

Phase-contrast imaging of multiply-scattering extended objects at atomic resolution by reconstruction of the scattering matrix

Philipp M. Pelz^{1,2,*} Hamish G. Brown,² Scott Stonemeyer^{3,4,5,6} Scott D. Findlay⁷ Alex Zettl,^{3,5,6} Peter Ercius^{6,2} Yaqian Zhang,¹ Jim Ciston,² M. C. Scott,^{1,2} and Colin Ophus^{2,†}

¹Department of Materials Science and Engineering, University of California, Berkeley, California 94720, USA

²National Center for Electron Microscopy, Molecular Foundry, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA


³Department of Physics, University of California, Berkeley, California 94720, USA

⁴Department of Chemistry, University of California, Berkeley, California 94720, USA

⁵Kavli Energy NanoSciences Institute, University of California, Berkeley, California 94720, USA

⁶Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA

⁷School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia

 (Received 29 December 2020; revised 5 January 2021; accepted 15 April 2021; published 28 May 2021)

Three-dimensional phase-contrast imaging of multiply-scattering samples in x-ray and electron microscopy is challenging due to small numerical apertures, the unavailability of wave front shaping optics, and the highly nonlinear inversion required from intensity-only measurements. In this work, we present an algorithm using the scattering matrix formalism to solve the scattering from a noncrystalline medium from scanning diffraction measurements and simultaneously recover the illumination aberrations. We demonstrate our method experimentally in a scanning transmission electron microscope, recovering the scattering matrix of a heterogeneous sample with two layers of multiwall carbon nanotubes filled with TaTe₂ core-shell structures, spaced 10 nm apart in the axial direction. Our work enables phase contrast imaging and materials characterization in multiply-scattering samples at high resolution for a wide range of materials.

DOI: [10.1103/PhysRevResearch.3.023159](https://doi.org/10.1103/PhysRevResearch.3.023159)

I. INTRODUCTION

Phase-contrast imaging is widely used in light [1,2], x-ray [3,4], and electron microscopy [5,6] due to its high efficiency and resolution. By using coherent radiation to illuminate a sample, we can resolve very small changes in a sample's local index of refraction through the interference of the illumination wave fronts that the accumulated phase shifts produce [7]. However, because we can directly measure only the probability density of an illumination wave function (given by the wave intensity, or amplitude squared), phase-contrast imaging is a fundamentally nonlinear measurement process: we must indirectly infer the underlying relative phase shifts induced by the sample [8].

Various approximations can make phase-contrast microscopy data easier to interpret. The first is by assuming that the sample is a pure phase object, i.e., that it does not modulate the illumination wave function amplitude directly, and so any variations in the measured intensity can be directly ascribed to changes in the sample's index of refraction [9]. However, this assumption does not guarantee uniqueness in all cases due

to the possibility of phase wrapping [10]. An even stronger assumption is the weak phase object approximation (WPOA), where the sample's transmission function is assumed to be a small imaginary perturbation to a known carrier wave [11]. When the WPOA holds, the linear relation implied between the specimen potential and measured intensity allows a constructive and unambiguous solution. Another commonly used simplification in phase-contrast microscopy is the projection approximation, in which all scattering is assumed to originate from an infinitesimally thin two-dimensional (2D) plane [12,13]. The various different approximations above hold for a wide range of samples of interest and are therefore very useful in practice [14].

However, phase-contrast imaging of many samples cannot be approximated by any of the above assumptions. Transmission electron microscopy (TEM) in particular often violates these assumptions due to the high-scattering cross section of electrons with matter [15]. Instead, these scattering processes can typically be modeled only by a framework that includes multiple scattering [16]. The equations describing multiple scattering for a paraxial wave function can be approximately solved with the multislice algorithm [17], which has also been used as a model for inverse scattering in many experimental configurations in light, x-ray, and electron microscopy. While the inverse multislice model has been successfully applied to image thick, multiply-scattering specimens in light microscopy [18–22], its use in x-ray [23–26] and electron microscopy [27–30] has been limited to proof-of-principle demonstrations with fewer than 10 slices or weakly scattering

*philipp.pelz@yahoo.de

†cophus@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

samples. This is mainly due to the fact that the optical systems in x-ray and electron microscopy have relatively small numerical apertures, such that the information recorded from a single view covers only a small fraction of reciprocal space [31–33]. This problem can be overcome either by enforcing strong prior knowledge about the underlying scattering potential in the form of sparsity constraints or the proper choice of slice separation [30] or by performing tomographic experiments [34–36].

Another framework that incorporates multiple scattering is the scattering matrix (\mathcal{S} -matrix) formalism [37,38]. In electron microscopy, the \mathcal{S} -matrix formalism has been used to efficiently calculate diffraction results with single crystals [38] and for scanning TEM (STEM) experiments [39,40] and to retrieve projected potentials of strongly scattering samples in a two-step approach. First, the \mathcal{S} matrix is retrieved from a series of intensity measurements. Second, the projected structure is retrieved. The proposed experimental methods for retrieval of the \mathcal{S} matrix from intensity measurements range from measurements with different crystal thicknesses and sample tilts [41] to different sample tilts alone [42–44], wavelength variation [45], large-angle rocking beam diffraction [46], and scanning diffraction with a convergent beam [47]. Only the last two of these approaches have been experimentally demonstrated [46,48].

In the visible light wavelengths, \mathcal{S} -matrix retrieval and subsequent singular value decomposition allow the identification of transmission eigenchannels [49] in strongly scattering materials and maximization of energy transport [50] through the system. Phase retrieval of the \mathcal{S} matrix is performed by real-space phase [51] or amplitude modulation [52,53], four-phase interferometry [49], or full-field Mach-Zehnder interferometry [54] with input and output channels in the plane-wave basis. The input and output channels of the \mathcal{S} matrix are often represented in real space, achieved by imaging the output plane with a CCD camera.

Our contribution in this work is threefold: first, we develop the measurement operator to calculate scanning diffraction intensities of arbitrary samples from a given \mathcal{S} matrix and derive its adjoint operator. Second, we formulate a phase retrieval algorithm that retrieves the \mathcal{S} matrix of arbitrary samples from a series of scanning diffraction measurements with different modulations of the illumination aperture (e.g., a defocus series). Third, we demonstrate our algorithm experimentally on a thick, heterogeneous sample, opening up possibilities of materials characterization and imaging with the scattering matrix.

While here we synthesize large-scale phase-contrast images from a strongly scattering sample with a thickness of 1.2 times the depth of focus, optical sectioning of the scattering matrix with a limited field of view was recently experimentally shown to provide three-dimensional information on the unit-cell scale in a strongly scattering sample that spans multiple depths of focus [55].

II. RECONSTRUCTING THE \mathcal{S} MATRIX

A. Theory of phase-contrast imaging

Phase-contrast microscopy with coherent light or matter waves defined by the wave function $|\psi\rangle_{\mathbf{r}}$ typically uses a

series of interferometric measurements to invert a partial differential equation of the form

$$i[a\nabla_{\perp}^2 + bV(\mathbf{r}')]| \psi \rangle_{\mathbf{r}'} = \frac{\partial |\psi\rangle_{\mathbf{r}'}}{\partial z}, \quad (1)$$

where i is the imaginary constant, ∇_{\perp}^2 is the two-dimensional Laplace operator, $V(\mathbf{r}')$ is the three-dimensional potential over the real-space coordinates $\mathbf{r}' = (\mathbf{r}, z)$, and a and b are real-valued constant prefactors. The formal operator solution to this equation for a wave function that has propagated a distance Δz through the potential is given by [56]

$$|\psi\rangle_{(\mathbf{r}, z+\Delta z)} = \exp[ia\Delta z\nabla_{\perp}^2 + ibV_{\Delta z}(\mathbf{r}, z)]|\psi\rangle_{\mathbf{r}'}. \quad (2)$$

In the scattering matrix formalism, the entire process of multiple scattering is modeled by multiplication with the complex-valued linear operator \mathcal{S} ,

$$|\psi\rangle_{\text{out}} = \mathcal{S}|\psi\rangle_{\text{in}}. \quad (3)$$

The \mathcal{S} -matrix formalism has a wide range of applications in describing the interaction of coherent waves with multiply scattering objects [57].

All the previously discussed methods for \mathcal{S} -matrix retrieval at high resolution have in common that they require a crystalline sample to solve for either the scattering matrix or the structure factors. The interferometric methods developed for light optics rely on the ability to precisely manipulate phases and/or amplitudes of the \mathcal{S} -matrix input channels, and such precise control of the electron and x-ray optics is not yet feasible. In the following section, we describe our iterative reconstruction scheme from scanning diffraction measurements for \mathcal{S} -matrix retrieval.

B. A real-space \mathcal{S} -matrix measurement model

Previous work for retrieving the \mathcal{S} matrix from scanning diffraction measurements modeled the formation of the diffraction pattern intensity in the far field of the sample, given a coherent probe $|\psi\rangle$ at position ρ ,

$$|\psi\rangle_{\mathbf{r}-\rho} = \sum_{|\mathbf{h}| < h_{\text{max}}} \Psi(\mathbf{h}) e^{2\pi i \mathbf{h} \cdot (\mathbf{r}-\rho)}, \quad (4)$$

with an intensity measurement given by [47]

$$I(\mathbf{q}, \rho, \Psi) = \left| \sum_{|\mathbf{h}| < h_{\text{max}}} \mathcal{S}_{\mathbf{q}, \mathbf{h}} \Psi(\mathbf{h}) e^{-2\pi i \mathbf{h} \cdot \rho} \right|^2. \quad (5)$$

In this work, we use the approximation that the wave function has finite support after propagating through the specimen potential. To use this approximation as a constraint in an inversion algorithm, we need to represent the \mathcal{S} matrix in real space:

$$I(\mathbf{q}, \rho, \Psi) = \left| \mathcal{F}_{\mathbf{r}} \left[\sum_{|\mathbf{h}| < h_{\text{max}}} \mathcal{S}_{\mathbf{r}, \mathbf{h}} \Psi(\mathbf{h}) e^{-2\pi i \mathbf{h} \cdot \rho} \right] \right|^2. \quad (6)$$

Here $\mathcal{S}_{\mathbf{r}, \mathbf{h}}$ is the \mathcal{S} matrix that maps Fourier-space input coefficients at wave vectors \mathbf{h} (we refer to these as the “beams” of the \mathcal{S} matrix) to real-space output coefficients at positions \mathbf{r} .

