN* are the complex conjugates of those in ANN. Because both produce the same values for the intensities and for E, application of the BP algorithm to ANN* yields the derivative of E with respect to the complex conjugates of the node inputs of ANN. On the other hand, inspection of Fig. 2(b) shows that the only operations O encountered in BP are addition and multiplication, both of which obey

$$O(z^*) = (O(z))^*, \quad \text{with} \quad z \in \mathbb{C}. \tag{B2}$$

Hence, BP of ANN* yields the complex conjugate of BP of ANN, and therefore,

$$\frac{\partial E}{\partial (\psi_{jk}t_{jk})^*} = \left(\frac{\partial E}{\partial \psi_{jk}t_{jk}}\right)^*. \tag{B3}$$

Substitution of Eq. (B3) in Eq. (B1), and using the definition in Eq. (1), yields the result in Eq. (5) through straightforward algebra.

2. Derivative with respect to the focus

In Fig. 16 the part of the ANN between the exit wave $\psi_{N+1,\cdot}$ and the wave emanating from the lens $\psi_{N+2,\cdot}$ is shown in detail. The elements of $\psi_{N+1,\cdot}$ and $\psi_{N+2,\cdot}$ are indexed with k and ℓ, respectively. The weights $u_{k\ell}$ of the interconnecting edges encode a real-space convolution with the lens function LF; hence, $u_{k\ell} = \mathrm{LF}_{k-\ell}$. To simplify further notations, define $s_{k\ell} = \psi_{N+1,k}u_{k\ell}$.

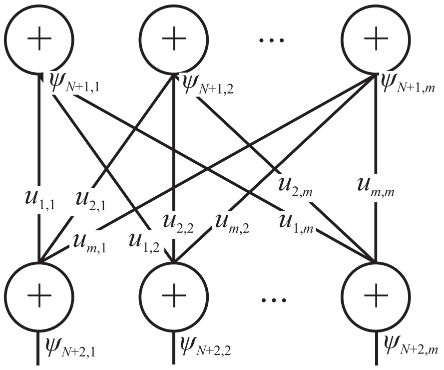

FIG. 16. Detail of the ANN, the weights $u_{k\ell}$ encode a convolution of the object exit wave with the lens function LF. The elements of the upper layer are indexed with k, those of the lower layer with ℓ. The weights $u_{k\ell}$ equal $\mathrm{LF}_{k-\ell}$.

Since each of the weights $u_{k\ell}$ depends on the focus C_1, the chain rule for complex analysis yields

$$\frac{\partial E}{\partial C_1} = \sum_{k\ell} \frac{\partial E}{\partial s_{k\ell}} \frac{\partial s_{k\ell}}{\partial C_1} + \frac{\partial E}{\partial s_{k\ell}^*} \frac{\partial s_{k\ell}^*}{\partial C_1} \tag{B4}$$

$$= 2\,\mathrm{Re}\left(\sum_{k\ell} \frac{\partial E}{\partial s_{k\ell}} \frac{\partial s_{k\ell}}{\partial C_1}\right), \tag{B5}$$

where $\mathrm{Re}(z)$ takes the real part of z and Eq. (B5) follows from Eq. (B3). The derivative $\partial E/\partial s_{k\ell}$ results from BP and can therefore be considered as known; furthermore, since $\partial E/\partial s_{k\ell}$ is independent of k one can set $k = \ell$,

$$\frac{\partial E}{\partial C_1} = 2\,\mathrm{Re}\left(\sum_{\ell} \frac{\partial E}{\partial s_{\ell\ell}} \sum_{k} \frac{\partial s_{k\ell}}{\partial C_1}\right), \tag{B6}$$

and invoking the definition for $s_{k\ell}$ yields

$$\sum_k \frac{\partial s_{k\ell}}{\partial C_1} = \sum_k \psi_{N+1,k} \frac{\partial \mathrm{LF}_{k-\ell}}{\partial C_1} \tag{B7}$$

$$= \left[\psi_{N+1,\cdot} \otimes \frac{\partial \mathrm{LF}_{\cdot}}{\partial C_1}\right]_\ell. \tag{B8}$$

Finally, by using the definitions in Appendix A, one can write

$$\frac{\partial \mathrm{LF}_\ell}{\partial C_1} = \frac{\imath\pi}{\lambda}[\mathcal{F}^{-1}(\nu^2\mathrm{CTF})]_\ell. \tag{B9}$$

The dependence of the error on C_1 through the partial spatial coherence, as quantified by Eq. (11), was neglected in this derivation. But a comparison with a numerical derivative showed that, for the microscope and imaging parameters given in Table I, this approximation produces a negligible error.

---

[*] wouter.vandenbroek@uni-ulm.de

[1] W. Van den Broek, A. Rosenauer, B. Goris, G. Martinez, S. Bals, S. Van Aert, and D. Van Dyck, Ultramicroscopy **116**, 8 (2012).