A previous experiment [48] used Eq. (5) and a series of defocus modulations to retrieve the phases of $\mathcal{S}_{\mathbf{q}, \mathbf{h}}$ for a set of \mathbf{h} vectors separately and then used symmetry relations of

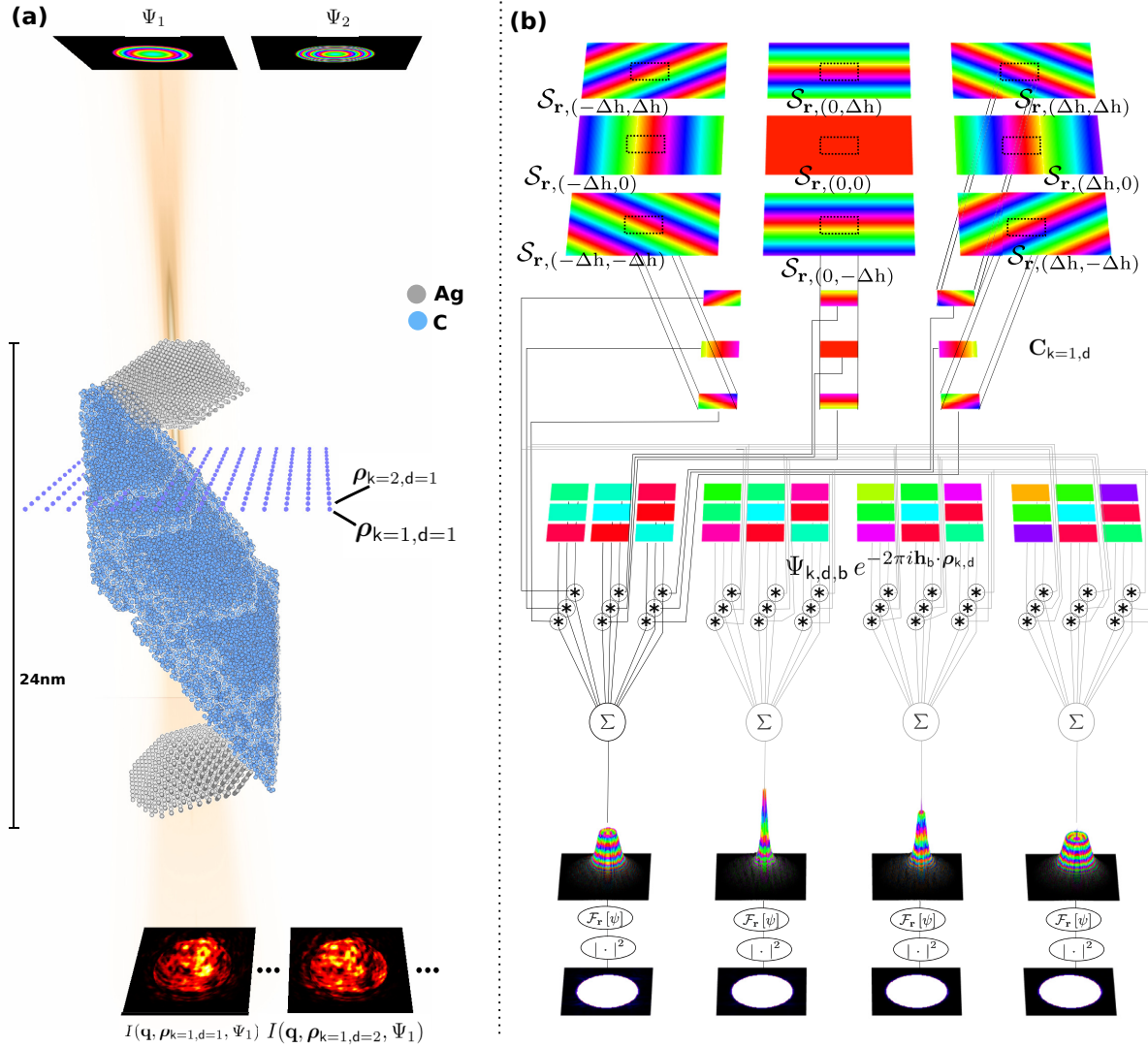


FIG. 1. Measurement scheme for S -matrix inversion. (a) A scanning diffraction series of a strongly scattering sample at atomic resolution, where the phases Ψ_d of the probe-forming aperture are varied after each scan, here by changing the defocus. (b) Computational graph of the S -matrix measurement operator for $D = 4$ different defocus aberrations. For each scan position, a patch with the size of the diffraction detector ($M_1 \times M_2$) is cropped out of each S -matrix beam. Then, each cropped beam is multiplied by the corresponding complex phase factor (indicated by the $*$ operation), depending on the phase and amplitude of the beam illumination aperture $\Psi_{d,b}$ and the scanning phase $e^{-2\pi i \mathbf{h}_b \cdot \rho_{k,d}}$ of the current position. Subsequently, all phase-shifted beams are coherently summed (the Σ operator) to form an exit wave. Then the exit wave is propagated to the far field (\mathcal{F}_r operation) and measured on the detector.

$\mathcal{S}_{q,h}$ to find the relative phases between the different S -matrix columns. Whereas that approach is valid only for crystalline samples, we use only self-consistency in the measured data and retrieve all amplitudes and phases of $\mathcal{S}_{r,h}$ simultaneously. We also introduce a real-space compactness constraint on the scattered probes produced by the scattering matrix, equivalent to the method of Fourier interpolating the S matrix [39]. We introduce the cropping operator

$$\mathbf{C}_{\rho,\Delta}(\mathbf{r}) = \begin{cases} 1 & \text{if } |r_x - \rho_x| \leq \Delta_x/2, |r_y - \rho_y| \leq \Delta_y/2, \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

a two-dimensional rectangular function of width Δ centered about each probe scan position ρ , which transforms Eq. (5) into

$$I(\mathbf{q}, \rho, \Psi) = \left| \mathcal{F}_r \left[\sum_{|\mathbf{h}| < h_{\max}} [\mathbf{C}_{\rho,\Delta}(\mathbf{r}) \mathcal{S}_{r,h}] \Psi(\mathbf{h}) e^{-2\pi i \mathbf{h} \cdot \rho} \right] \right|^2. \quad (8)$$

The fact that the cropping operator acts on all S -matrix beams equally leads to a self-consistent solution when measurements are taken with overlapping probe positions.

C. Phase retrieval of the \mathcal{S} matrix

We now describe an algorithm to retrieve all amplitudes and phases of $\mathcal{S}_{r,h}$ simultaneously, given a set of phase modulations $\{\chi_d(\mathbf{h})\}_{d=1,\dots,D}$ of the probe-forming aperture, using only self-consistency in the measured data. Let the detector be sampled with $M_1 \times M_2$ pixels. We perform a scan with K positions and D different probes and label a single position with k and a single defocus with d . Then the measured intensities have the dimension $\mathbf{I} \in \mathbb{R}^{KDM_1M_2}$. For ease of notation, we enumerate all B samples in $|\mathbf{h}| < h_{\max}$ with indices $b = 1, \dots, B$. The \mathcal{S} -matrix measurement operator maps the B beams of the \mathcal{S} matrix sampled on a discrete grid of $N_1 \times N_2$ pixels and the D probes to KD diffraction patterns of size $M_1 \times M_2$. $\mathcal{A} : \mathbb{C}^{B \times N_1 \times N_2} \times \mathbb{C}^{D \times M_1 \times M_2} \rightarrow \mathbb{C}^{KDM_1M_2}$. For better readability, we first define the measurement operator for position k and probe d , $\mathcal{A}_{k,d} : \mathbb{C}^{B \times N_1 \times N_2} \times \mathbb{C}^{M_1 \times M_2} \rightarrow \mathbb{C}^{M_1M_2}$:

$$\mathcal{A}_{k,d}(\mathcal{S}, \Psi_d) := \left[\mathcal{F}_r \left[\sum_{b=1}^B \Psi_{d,b} e^{-2\pi i \mathbf{h}_b \cdot \rho_{k,d}} [\mathbf{C}_{k,d} \mathcal{S}]_b \right] \right]^V, \quad (9)$$

where $[\cdot]^V$ is a vectorization from two dimensions to one dimension. We have also introduced the linear cropping operator $\mathbf{C}_{k,d} := \mathbf{C}_{\rho_{k,d}} : \mathbb{C}^{B \times N_1 \times N_2} \rightarrow \mathbb{C}^{B \times M_1 \times M_2}$, which extracts a real-space patch of size $M_1 \times M_2$ from each beam of a given \mathcal{S} matrix at the position with index k for the phase modulation d . The measurement operator for the full experiment is just the operators for each probe and position stacked on top of each other: $\mathcal{A}(\mathcal{S}, \Psi) = [\mathcal{A}_{1,1}(\mathcal{S}, \Psi_1), \mathcal{A}_{2,1}(\mathcal{S}, \Psi_1), \dots, \mathcal{A}_{K,D}(\mathcal{S}, \Psi_D)]^T$. We can then write the forward model for the measured intensities of a series of D scanning diffraction experiments taken with different probes as

$$\mathbf{y} = |\mathcal{A}(\mathcal{S}, \Psi)|^2. \quad (10)$$

Given this forward model and a set of intensity measurements \mathbf{I} , we can formulate the phase retrieval problem for blind \mathcal{S} -matrix inversion as

$$\begin{aligned} \text{Find } \mathcal{S} \in \mathbb{C}^{B \times N_1 \times N_2} \quad \text{and} \quad \Psi \in \mathbb{C}^{D \times B} \\ \text{Subject to } |\mathcal{A}(\mathcal{S}, \Psi)|^2 = \mathbf{I}. \end{aligned}$$

If the wave functions Ψ are known, the problem of finding \mathcal{S} from a set of measurements \mathbf{I} is a classical phase retrieval problem. There is a rich history of algorithmic developments to solve the phase retrieval problem. Historically, the first were algorithms based on alternating projections onto non-convex constraint sets [58–60]. Since these algorithms lack theoretical convergence guarantees, more recently, convex relaxations were developed [61,62] which provide a convergence guarantee but use a prohibitive amount of memory. More recently, Bayesian accelerated gradient methods [63] and methods based on the alternating direction method of multipliers (ADMM) [64] have become popular.

Since the wave functions Ψ_d are usually not known precisely in advance, the problem turns into multiobjective

optimization. Additionally, in the presence of noise, it is beneficial to the reconstruction quality to include a model of the detector-specific noise characteristics in the optimization. For most advanced detectors in x-ray and electron microscopy the noise statistics follow a Poisson distribution: $\mathbf{I} \sim \text{Poisson}(\mathbf{y})$. Therefore, usually, either the negative Poisson log likelihood or a first-order approximation to it, the least-squares loss of the measured amplitudes [65], is minimized in phase-retrieval problems dealing with measurement noise. These are given by

$$\mathcal{D}_{\text{Pois}}(\mathbf{y}, \mathbf{I}) := \langle \mathbf{y} - \mathbf{I} \ln(\mathbf{y}) | \mathbf{1} \rangle, \quad (11)$$

$$\mathcal{D}_{\text{Amp}}(\mathbf{y}, \mathbf{I}) := \|\sqrt{\mathbf{y}} - \sqrt{\mathbf{I}}\|_2, \quad (12)$$

where $\|\cdot\|_2$ is the l^2 -norm, $\langle \cdot | \cdot \rangle$ is the L^2 inner product of two vectors, and \mathbf{y} are far-field intensities of the current model. We use block coordinate descent to solve the joint optimization problem of \mathcal{S} and Ψ . The loss function of the \mathcal{S} -matrix retrieval problem is

$$\mathcal{L}_{\square}(\mathcal{S}, \Psi) = \mathcal{D}_{\square}(|\mathcal{A}(\mathcal{S}, \Psi)|^2), \quad (13)$$

where \square is a placeholder for the chosen loss function. We seek to solve for \mathcal{S} and Ψ such that $\mathcal{L}(\mathcal{S}, \Psi)$ is minimized:

$$(\mathcal{S}^*, \Psi^*) = \arg \min_{\mathcal{S}, \Psi} \mathcal{L}_{\square}(\mathcal{S}, \Psi) \quad (14)$$

Block coordinate descent solves the joint problem by alternating gradient updates for one variable with the other variable fixed: (1) $\mathcal{S}^{l+1} = \arg \min_{\mathcal{S}} \mathcal{L}_{\square}(\mathcal{S}^l, \Psi^l)$ and (2) $\Psi^{l+1} = \arg \min_{\Psi} \mathcal{L}_{\square}(\mathcal{S}^{l+1}, \Psi^l)$.

D. Gradients with respect to Ψ and \mathcal{S}

The gradients with respect to Ψ and \mathcal{S} both involve the adjoint of the measurement operator \mathcal{A} , which for a single measurement is given by $\mathcal{A}_{k,d}^{\dagger} : \mathbb{C}^{M_1M_2} \rightarrow \mathbb{C}^{B \times N_1 \times N_2}$,

$$\mathcal{A}_{k,d}^{\dagger}(\mathbf{z}) = \mathcal{C}_{k,d}^T [\Psi_{d,b}^* e^{2\pi i \mathbf{h}_b \cdot \rho_{k,d}} \mathcal{F}_q^{\dagger}[\mathbf{z}_{k,d}]] \quad (15)$$

for a fixed Ψ , and $\mathcal{A}_{k,d}^{\Psi_{d,b} \dagger} : \mathbb{C}^{M_1M_2} \rightarrow \mathbb{C}^{M_1 \times M_2}$,

$$\begin{aligned} \mathcal{A}_{k,d}^{\Psi_{d,b} \dagger}(\mathbf{z}) &= \frac{1}{M_1M_2} \\ &\times \sum_{m_1}^{M_1} \sum_{m_2}^{M_2} \left[\sum_{k=1}^K [\mathbf{C}_{k,d} \mathcal{S}]_b^* e^{2\pi i \mathbf{h}_b \cdot \rho_{k,d}} \mathcal{F}_q^{\dagger}[\mathbf{z}_{k,d}] \right]_{m_1, m_2} \end{aligned} \quad (16)$$

for a fixed \mathcal{S}_b . We solve the subproblems with respect to Ψ and \mathcal{S} with gradient descent:

$$\Psi^{l+1} = \Psi^l + \gamma_1 \frac{\partial \mathcal{L}_{\square}^{\Psi}}{\partial \Psi}, \quad (17)$$

$$\mathcal{S}^{l+1} = \mathcal{S}^l + \gamma_2 \frac{\partial \mathcal{L}_{\square}^{\mathcal{S}}}{\partial \mathcal{S}}, \quad (18)$$

where $\gamma_1, \gamma_2 \in \mathbb{R}$ are gradient descent step sizes. We found that one gradient step per iteration is usually enough for fast convergence. The gradient is given by

$$\frac{\partial \mathcal{L}_{\text{Pois}}^{\Psi}}{\partial \Psi_{d,b}} = \mathcal{A}_{k,d}^{\Psi_{d,b}^\dagger} \left(\mathcal{A}_{k,d}(\mathcal{S}, \Psi_d) \cdot \left(\mathbf{1} - \frac{\mathbf{I}}{\mathbf{y}} \right) \right), \quad (19)$$

$$\frac{\partial \mathcal{L}_{\text{Pois}}^{\mathcal{S}}}{\partial \mathcal{S}_b} = \sum_{k=1}^K \sum_{d=1}^D \mathcal{A}_{k,d}^{\mathcal{S}_b^\dagger} \left(\mathcal{A}_{k,d}(\mathcal{S}, \Psi_d) \cdot \left(\mathbf{1} - \frac{\mathbf{I}}{\mathbf{y}} \right) \right), \quad (20)$$

$$\frac{\partial \mathcal{L}_{\text{Amp}}^{\Psi}}{\partial \Psi_{d,b}} = \mathcal{A}_{k,d}^{\Psi_{d,b}^\dagger} \left(\mathcal{A}_{k,d}(\mathcal{S}, \Psi_d) \cdot \left(\mathbf{1} - \frac{\sqrt{\mathbf{I}}}{\sqrt{\mathbf{y}}} \right) \right), \quad (21)$$

$$\frac{\partial \mathcal{L}_{\text{Amp}}^{\mathcal{S}}}{\partial \mathcal{S}_b} = \sum_{k=1}^K \sum_{d=1}^D \mathcal{A}_{k,d}^{\mathcal{S}_b^\dagger} \left(\mathcal{A}_{k,d}(\mathcal{S}, \Psi_d) \cdot \left(\mathbf{1} - \frac{\sqrt{\mathbf{I}}}{\sqrt{\mathbf{y}}} \right) \right). \quad (22)$$

See the Supplemental Material [66] for a detailed derivation.

E. Probe modeling with Zernike polynomials

We found that the stability of probe refinement was improved by modeling the phase modulations of the probes in the low-dimensional space of Zernike polynomials up to fourth order, using the formalism in [67]: $\Psi_d(\mathbf{q}) = A(\mathbf{q})e^{-i\chi_d(\mathbf{q})}$, where $A(\mathbf{q})$ is a circular top-hat function $A(\mathbf{q}) = \text{Circ}(\mathbf{q})$ and $\chi_d(\mathbf{q})$ is the aberration function,

$$\chi_d(\alpha, \phi) = \frac{2\pi}{\lambda} \sum_{m,n} \frac{C_{n,m} \alpha^{n+1} \cos[m(\phi - \phi_{n,m})]}{n+1}, \quad (23)$$

where α is the numerical aperture and ϕ is the azimuthal angle. Here $C_{1,0}$ is the defocus, $C_{1,2}$ is the astigmatism aberration coefficient, $C_{2,1}$ are the coma aberration coefficients, $C_{2,3}$ is the threefold astigmatism aberration coefficient, and $C_{3,0}$ is the spherical aberration coefficient. At each iteration, we compute the derivatives $\frac{\partial \Psi}{\partial C_{n,m}}$ with the automatic gradient calculation of the PYTORCH library [68].

F. Ambiguities and initialization

Since the forward model in Eq. (10) contains the product of two complex numbers, $\Psi_{d,b}$ and $[\mathcal{C}_{k,d}\mathcal{S}]_b$, any combination of phases that cancel each other out could be added to these two terms without influencing the measured intensities. Therefore, a meaningful initialization is key to a successful reconstruction. \mathcal{S} -matrix retrieval shares with other phase retrieval methods the property that the focal plane of the reconstruction can be chosen freely. The initialization of the \mathcal{S} matrix with plane waves $\mathcal{S}_b^0 \leftarrow e^{i\mathbf{h}_b \cdot \mathbf{r}}$ corresponds to a choice of the zero-focus position at the top of the sample and worked well for all our tests, but a different focal plane can be selected by shifting the relative defocus of the probes

$\Psi^0 \leftarrow \Psi^0 e^{-i\pi\lambda h^2 z_0}$. The full block coordinate descent algorithm is then given as follows:

Algorithm 1: Joint \mathcal{S} matrix and probe retrieval via block coordinate descent.

Input:

measured intensities $\mathbf{I} \in \mathbb{R}^{K \times D \times M_1 \times M_2}$

scan positions $\boldsymbol{\rho} \in \mathbb{R}^{K \times D \times 2}$

initial aberration coefficients $C_{n,m} \in \mathbb{R}^{D \times 12}$

step sizes $\gamma_1, \gamma_2 \in \mathbb{R}$

Initialize:

initialize Fourier space probe phases $\chi^0 \in \mathbb{C}^{D \times B}$ according to Eq. (23)

set $(N_1, N_2) = \lceil \frac{\max(\mathbf{r}_2) + M}{M} \rceil \cdot M$ such that the plane waves $e^{i\mathbf{h} \cdot \mathbf{r}}$

have periodic boundary conditions

calculate $\mathbf{I}^{\text{mean}} = \frac{1}{K} \sum_{k=1}^K \mathbf{I}_k$ and

$a_{\max} = \max(\|\mathbf{I}_k\|_1 \forall k = \{1, \dots, K\})$

$\Psi^0 \leftarrow \frac{a_{\max}}{\sqrt{\|\mathbf{I}^{\text{mean}}\|_1}} \mathbf{I}^{\text{mean}} e^{i\chi^0}$

$\mathcal{S}_b^0 \leftarrow e^{i\mathbf{h}_b \cdot \mathbf{r}}, \mathcal{S} \in \mathbb{C}^{B \times N_1 \times N_2}$

Run: for $l = 0$ to L do

$\mathcal{S}^{l+1} \leftarrow \mathcal{S}^l + \gamma_1 \frac{\partial \mathcal{L}_{\square}}{\partial \mathcal{S}}(\mathcal{S}^l, \Psi^{l+1});$

$\frac{\partial \mathcal{L}_{\square}}{\partial \Psi}^{l+1} \leftarrow \Psi^l + \gamma_1 \frac{\partial \mathcal{L}_{\square}}{\partial \Psi}(\mathcal{S}^l, \Psi^l);$

$C_{n,m}^{l+1} \leftarrow C_{n,m}^l + \gamma_2 \frac{\partial \mathcal{L}_{\square}}{\partial C_{n,m}} \frac{\partial \Psi}{\partial C_{n,m}}(\mathcal{S}^{l+1}, \Psi^l);$

$\Psi_d^{l+1} \leftarrow A(\mathbf{q})e^{-i\chi_d(\mathbf{q})};$

Output: $\mathcal{S}^* = \mathcal{S}^L$

We have also implemented the ADMM algorithm to solve the joint optimization problem of \mathcal{S} and Ψ and found roughly similar results in the simulation. ADMM may be better suited in the future to implement additional constraints and regularizers; therefore, we give a description in Appendix D for future reference.

III. SIMULATED \mathcal{S} -MATRIX PHASE RETRIEVAL

In this section, we use forward simulations to validate our \mathcal{S} -matrix phase retrieval algorithm. We also examine the algorithm dependence on the sampling density and calibration.