[2] P. Hartel, H. Rose, and C. Dinges, Ultramicroscopy **63**, 93 (1996).

[3] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (Society of Industrial and Applied Mathematics, Philadelphia, 2001).

[4] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction* (SIAM, Philadelphia, 2001).

[5] P. A. Midgley and M. Weyland, Ultramicroscopy **96**, 413 (2003).

[6] S. P. Frigo, Z. H. Levine, and N. J. Zaluzec, Appl. Phys. Lett. **81**, 2112 (2002).

[7] P. D. Nellist, G. Behan, A. I. Kirkland, and C. J. D. Hetherington, Appl. Phys. Lett. **89**, 124105 (2006).

[8] E. C. Cosgriff, A. J. D'Alfonso, L. J. Allen, S. D. Findlay, A. I. Kirkland, and P. D. Nellist, Ultramicroscopy **108**, 1558 (2008).

[9] S. Van Aert, K. J. Batenburg, M. D. Rossell, R. Erni, and G. Van Tendeloo, Nature (London) **470**, 374 (2011).

[10] M. C. Scott, C. Chen, M. Mecklenburg, C. Zhu, R. Xu, P. Ercius, U. Dahmen, B. C. Regan, and J. Miao, Nature (London) **483**, 444 (2012).

[11] M. Haider, H. Rose, S. Uhlemann, E. Schwan, B. Kabius, and K. Urban, Ultramicroscopy **75**, 53 (1998).

[12] O. L. Krivanek, N. Dellby, and A. R. Lupini, Ultramicroscopy **78**, 1 (1999).

[13] W. Van den Broek and C. T. Koch, Phys. Rev. Lett. **109**, 245502 (2012).

[14] J. C. H. Spence, in *Science of Microscopy*, edited by P. W. Hawkes and J. C. H. Spence (Springer, New York, 2007), pp. 1196–1227.

[15] M. J. Humphry, B. Kraus, A. C. Hurst, A. M. Maiden, and J. M. Rodenburg, Nat. Commun. **3**, 730 (2012).

[16] J. M. Cowley and A. F. Moodie, Acta Crystallogr. **10**, 609 (1957).

[17] J. G. Allpress, E. A. Hewat, A. F. Moodie, and J. V. Sanders, Acta Crystallogr. Sect. A **28**, 528 (1972).

[18] D. F. Lynch and M. A. O'Keefe, Acta Crystallogr. Sect. A **28**, 536 (1972).

[19] P. Goodman and A. F. Moodie, Acta Crystallogr. Sect. A **30**, 280 (1974).

[20] E. J. Kirkland, *Advanced Computing in Electron Microscopy*, 2nd ed. (Springer, New York, 2010).

[21] R. Rojas, *Neural Networks—A Systematic Introduction* (Springer, Berlin, 1996).

[22] M. Lentzen and K. Urban, Acta Crystallogr. Sect. A **56**, 235 (2000).

[23] An epoch is the time taken to go through all the measurements once.

[24] J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer, New York, 1999).

[25] R. G. Baraniuk, IEEE Signal Process. Mag. **24**, 118 (2007).

[26] J. Romberg, IEEE Signal Process. Mag. **25**, 14 (2008).

[27] S. J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, IEEE J-STSP **1**, 606 (2007).

[28] G. Oszlányi and A. Sütő, Acta Crystallogr. Sect. A **64**, 123 (2008).

[29] A. M. Maiden and J. M. Rodenburg, Ultramicroscopy **109**, 1256 (2009).

[30] A. Maiden, M. Humphry, M. Sarahan, B. Kraus, and J. Rodenburg, Ultramicroscopy **120**, 64 (2012).

[31] K. Ishizuka, Ultramicroscopy **5**, 55 (1980).

[32] C. T. Koch, Ultramicroscopy **108**, 141 (2008).

[33] I. Daberkow, K. H. Herrmann, L. Liu, and W. D. Rau, Ultramicroscopy **38**, 215 (1991).

[34] W. Van den Broek, S. Van Aert, and D. Van Dyck, Microsc. Microanal. **18**, 336 (2012).

[35] K. Ishizuka, Ultramicroscopy **90**, 71 (2002).

[36] W. Van den Broek, S. Van Aert, and D. Van Dyck, Ultramicroscopy **109**, 1485 (2009).

[37] W. Van den Broek, S. Van Aert, and D. Van Dyck, Ultramicroscopy **110**, 548 (2010).

[38] D. Paganin, S. C. Mayo, T. E. Gureyev, P. R. Miller, and S. W. Wilkins, J. Microsc. **206**, 33 (2002).

[39] B. Goris, S. Bals, W. Van den Broek, E. Carbó-Argibay, S. Gómez-Graña, L. M. Liz-Marzán, and G. Van Tendeloo, Nat. Mater. **11**, 930 (2012).

[40] M. Haider, H. Müller, S. Uhlemann, J. Zach, U. Loebau, and R. Hoeschen, Ultramicroscopy **108**, 167 (2008).