A. Sampling and calibration dependence

To demonstrate that our algorithm can reconstruct \mathcal{S} matrices of realistic samples, we simulate a four-dimensional (4D) STEM focal series of the sample shown in Fig. 1(a), as it may appear in a tomography experiment. The sample contains two decahedral Ag nanoparticles with a 3.3-nm diameter, placed on the top and bottom sides of an amorphous carbon substrate, tilted by 67° , giving it an axial extent of 24 nm. The probe convergence angle is chosen to be 26 mrad, and the electron energy is 300 kV, resulting in a depth of focus (DOF) of 5.8 nm and a sample depth of 4.1 times the DOF. The detector was set to record the diffraction signal up to 40 mrad, resulting in a sampling grid with steps of 25 pm. The field of view was scanned with 129×129 positions on a 2D

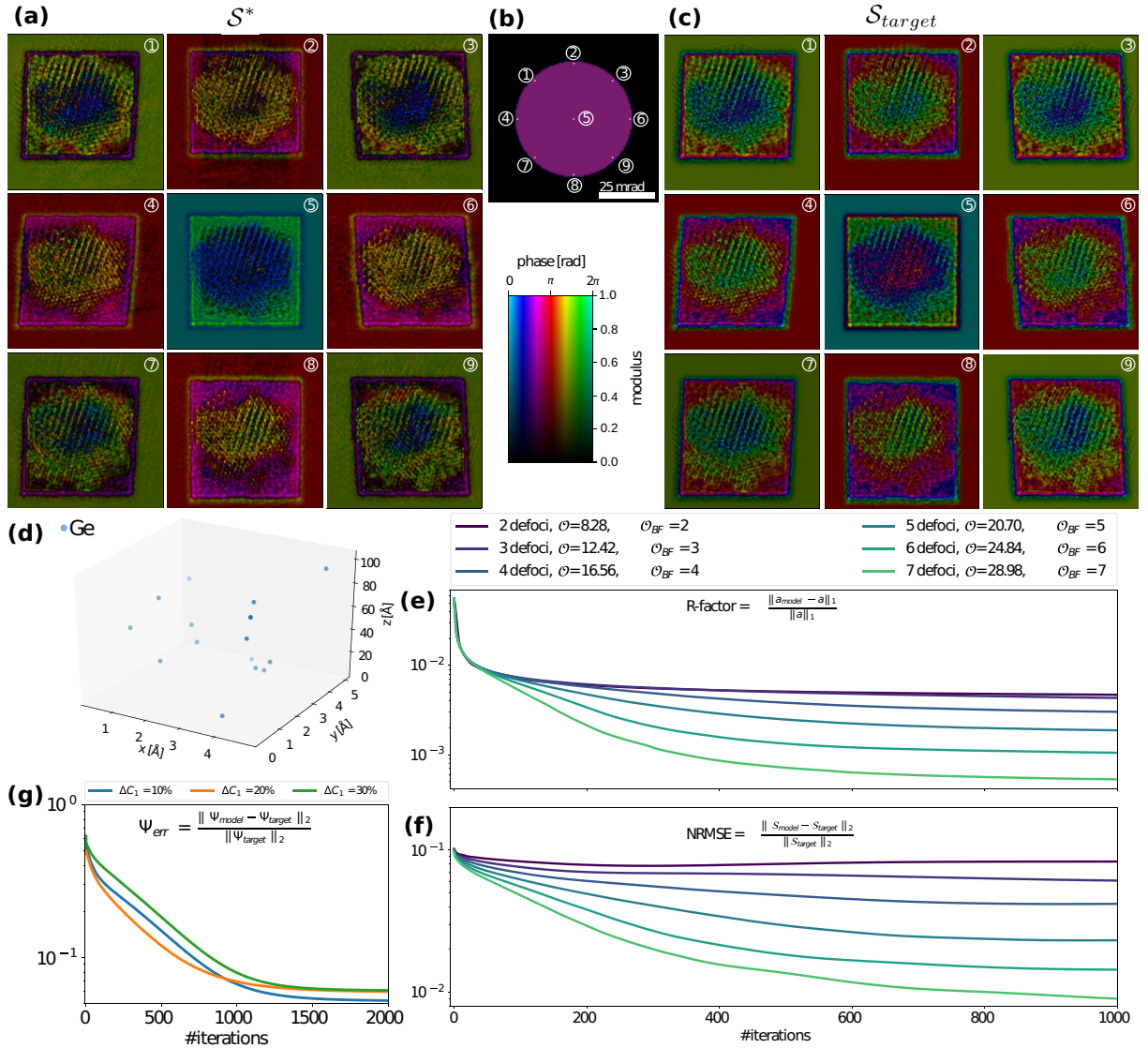


FIG. 2. (a) Simulated experiment with (b) the probe-forming aperture used for simulating the experiment shown in Fig. 1(a). Selected beams numbered in (b) are shown from the reconstructed S matrix in (a) and the ground-truth S matrix in (c). The skew effect of the exit waves in different beams comes from the three-dimensional structure and is a parallax effect of the different propagation directions of the beams. (d) Test sample of randomly distributed germanium atoms. (e) R factor vs number of iterations for different numbers of defoci and oversampling rates used in the simulations. (f) Normalized rms error of the model S matrix vs the number of iterations for different numbers of defoci and oversampling rates. (g) Mean probe error vs the number of iterations for defocus miscalibration levels of 10%, 20%, and 30% of the defocus step and random higher-order aberrations.

grid with the half-period resolution. The reconstruction shown in Fig. 2 used six defoci with a step of 4.6 nm, with the first defocus at the top of the sample. The detector size was set to 128×128 pixels, yielding an angular resolution of 0.31 mrad and S -matrix dimensions of $S \in \mathbb{C}^{5973 \times 256 \times 256}$.

We ran Algorithm 1 for 500 iterations, utilizing 48 NVIDIA V-100 graphics processing units (GPUs). After 200 min, the reconstruction converged to a normalized rms error (NRMSE) of 4% and an R factor of 0.1%. Nine selected S -matrix beams from the reconstruction are shown in

Fig. 2(a), and the ground-truth S matrix is shown in Fig. 2(c). To investigate the convergence properties under varying numbers of measurements and calibration errors, we used a smaller test sample, consisting of 16 randomly distributed germanium atoms in a volume of $5 \times 5 \times 100 \text{ \AA}^3$, shown in Fig. 2(d). The convergence angle for the following tests was chosen to be 30 mrad, with a detector spanning 60 mrad, and the diffraction patterns were sampled on a 20×20 pixel detector, yielding S -matrix dimensions of $S \in \mathbb{C}^{177 \times 60 \times 60}$; the defocus step was chosen to be 10 nm.

For the following investigations we fix the scan step to Nyquist sampling. First, we investigate the convergence behavior with respect to the number of measured defoci. Figures 2(e) and 2(f) show the R factor and the NRMSE as a function of iterations and the number of defoci measured. We define the oversampling factor as

$$\mathcal{O} = \frac{\text{number of nonzero measurements}}{\text{number of variables in } \mathcal{S} \text{ matrix}} \quad (24)$$

and the bright-field oversampling factor as

$$\mathcal{O}_{BF} = \frac{\text{number of nonzero measurements in bright field}}{\text{number of variables in } \mathcal{S} \text{ matrix}}. \quad (25)$$

One can see that for two defocus measurements, the NRMSE diverges slowly, and for three measurements the NRMSE does not converge monotonically with the R factor. While the oversampling factor \mathcal{O} lies above the number 4 typically needed for successful phase retrieval, the number of phase modulations that each beam receives \mathcal{O}_{BF} is below the threshold. In this case, a more heterogeneous sample than the crystalline objects considered in previous work, the reconstruction does not stably converge. This could be due to the small defocus steps used and will be investigated in the future.

We also investigate the dependence of the probe refinement on the level of defocus miscalibration and residual uncorrected probe aberrations. Figure 2(h) shows the mean errors of 30 reconstructions performed with defocus errors ΔC_1 drawn from a normal distribution with a standard deviation of 10%, 20%, and 30% of the defocus step, axial coma with a standard deviation of 100 nm, threefold astigmatism with a standard deviation of 20 nm, spherical aberration with a standard deviation of 4 μm , and star aberration with a standard deviation of 4 μm . Although convergence takes roughly twice as many iterations of \mathcal{S} -matrix reconstruction with miscalibrated aberrations, for all miscalibration values a probe reconstruction error of less than 10% was achieved.

The high number of possible experimental configurations for this type of experiment leaves some areas unexplored by our simulated reconstructions: we did not explore the dependence on experimental noise and how an available dose budget should be distributed over a number of scans with different aberrations, but our experimental demonstration in the next section with noise in the Poisson regime demonstrates that \mathcal{S} -matrix retrieval is feasible with roughly 1700 counts per probe position, and an even distribution of the dose over several scans yields good reconstructions. We also did not explore aberrations other than defocus and how the amount of crystallinity of a sample influences the convergence.

IV. EXPERIMENTAL DEMONSTRATION

As an experimental proof of principle for the \mathcal{S} -matrix reconstruction method, we used the TEAM 0.5 microscope at the National Center for Electron Microscopy to perform 4D-STEM experiments [69]. The TEAM 0.5 microscope is a double aberration-corrected Thermo Fisher/FEI Titan microscope, equipped with the 4D Camera, a fast-framing direct electron detector that operates at 87-kHz frame rate, which

leads to a dwell time of 11 μs [70]. We used an acceleration voltage of 80 kV and a numerical aperture (NA) of 25 mrad and recorded the 4D-STEM data sets with multiple STEM probe defocus values. The sample consisted of multiwall carbon nanotubes filled tantalum and tellurium, arranged in a complex core-shell geometry, with sample preparation similar to previous encapsulated transition metal chalcogenides. Figure 3(a) shows a high-angle annular dark field (HAADF)-STEM image of a single filled nanotube. HAADF-STEM imaging is not sensitive to weakly scattering elements such as carbon [71], and therefore, only the Ta and Te structures are visible. An approximate structural model cross section is given in the inset in Fig. 3(a), showing both core and shell structures that consist of hexagonal TaTe_2 [72–74].

For the \mathcal{S} -matrix reconstructions, we chose a region of the sample consisting of two overlapping bundles of nanotubes, spaced roughly 10 nm apart from each other along the beam direction. We recorded four 4D-STEM data sets with 512×512 probe positions, a raster step of 31 pm, defocus steps of 10 nm, and a fluence of $3.05 \times 10^4 \text{ e}^-/\text{\AA}^2$ per data set, estimated from the counts on the 4D Camera. Simultaneously collected HAADF-STEM images are shown in Fig. 1 in the Supplemental Material [66]. The $512 \times 512 \times 576 \times 576$ raw 4D data cube was counted and converted to a sparse electron event format with the STEMpy open-source toolkit [75]. Mean diffraction patterns and intensity histograms are shown in Fig. 2 in the Supplemental Material [66], and example diffraction patterns are shown in Fig. 4 in the Supplemental Material [66]. The sparse data were then cropped and centered around the bright field disk and binned by a factor of 6 in reciprocal space, yielding an \mathcal{S} matrix with 1371 beams and spatial dimensions of $15.8 \times 15.8 \text{ nm}^2$. For each data set we performed a single-sideband ptychography reconstruction [76] focused at the thick nanotube and aligned the single-scan reconstructions to each other with subpixel rigid cross correlation. The averaged and aligned single-sideband reconstructions are shown in Fig. 3 in the Supplemental Material [66]. The \mathcal{S} -matrix reconstruction code was adapted to operate directly on sparse electron event coordinates instead of the dense 4D data set, which resulted in a 10-fold reduction in memory usage for the experimental data. We obtain a solution to the joint \mathcal{S} matrix and probe retrieval problem with an R factor of 0.159 after 100 iterations of Algorithm 1 with Poisson loss function, with step sizes of $\gamma_1 = 10^{-4}$ and $\gamma_2 = 10^2$. Figures 3(b) and 3(c) show phase-contrast images focused on the thick nanotube, generated from the reconstructed \mathcal{S} matrix by performing a coherent sum of all beams, with the incident beam tilt removed from the output beams, $\sum_{\mathbf{h}} \mathcal{F}_{\mathbf{q}}^\dagger[\mathcal{S}_{\mathbf{q}-\mathbf{h},\mathbf{h}}]$. In both of the \mathcal{S} -matrix reconstructions, we see substantially higher contrast than in the HAADF-STEM images due to the phase-contrast nature of the imaging. Additionally, this method is also highly sensitive to the weakly scattering carbon portions of the sample, including the multiwalled nanotubes wrapped around the Ta and Te structures.

Figures 3(d)–3(f) show average unit cells taken from the three images shown in Figs. 3(a)–3(c). The wires all show nearly ideal periodicity with a length of 17.1 \AA . The ratio of the periodicity measured in the core structure to that of the shell structure is almost exactly 2:5, corresponding to

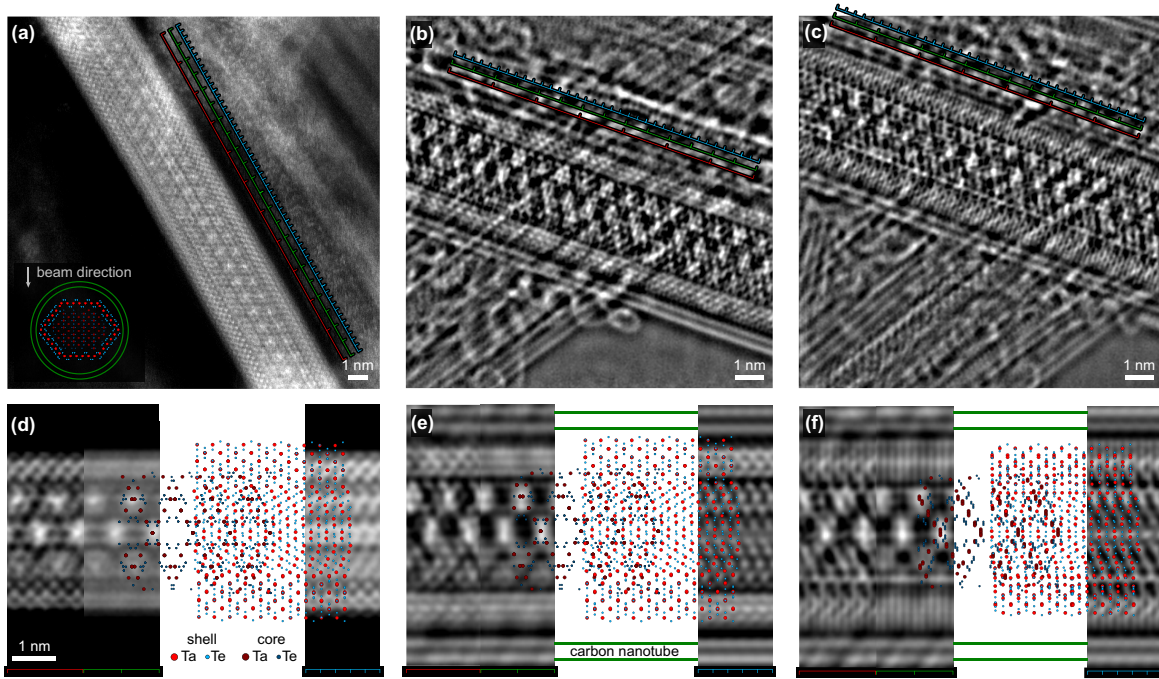


FIG. 3. (a) HAADF-STEM image of a carbon nanotube containing a core-shell Ta and Te structure. An approximate structural model is shown in the inset, with the beam direction indicated. (b) and (c) Phase images of overlapping nanotubes, recorded using the S -matrix reconstruction method described in this work, at two tilt angles separated by $\approx 5^\circ$. (d)–(f) Three average unit cells, corresponding to the images shown in (a)–(c), respectively. Three unit cells are shown in each panel. The leftmost one has a periodicity of 17.1 Å and corresponds to the overall structure from the experiment. The center and rightmost panels are simulated from the proposed core-shell structure and match the overall periodicity of the experimental structure. The center has a periodicity of 8.55 Å to highlight the core structure, and the rightmost unit cell has a periodicity of 3.42 Å to highlight the shell structure. A potential structural model of core-shell TaTe₂ has been overlaid over the images.

periodicities of 3.42 and 8.55 Å, respectively. The average unit cells for these structure lengths are also shown in Fig. 3. We have overlaid some possible core-shell structures onto the average unit cells, corresponding to different orientations of TaTe₂. The shell structures are an excellent match to a hexagonal shape, where the TaTe₂ edges are connected with slight offsets. The core structures are more ambiguous. They could correspond to a rotated TaTe₂ structure where the layer stacking vector is nearly perpendicular to the beam direction (as in the model shown here) or a high-angle lattice rotation leading to a moiré structure [77] or some even more complex reconstruction such as a 2D quasicrystalline arrangement or its periodic approximant [78,79]. By assuming that the core structure consists of overlapping layers of TaTe₂, we are able to roughly match the structures with 8.55-Å periodicity. Because this is a proof-of-principle experiment, we have not attempted to solve the structure exactly or precisely refine the atomic coordinates.

V. CONCLUSION AND OUTLOOK

We have introduced a method for S -matrix retrieval that converges for samples which span four or more depths of focus and numerical apertures which are experimentally accessible and can refine aberration miscalibrations of up to 30%. We have demonstrated this algorithm experimentally

with STEM at atomic resolution with a bundle of multiwall carbon nanotubes filled with a TaTe core-shell structure and synthesized two-dimensional phase images from the scattering matrix, demonstrating its use in imaging light atoms in thick, heterogeneous nanostructures. In future work, we will compare S -matrix inversion to mixed-state ptychography and multislice ptychography since both offer alternative methods for reconstructing phase-contrast images from scanning diffraction data and move beyond the simple model of single-mode ptychography.

Our algorithm and experimental configuration can be readily applied at scanning x-ray diffraction microscopes as well. The main limitation of the implementation at synchrotron sources is the relatively strict requirement for the scan step size and the relatively slow frame rate compared to STEM instruments. The advent of high-NA diffractive x-ray optics with NAs of 7.5 mrad experimentally demonstrated [80] and NAs of 16 mrad projected to be feasible [81] and efforts to increase spatial resolution of x-ray microscopes to 1 nm will require new algorithms to move beyond the thickness limits of current methods.

The S -matrix-retrieval methods developed here could be used for a number of advancements in imaging through and with strongly scattering materials in x-ray and electron microscopy. In combination with adaptive electron optics [82],

selective focusing through crystalline materials may become possible in a vein similar to light optical experiments [83]. The retrieved \mathcal{S} matrix can be used for optical depth sectioning, which proved robust against multiple scattering in a proof-of-principle experiment [55].

\mathcal{S} -matrix retrieval can also form the basis of inverse multislice algorithms for phase-contrast tomography in scanning diffraction microscopes. The fine angular decomposition of exit waves in the \mathcal{S} matrix may be useful for *ab initio* angular and transverse alignment of different tilt angles for phase-contrast tomography. The retrieved \mathcal{S} matrix is identical to the result of performing many tilt-focal series reconstructions with plane-wave illumination in parallel [84].

To allow optimal image quality, future refinements of the algorithm could include experimental uncertainties like position errors and modeling of nuisance parameters like partial spatial and temporal coherence, similar to their treatment in ptychographic reconstruction algorithms [85–89].

ACKNOWLEDGMENTS

We thank T. Pelz and N. and P. Hosemann for support during the COVID-19 pandemic. Without them, this research would not have been possible. P.M.P. acknowledges financial support from STROBE: A National Science Foundation Science and Technology Center under Grant No. DMR 1548924. H.G.B. and J.C. acknowledge support from the Presidential Early Career Award for Scientists and Engineers (PECASE) through the U.S. Department of Energy. C.O. acknowledges support from the U.S. Department of Energy Early Career Research Award program. S.S. and A.Z. acknowledge support from the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, U.S. Department of Energy, under Contract No. DEAC02-05-CH11231 within the sp2-Bonded Materials Program (KC-2207), which provided for high-resolution TEM imaging, and within the van der Waals Heterostructures program (KCWF16), which provided for synthesis of the encapsulated material, and the National Science Foundation under Grant No. DMR-1807233, which provided for preliminary conventional TEM imaging. S.D.F. acknowledges support from the Discovery Projects funding scheme of the Australian Research Council (Project No. FT190100619). C.O. acknowledges support from the U.S. Department of Energy Early Career Research Program. Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. We are grateful for early access to the computer cluster of the NERSC Exascale Science Applications Program.

APPENDIX A: DERIVATION OF THE ADJOINT \mathcal{S} -MATRIX MEASUREMENT OPERATOR

We derive the adjoint operators $\mathcal{A}_{k,d}^{S_b^\dagger}(\mathbf{z})$ and $\mathcal{A}_{k,d}^{\Psi_{a,b}^\dagger}(\mathbf{z})$ with matrix algebra. In matrix notation, the \mathcal{S} -matrix forward model to generate $\mathbf{I} \in \mathbb{R}^{\text{KDM}}$ with $M = M_1 M_2$ from $\mathcal{S} \in \mathbb{C}^{\text{BN}}$

with $N = N_1 N_2$ can be written as $\mathcal{A}_S : \mathbb{C}^{\text{BN}} \rightarrow \mathbb{C}^{\text{KDM}}$

$$\mathbf{I} = |\mathbf{F}\Sigma\mathcal{C}\mathcal{S}|^2 = |\mathcal{A}_S\mathcal{S}|^2, \quad (\text{A1})$$

with $\mathbf{F} \in \mathbb{C}^{\text{KDM} \times \text{KDM}}$ a block-diagonal matrix representing a batched Fourier transform acting on KD exit waves, $\mathbf{C} \in \mathbb{R}^{\text{KDBM} \times \text{BN}}$ the cropping matrix that extracts KD patches centered at the scanning positions out of the \mathbf{B} beams of the \mathcal{S} -matrix, and $\Sigma \in \mathbb{C}^{\text{KDM} \times \text{KDBM}}$ the coherent summation operator over all beams. Written out in block matrices with diagonal entries, Σ is shown in Fig. 4. The adjoint (hermitian transpose) operator \mathcal{A}_S^\dagger is then

$$\mathcal{A}_S^\dagger = \mathbf{C}^T \Sigma^\dagger \mathbf{F}^\dagger, \quad (\text{A2})$$

which, written out for a single diffraction pattern with defocus index \mathbf{d} and position index \mathbf{k} is

$$\mathcal{A}_{k,d}^{S_b^\dagger}(\mathbf{z}) = \mathbf{C}_{k,d}^T [\Psi_{d,b}^* e^{2\pi i \mathbf{h}_b \cdot \rho_{k,d}} \mathcal{F}_q^\dagger[\mathbf{z}_{k,d}]] \quad (\text{A3})$$

In the same vein, the forward model to generate $\mathbf{I} \in \mathbb{R}^{\text{KDM}}$ from $\Psi \in \mathbb{C}^{\text{DB}}$ can be written as $\mathcal{A}_\Psi : \mathbb{C}^{\text{DB}} \rightarrow \mathbb{C}^{\text{KDM}}$

$$\mathbf{I} = |\mathbf{F}\Sigma\Psi|^2 = |\mathcal{A}_\Psi\Psi|^2 \quad (\text{A4})$$

The adjoint operator \mathcal{A}_Ψ^\dagger is then

$$\mathcal{A}_\Psi^\dagger = \Sigma^\dagger \mathbf{F}^\dagger, \quad (\text{A5})$$

which, written out for a single diffraction pattern with defocus index \mathbf{d} and position index \mathbf{k} is

$$\mathcal{A}_{k,d}^{\Psi_{a,b}^\dagger}(\mathbf{z}) = \frac{1}{M_1 M_2} \times \sum_{m_1}^{M_1} \sum_{m_2}^{M_2} \left[\sum_{k=1}^K [\mathbf{C}_{k,d} \mathcal{S}]_b^* e^{2\pi i \mathbf{h}_b \cdot \rho_{k,d}} \mathcal{F}_q^\dagger[\mathbf{z}_{k,d}] \right]_{m_1, m_2} \quad (\text{A6})$$

APPENDIX B: COMPLEXITY ANALYSIS

Both the forward calculation and the backward calculation have the following computational complexity: per the diffraction pattern the forward pass has a complexity of $O(M_1 M_2 B + M_1 M_2 \log_2(M_1 M_2))$, where the factor $O(M_1 M_2 \log_2(M_1 M_2))$ comes from the fast Fourier transform operation. The forward and backward calculations on the full data set then have a complexity of $O(KD(M_1 M_2 B + M_1 M_2 \log_2(M_1 M_2)))$. Since the number of beams scales quadratically with the size of the detector, the overall complexity scales with $O(KD(M^4 + 2M^2 \log_2(M)))$ for a square detector of size M . While this might seem intractable for currently available large detectors, it is offset by the fact that KDM^2 of these computations are embarrassingly parallel batched complex matrix multiplications and can be carried out very efficiently on commonly available hardware accelerators.

$$\begin{array}{c}
\mathbf{\Sigma}[:,1:\mathbf{BM}] := \\
\begin{array}{ccccc}
\mathbf{b}=1 & \mathbf{b}=2 & & \mathbf{b}=\mathbf{B} & \\
\left[\begin{array}{c} \text{diag}(\Psi_{1,1}e^{-2\pi i \mathbf{h}_1 \cdot \boldsymbol{\rho}_{1,1}}) \\ \vdots \\ \text{diag}(\Psi_{1,1}e^{-2\pi i \mathbf{h}_1 \cdot \boldsymbol{\rho}_{1,1}}) \end{array} \right] \left[\begin{array}{c} \text{diag}(\Psi_{1,2}e^{-2\pi i \mathbf{h}_2 \cdot \boldsymbol{\rho}_{1,1}}) \\ \vdots \\ \text{diag}(\Psi_{1,2}e^{-2\pi i \mathbf{h}_2 \cdot \boldsymbol{\rho}_{1,1}}) \end{array} \right] & \cdots & \left[\begin{array}{c} \text{diag}(\Psi_{1,\mathbf{B}}e^{-2\pi i \mathbf{h}_{\mathbf{B}} \cdot \boldsymbol{\rho}_{1,1}}) \\ \vdots \\ \text{diag}(\Psi_{1,\mathbf{B}}e^{-2\pi i \mathbf{h}_{\mathbf{B}} \cdot \boldsymbol{\rho}_{1,1}}) \end{array} \right] & \begin{array}{l} \mathbf{k}=1 \text{ d}=1 \\ \mathbf{k}=2 \text{ d}=1 \end{array} \\
\vdots & \vdots & \ddots & \vdots & \\
\left[\begin{array}{c} \text{diag}(\Psi_{2,1}e^{-2\pi i \mathbf{h}_1 \cdot \boldsymbol{\rho}_{1,2}}) \\ \vdots \\ \text{diag}(\Psi_{2,1}e^{-2\pi i \mathbf{h}_1 \cdot \boldsymbol{\rho}_{2,2}}) \end{array} \right] \left[\begin{array}{c} \text{diag}(\Psi_{2,2}e^{-2\pi i \mathbf{h}_2 \cdot \boldsymbol{\rho}_{1,2}}) \\ \vdots \\ \text{diag}(\Psi_{2,2}e^{-2\pi i \mathbf{h}_2 \cdot \boldsymbol{\rho}_{2,2}}) \end{array} \right] & \cdots & \left[\begin{array}{c} \text{diag}(\Psi_{2,\mathbf{B}}e^{-2\pi i \mathbf{h}_{\mathbf{B}} \cdot \boldsymbol{\rho}_{1,2}}) \\ \vdots \\ \text{diag}(\Psi_{2,\mathbf{B}}e^{-2\pi i \mathbf{h}_{\mathbf{B}} \cdot \boldsymbol{\rho}_{2,2}}) \end{array} \right] & \begin{array}{l} \mathbf{k}=1 \text{ d}=2 \\ \mathbf{k}=2 \text{ d}=2 \end{array} \\
\vdots & \vdots & \vdots & \vdots & \\
\left[\begin{array}{c} \text{diag}(\Psi_{\mathbf{D},1}e^{-2\pi i \mathbf{h}_1 \cdot \boldsymbol{\rho}_{\mathbf{K},\mathbf{D}}}) \\ \vdots \\ \text{diag}(\Psi_{\mathbf{D},1}e^{-2\pi i \mathbf{h}_1 \cdot \boldsymbol{\rho}_{\mathbf{K},\mathbf{D}}}) \end{array} \right] \left[\begin{array}{c} \text{diag}(\Psi_{\mathbf{D},2}e^{-2\pi i \mathbf{h}_2 \cdot \boldsymbol{\rho}_{\mathbf{K},\mathbf{D}}}) \\ \vdots \\ \text{diag}(\Psi_{\mathbf{D},2}e^{-2\pi i \mathbf{h}_2 \cdot \boldsymbol{\rho}_{\mathbf{K},\mathbf{D}}}) \end{array} \right] & \cdots & \left[\begin{array}{c} \text{diag}(\Psi_{\mathbf{D},\mathbf{B}}e^{-2\pi i \mathbf{h}_{\mathbf{B}} \cdot \boldsymbol{\rho}_{\mathbf{K},\mathbf{D}}}) \\ \vdots \\ \text{diag}(\Psi_{\mathbf{D},\mathbf{B}}e^{-2\pi i \mathbf{h}_{\mathbf{B}} \cdot \boldsymbol{\rho}_{\mathbf{K},\mathbf{D}}}) \end{array} \right] & \mathbf{k}=\mathbf{K} \text{ d}=\mathbf{D}
\end{array}
\end{array}
\in \mathbb{C}^{\mathbf{KDM} \times \mathbf{BM}}$$

FIG. 4. The first \mathbf{BM} columns of $\mathbf{\Sigma}$ written out explicitly with diagonal matrix blocks of size $\mathbf{M} \times \mathbf{M}$.

APPENDIX C: COMPUTING INFRASTRUCTURE

All reconstructions in this work utilized the National Energy Research Scientific Computing Center Exascale Science Applications Program preparation cluster for the Perlmutter

supercomputer. A single node in the cluster is equipped with two 20-core Intel Xeon processors at 2.40 GHz, with 384 GB DDR4 memory, 930 GB on-node NVMe storage, 8 NVIDIA Tesla V100 GPUs, each with 16 GB HBM2 memory connected with NVLink interconnect, and four dual-port

Mellanox MT27800 (ConnectX-5) EDR InfiniBand HCAs. The reconstruction shown in Fig. 2(a) utilized six nodes with a total of 48 V100 GPUs. The reconstruction shown in Figs. 2(d)–2(f) utilized one node with one V100 GPU. The reconstructions shown in Figs. 3(b) and 3(c) utilized 10 nodes with a total of 80 V100 GPUs.

APPENDIX D: AN ALTERNATING DIRECTION METHOD OF MULTIPLIERS TO SOLVE THE JOINT \mathcal{S} -MATRIX AND PROBE RETRIEVAL PROBLEM

We have also investigated the use of the ADMM algorithm to solve the joint optimization problem of \mathcal{S} and Ψ . While we do not report the convergence results here, the basic ADMM algorithm shown below did not significantly improve convergence speed but added some complexity in the implementation and initialization and has higher memory usage than the block coordinate descent algorithm shown above. It is, however, more flexible in adding additional reconstruction constraints and regularization terms; therefore, we state it here as a reference for future developments. The augmented Lagrangian of the \mathcal{S} -matrix retrieval problem is

$$\mathcal{L}_\beta(\mathcal{S}, \Psi, \mathbf{z}, \mathbf{\Lambda}) = \mathcal{D}(\mathbf{z}) + \text{Re}\mathbf{\Lambda}^\dagger[\mathcal{A}(\mathcal{S}, \Psi) - \mathbf{z}] + \frac{\beta}{2}\|\mathcal{A}(\mathcal{S}, \Psi) - \mathbf{z}\|_2^2, \quad (\text{D1})$$

where $\|\cdot\|_2$ is the l_2 norm and we have introduced the auxiliary variables $\mathbf{z} \in \mathbb{C}^{\text{KDM}_1\text{M}_2}$ and $\mathbf{\Lambda} \in \mathbb{C}^{\text{KDM}_1\text{M}_2}$. We seek to solve for \mathcal{S} and Ψ such that $\mathcal{L}(\mathcal{S}, \Psi, \mathbf{z}, \mathbf{\Lambda})$ is minimized:

$$(\mathcal{S}^*, \Psi^*, \mathbf{z}^*, \mathbf{\Lambda}^*) = \arg \max_{\mathbf{\Lambda}} \arg \min_{\mathcal{S}, \Psi, \mathbf{z}} \mathcal{L}(\mathcal{S}, \Psi, \mathbf{z}, \mathbf{\Lambda}). \quad (\text{D2})$$

ADMM decouples the joint problem into subproblems and solves them step by step:

$$\begin{aligned} \text{Step 1: } \Psi^{l+1} &= \arg \min_{\Psi} \mathcal{L}_\beta^\Psi \\ &:= \arg \min_{\Psi} \mathcal{L}_\beta(\mathcal{S}^l, \Psi, \mathbf{z}^l, \mathbf{\Lambda}^l), \end{aligned} \quad (\text{D3})$$

$$\begin{aligned} \text{Step 2: } \mathcal{S}^{l+1} &= \arg \min_{\mathcal{S}} \mathcal{L}_\beta^\mathcal{S} \\ &:= \arg \min_{\mathcal{S}} \mathcal{L}_\beta(\mathcal{S}, \Psi^{l+1}, \mathbf{z}^l, \mathbf{\Lambda}^l), \end{aligned} \quad (\text{D4})$$

$$\text{Step 3: } \mathbf{z}^{l+1} = \arg \min_{\mathbf{z}} \mathcal{L}_\beta(\mathcal{S}^{l+1}, \Psi^{l+1}, \mathbf{z}, \mathbf{\Lambda}^l), \quad (\text{D5})$$

$$\text{Step 4: } \mathbf{\Lambda}^{l+1} = \mathbf{\Lambda}^l + \beta[\mathbf{z}^{l+1} - \mathcal{A}(\mathcal{S}^{l+1}, \Psi^{l+1})]. \quad (\text{D6})$$

1. Subproblem with respect to Ψ and \mathcal{S}

We solve the subproblems with respect to Ψ and \mathcal{S} with accelerated gradient descent:

$$\Psi^{l+1} = \Psi^l + \gamma_1 \frac{\partial \mathcal{L}_\beta^\Psi}{\partial \Psi}, \quad (\text{D7})$$

$$\mathcal{S}^{l+1} = \mathcal{S}^l + \gamma_2 \frac{\partial \mathcal{L}_\beta^\mathcal{S}}{\partial \mathcal{S}}. \quad (\text{D8})$$

The gradient is

$$\begin{aligned} \frac{\partial \mathcal{L}_\beta^\Psi}{\partial \Psi_{\mathbf{d},\mathbf{b}}}(\mathcal{S}, \Psi, \mathbf{z}) &= \frac{\beta}{M_1 M_2} \sum_{m_1}^{M_1} \sum_{m_2}^{M_2} \left(\sum_{k=1}^K [\mathbf{C}_{k,\mathbf{d}} \mathcal{S}]_{\mathbf{b}}^* e^{2\pi i \mathbf{h}_{\mathbf{b}} \cdot \boldsymbol{\rho}_{k,\mathbf{d}}} \mathcal{F}_{\mathbf{q}}^\dagger \right. \\ &\quad \left. \times \left[\mathbf{z}^l - \mathcal{A}_{k,\mathbf{d}}(\mathcal{S}, \Psi_{\mathbf{d}}) - \frac{\mathbf{\Lambda}^l}{\beta} \right] \right)_{m_1, m_2}, \end{aligned} \quad (\text{D9})$$

$$\begin{aligned} \frac{\partial \mathcal{L}_\beta^\mathcal{S}}{\partial \mathcal{S}_{\mathbf{b}}}(\mathcal{S}, \Psi, \mathbf{z}) &= \frac{\beta}{KD} \sum_{k=1}^K \sum_{\mathbf{d}=1}^D \mathbf{C}_{k,\mathbf{d}}^T \\ &\quad \times \left[\frac{\Psi_{\mathbf{d},\mathbf{b}}^* e^{2\pi i \mathbf{h}_{\mathbf{b}} \cdot \boldsymbol{\rho}_{k,\mathbf{d}}} \mathcal{F}_{\mathbf{q}}^\dagger \left[\mathbf{z}^l - \mathcal{A}_{k,\mathbf{d}}(\mathcal{S}, \Psi_{\mathbf{d}}) - \frac{\mathbf{\Lambda}^l}{\beta} \right]}{\frac{1}{M_1 \cdot M_2} \sum_{\mathbf{b}=1}^B |\Psi_{\mathbf{d},\mathbf{b}}|^2} \right]. \end{aligned} \quad (\text{D10})$$

2. Subproblem with respect to \mathbf{z}

The subproblem with respect to \mathbf{z} was solved elsewhere [90]. The solution is

$$\mathbf{z}^{l+1} = \frac{\text{sgn}(\hat{\mathbf{z}})[\sqrt{\mathbf{I}} + \beta|\hat{\mathbf{z}}|]}{(1 + \beta)}. \quad (\text{D11})$$

The full ADMM algorithm is then given as follows:

Algorithm 2: Joint \mathcal{S} matrix and probe retrieval via ADMM.

text

Input:

measured intensities $\mathbf{I} \in \mathbb{R}^{K \times D \times M_1 \times M_2}$

scan positions $\boldsymbol{\rho} \in \mathbb{R}^{K \times D \times 2}$

initial Fourier space probe phases $\chi^0 \in \mathbb{C}^{D \times M_1 \times M_2}$

step sizes $\gamma_1, \gamma_2, \beta \in \mathbb{R}$

Initialize:

set $(N_1, N_2) = \lceil \frac{\max(\mathbf{r}_2) + M}{M} \rceil \cdot M$ such that the plane waves $e^{i\mathbf{h} \cdot \mathbf{r}}$

have periodic boundary conditions

calculate $\mathbf{I}^{\text{mean}} = \frac{1}{K} \sum_{k=1}^K \mathbf{I}_k$ and

$a_{\max} = \max\{|\mathbf{I}_k|_1 \mid \forall k = \{1, \dots, K\}\}$

set $\Psi^0 \leftarrow \frac{a_{\max}}{\sqrt{\|\mathbf{I}^{\text{mean}}\|_1}} \mathbf{I}^{\text{mean}} e^{i\chi^0}$

$\mathcal{S}_{\mathbf{b}}^0 \leftarrow e^{i\mathbf{h}_{\mathbf{b}} \cdot \mathbf{r}}, \mathcal{S} \in \mathbb{C}^{B \times N_1 \times N_2}$

for $l = 0$ to L **do**

$$\hat{\mathbf{z}} = \mathbf{z}^l + \frac{\mathbf{\Lambda}^l}{\beta};$$

$$\mathcal{S}^{l+1} \leftarrow \mathcal{S}^l + \gamma_2 \cdot \frac{\partial \mathcal{L}_\beta^\mathcal{S}}{\partial \mathcal{S}}(\mathcal{S}^l, \Psi^{l+1}, \hat{\mathbf{z}});$$

$$\Psi^{l+1} \leftarrow \Psi^l + \gamma_1 \cdot \frac{\partial \mathcal{L}_\beta^\Psi}{\partial \Psi}(\mathcal{S}^l, \Psi^l, \hat{\mathbf{z}});$$

$$\hat{\mathbf{z}} \leftarrow \mathcal{A}(\mathcal{S}^{l+1}, \Psi^{l+1}) - \frac{\mathbf{\Lambda}^l}{\beta};$$

$$\mathbf{z}^{l+1} \leftarrow \frac{\text{sgn}(\hat{\mathbf{z}})[\sqrt{\mathbf{I}} + \beta|\hat{\mathbf{z}}|]}{(1 + \beta)};$$

$$\mathbf{\Lambda}^{l+1} \leftarrow \mathbf{\Lambda}^l + \beta[\mathbf{z}^{l+1} - \mathcal{A}(\mathcal{S}^{l+1}, \Psi^{l+1})];$$

Output: $\mathcal{S}^* = \mathcal{S}^L$

- [1] M. Pluta, *Advanced Light Microscopy* (Elsevier, Amsterdam, 1988), Vol. 1.
- [2] A. Clarke and C. N. Eberhardt, *Microscopy Techniques for Materials Science* (Woodhead, 2002).
- [3] J. Kirz, C. Jacobsen, and M. Howells, Soft X-ray microscopes and their biological applications, *Q. Rev. Biophys.* **28**, 33 (1995).
- [4] S. C. Mayo, A. W. Stevenson, and S. W. Wilkins, In-line phase-contrast X-ray imaging and tomography for materials science, *Materials* **5**, 937 (2012).
- [5] J. C. H. Spence, The future of atomic resolution electron microscopy for materials science, *Mater. Sci. Eng., R* **26**, 1 (1999).
- [6] R. M. Glaeser, Invited review article: Methods for imaging weak-phase objects in electron microscopy, *Rev. Sci. Instrum.* **84**, 111101 (2013).
- [7] F. Zernike, Phase contrast, a new method for the microscopic observation of transparent objects, *Physica* **9**, 686 (1942).
- [8] L. E. Ballentine, The statistical interpretation of quantum mechanics, *Rev. Mod. Phys.* **42**, 358 (1970).
- [9] R. Barer, A vector theory of phase contrast and interference contrast. I. Positive phase contrast, *J. R. Microsc. Soc.* **72**, 10 (1952).
- [10] S. Maretzke, A uniqueness result for propagation-based phase contrast imaging from a single measurement, *Inverse Probl.* **31**, 065003 (2015).
- [11] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Elsevier, 2013).
- [12] H. A. Cohen, M. F. Schmid, and W. Chiu, Estimates of validity of projection approximation for three-dimensional reconstructions at high resolution, *Ultramicroscopy* **14**, 219 (1984).
- [13] A. Burvall, U. Lundström, P. A. C. Takman, D. H. Larsson, and H. M. Hertz, Phase retrieval in X-ray phase-contrast imaging suitable for tomography, *Opt. Express* **19**, 10359 (2011).
- [14] M. Vulović, L. M. Voortman, L. J. van Vliet, and B. Rieger, When to use the projection assumption and the weak-phase object approximation in phase contrast cryo-EM, *Ultramicroscopy* **136**, 61 (2014).
- [15] A. V. Crewe and T. Groves, Thick specimens in the CEM and STEM. I. Contrast, *J. Appl. Phys.* **45**, 3662 (1974).
- [16] H. Bethe, The theory of electron diffraction on crystals, *Ann. Phys. (Berlin, Ger.)* **392**, 55 (1928).
- [17] J. M. Cowley and A. F. Moodie, The scattering of electrons by atoms and crystals. I. A new theoretical approach, *Acta Crystallogr.* **10**, 609 (1957).
- [18] S. Chowdhury, M. Chen, R. Eckert, D. Ren, F. Wu, N. Repina, and L. Waller, High-resolution 3D refractive index microscopy of multiple-scattering samples from intensity images, *Optica* **6**, 1211 (2019).
- [19] T. M. Godden, R. Suman, M. J. Humphry, J. M. Rodenburg, and A. M. Maiden, Ptychographic microscope for three-dimensional imaging, *Opt. Express* **22**, 12513 (2014).
- [20] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, Learning approach to optical tomography, *Optica* **2**, 517 (2015).
- [21] P. Li and A. Maiden, Multi-slice ptychographic tomography, *Sci. Rep.* **8**, 2049 (2018).
- [22] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg, Ptychographic transmission microscopy in three dimensions using a multi-slice approach, *J. Opt. Soc. Am. A* **29**, 1606 (2012).
- [23] A. Suzuki, S. Furutaku, K. Shimomura, K. Yamauchi, Y. Kohmura, T. Ishikawa, and Y. Takahashi, High-Resolution Multislice X-ray Ptychography of Extended Thick Objects, *Phys. Rev. Lett.* **112**, 053903 (2014).
- [24] K. Shimomura, A. Suzuki, M. Hirose, and Y. Takahashi, Precession x-ray ptychography with multislice approach, *Phys. Rev. B* **91**, 214114 (2015).
- [25] E. H. R. Tsai, I. Usov, A. Diaz, A. Menzel, and M. Guizar-Sicairos, X-ray ptychography with extended depth of field, *Opt. Express* **24**, 29089 (2016).
- [26] H. Öztürk, H. Yan, Y. He, M. Ge, Z. Dong, M. Lin, E. Nazaretski, I. K. Robinson, Y. S. Chu, and X. Huang, Multislice ptychography with large numerical aperture multilayer Laue lenses, *Optica* **5**, 601 (2018).
- [27] W. Van den Broek and C. T. Koch, Method for Retrieval of the Three-Dimensional Object Potential by Inversion of Dynamical Electron Scattering, *Phys. Rev. Lett.* **109**, 245502 (2012).
- [28] W. Van den Broek and C. T. Koch, General framework for quantitative three-dimensional reconstruction from arbitrary detection geometries in TEM, *Phys. Rev. B* **87**, 184108 (2013).
- [29] S. Gao, P. Wang, F. Zhang, G. T. Martinez, P. D. Nellist, X. Pan, and A. I. Kirkland, Electron ptychographic microscopy for three-dimensional imaging, *Nat. Commun.* **8**, 163 (2017).
- [30] M. Schloz, T. C. Pekin, Z. Chen, W. Van den Broek, D. A. Muller, and C. T. Koch, Overcoming information reduced data and experimentally uncertain parameters in ptychography with regularized optimization, *Opt. Express* **28**, 28306 (2020).
- [31] E. H. R. Tsai, F. Marone, and M. Guizar-Sicairos, Gridrec-MS: An algorithm for multi-slice tomography, *Opt. Lett.* **44**, 2181 (2019).
- [32] C. Jacobsen, Relaxation of the Crowther criterion in multislice tomography, *Opt. Lett.* **43**, 4811 (2018).
- [33] H. L. Xin and D. A. Muller, Aberration-corrected ADF-STEM depth sectioning and prospects for reliable 3D imaging in S/TEM, *J. Electron Microsc.* **58**, 157 (2009).
- [34] M. A. Gilles, Y. S. G. Nashed, M. Du, C. Jacobsen, and S. M. Wild, 3D x-ray imaging of continuous objects beyond the depth of focus limit, *Optica* **5**, 1078 (2018).
- [35] M. Du, Y. S. G. Nashed, S. Kandel, D. Gürsoy, and C. Jacobsen, Three dimensions, two microscopes, one code: Automatic differentiation for x-ray nanotomography beyond the depth of focus limit, *Sci. Adv.* **6**, eaay3700 (2020).
- [36] D. Ren, C. Ophus, M. Chen, and L. Waller, A multiple scattering algorithm for three dimensional phase contrast atomic electron tomography, *Ultramicroscopy* **208**, 112860 (2020).
- [37] F. Fujimoto, Dynamical theory of electron diffraction in Laue-case, I. General theory, *J. Phys. Soc. Jpn.* **14**, 1558 (1959).
- [38] L. Sturkey, The calculation of electron diffraction intensities, *Proc. Phys. Soc.* **80**, 321 (1962).
- [39] C. Ophus, A fast image simulation algorithm for scanning transmission electron microscopy, *Adv. Struct. Chem. Imaging* **3**, 13 (2017).
- [40] H. G. Brown, J. Ciston, and C. Ophus, Linear-scaling algorithm for rapid computation of inelastic transitions in the presence of multiple electron scattering, *Phys. Rev. Research* **1**, 033186 (2019).

- [41] J. C. H. Spence, Direct inversion of dynamical electron diffraction patterns to structure factors, *Acta Crystallogr., Sect. A* **54**, 7 (1998).
- [42] L. J. Allen, T. W. Josefsson, and H. Leeb, Obtaining the crystal potential by inversion from electron scattering intensities, *Acta Crystallogr., Sect. A* **54**, 388 (1998).
- [43] L. J. Allen, H. M. L. Faulkner, and H. Leeb, Inversion of dynamical electron diffraction data including absorption, *Acta Crystallogr., Sect. A* **56**, 119 (2000).
- [44] J. J. Donatelli and J. C. H. Spence, Inversion of Many-Beam Bragg Intensities for Phasing by Iterated Projections: Removal of Multiple Scattering Artifacts from Diffraction Data, *Phys. Rev. Lett.* **125**, 065502 (2020).
- [45] P. Rez, Schemes to determine the crystal potential under dynamical conditions using voltage variation, *Acta Crystallogr., Sect. A* **55**, 160 (1999).
- [46] F. Wang, R. S. Pennington, and C. T. Koch, Inversion of Dynamical Scattering from Large-Angle Rocking-Beam Electron Diffraction Patterns, *Phys. Rev. Lett.* **117**, 015501 (2016).
- [47] S. D. Findlay, Quantitative structure retrieval using scanning transmission electron microscopy, *Acta Crystallogr., Sect. A* **61**, 397 (2005).
- [48] H. G. Brown, Z. Chen, M. Weyland, C. Ophus, J. Ciston, L. J. Allen, and S. D. Findlay, Structure Retrieval at Atomic Resolution in the Presence of Multiple Scattering of the Electron Probe, *Phys. Rev. Lett.* **121**, 266102 (2018).
- [49] S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, Measuring the Transmission Matrix in Optics: An Approach to the Study and Control of Light Propagation in Disordered Media, *Phys. Rev. Lett.* **104**, 100601 (2010).
- [50] M. Kim, Y. Choi, C. Yoon, W. Choi, J. Kim, Q.-H. Park, and W. Choi, Maximal energy transport through disordered media with the implementation of transmission eigenchannels, *Nat. Photonics* **6**, 581 (2012).
- [51] C. A. Metzler, M. K. Sharma, S. Nagesh, R. G. Baraniuk, O. Cossairt, and A. Veeraraghavan, Coherent inverse scattering via transmission matrices: Efficient phase retrieval algorithms and a public dataset, in *2017 IEEE International Conference on Computational Photography (ICCP)* (IEEE, Piscataway, NJ, 2017), pp. 1–16.
- [52] A. Drémeau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques, *Opt. Express* **23**, 11898 (2015).
- [53] B. Rajaei, E. W. Tramel, S. Gigan, F. Krzakala, and L. Daudet, Intensity-only optical compressive imaging using a multiply scattering material and a double phase retrieval approach, in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, Piscataway, NJ, 2016), pp. 4054–4058.
- [54] H. Yu, T. R. Hillman, W. Choi, J. O. Lee, M. S. Feld, R. R. Dasari, and Y. Park, Measuring Large Optical Transmission Matrices of Disordered Media, *Phys. Rev. Lett.* **111**, 153902 (2013).
- [55] H. G. Brown, P. M. Pelz, S.-L. Hsu, Z. Zhang, R. Ramesh, K. Inzani, E. Sheridan, S. M. Griffin, S. D. Findlay, L. J. Allen *et al.*, A three-dimensional reconstruction algorithm for scanning transmission electron microscopy data from thick samples, [arXiv:2011.07652](https://arxiv.org/abs/2011.07652).
- [56] E. J. Kirkland, *Advanced Computing in Electron Microscopy* (Springer, 1998).
- [57] S. Rotter and S. Gigan, Light fields in complex media: Mesoscopic scattering meets wave control, *Rev. Mod. Phys.* **89**, 015005 (2017).
- [58] J. Miao, D. Sayre, and H. N. Chapman, Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects, *J. Opt. Soc. Am. A* **15**, 1662 (1998).
- [59] J. R. Fienup, Phase retrieval algorithms: A comparison, *Appl. Opt.* **21**, 2758 (1982).
- [60] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, Phase retrieval with application to optical imaging: A contemporary overview, *IEEE Signal Process. Mag.* **32**, 87 (2015).
- [61] E. J. Candes, T. Strohmer, and V. Voroninski, Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming, *Commun. Pure Appl. Math.* **66**, 1241 (2013).
- [62] I. Waldspurger, A. d'Aspremont, and S. Mallat, Phase recovery, MaxCut and complex semidefinite programming, *Math. Program.* **149**, 47 (2015).
- [63] E. Bostan, M. Soltanolkotabi, D. Ren, and L. Waller, Accelerated Wirtinger flow for multiplexed fourier ptychographic microscopy, in *2018 25th IEEE International Conference on Image Processing (ICIP)* (IEEE, Piscataway, NJ, 2018), pp. 3823–3827.
- [64] V. Nikitin, S. Aslan, Y. Yao, T. Biçer, S. Leyffer, R. Mokso, and D. Gürsoy, Photon-limited ptychography of 3D objects via Bayesian reconstruction, *OSA Continuum* **2**, 2948 (2019).
- [65] P. Thibault and M. Guizar-Sicairos, Maximum-likelihood refinement for coherent diffractive imaging, *New J. Phys.* **14**, 063004 (2012).
- [66] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.3.023159> for supplementary figures showing simultaneous HAADF-STEM images, mean diffraction patterns of the data set and count histogram, aligned single-sideband ptychography reconstructions, and example diffraction patterns.
- [67] O. L. Krivanek, N. Dellby, and A. R. Lupini, Towards sub-Å electron beams, *Ultramicroscopy* **78**, 1 (1999).
- [68] A. Paszke *et al.*, Pytorch: An imperative style, high-performance deep learning library, [arXiv:1912.01703](https://arxiv.org/abs/1912.01703).
- [69] C. Ophus, Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond, *Microsc. Microanal.* **25**, 563 (2019).
- [70] J. Ciston, I. J. Johnson, B. R. Draney, P. Ercius, E. Fong, A. Goldschmidt, J. M. Joseph, J. R. Lee, A. Mueller, C. Ophus, A. Selvarajan, D. E. Skinner, T. Stezelberger, C. S. Tindall, A. M. Minor, and P. Denes, The 4D Camera: Very high speed electron counting for 4D-STEM, *Microsc. Microanal.* **25**, 1930 (2019).
- [71] S. J. Pennycook and P. D. Nellist, *Scanning Transmission Electron Microscopy: Imaging and Analysis* (Springer Science & Business Media, 2011).
- [72] B. E. Brown, The crystal structures of NbTe₂ and TaTe₂, *Acta Crystallogr.* **20**, 264 (1966).
- [73] J. Su, K. Liu, F. Wang, B. Jin, Y. Guo, G. Liu, H. Li, and T. Zhai, Van der Waals 2D transition metal tellurides, *Adv. Mater. Interfaces* **6**, 1900741 (2019).

- [74] K. M. Siddiqui, D. B. Durham, F. Cropp, C. Ophus, S. Rajpurohit, Y. Zhu, J. D. Carlström, C. Stavrakas, Z. Mao, A. Raja, P. Musumeci, L. Z. Tan, A. M. Minor, D. Filippetto, and R. A. Kaindl, Ultrafast optical melting of trimer superstructure in layered 1T'-TaTe₂, [arXiv:2009.02891](https://arxiv.org/abs/2009.02891).
- [75] P. Ercius, C. Harris, and P. Avery, STEMPY: Toolkit for processing 4D STEM image data on hpc, <https://zenodo.org/record/4383275#.X-EwtNhKguU>.
- [76] J. M. Rodenburg, B. C. McCallum, and P. D. Nellist, Experimental tests on double-resolution coherent imaging via stem, *Ultramicroscopy* **48**, 304 (1993).
- [77] P. Pochet, B. C. McGuigan, J. Coraux, and H. T. Johnson, Toward moiré engineering in 2D materials via dislocation theory, *Appl. Mater. Today* **9**, 240 (2017).
- [78] M. Uchida and S. Horiuchi, New approximant of twelvefold quasicrystal in Ta-Te: Electron diffraction study, *Jpn. J. Appl. Phys.* **37**, L531 (1998).
- [79] J. D. Cain, A. Azizi, M. Conrad, S. M. Griffin, and A. Zettl, Layer-dependent topological phase in a two-dimensional quasicrystal and approximant, *Proc. Natl. Acad. Sci. USA* **117**, 26135 (2020).
- [80] S. Bajt *et al.*, X-ray focusing with efficient high-NA multilayer Laue lenses, *Light: Sci. Appl.* **7**, 17162 (2018).
- [81] H. N. Chapman and S. Bajt, A ray-trace analysis of x-ray multilayer Laue lenses for nanometer focusing, *J. Opt.* **22**, 115610 (2020).
- [82] J. Verbeeck, A. Béch  , K. M  ller-Caspary, G. Guzzinati, M. A. Luong, and M. Den Hertog, Demonstration of a 2×2 programmable phase plate for electrons, *Ultramicroscopy* **190**, 58 (2018).
- [83] F. Kong, R. H. Silverman, L. Liu, P. V. Chitnis, K. K. Lee, and Y. C. Chen, Photoacoustic-guided convergence of light through optically diffusive media, *Opt. Lett.* **36**, 2053 (2011).
- [84] S. J. Haigh, H. Sawada, and A. I. Kirkland, Atomic Structure Imaging Beyond Conventional Resolution Limits in the Transmission Electron Microscope, *Phys. Rev. Lett.* **103**, 126101 (2009).
- [85] A. M. Maiden, M. J. Humphry, M. C. Sarahan, B. Kraus, and J. M. Rodenburg, An annealing algorithm to correct positioning errors in ptychography, *Ultramicroscopy* **120**, 64 (2012).
- [86] M. Odstr  il, A. Menzel, and M. Guizar-Sicairos, Iterative least-squares solver for generalized maximum-likelihood ptychography, *Opt. Express* **26**, 3108 (2018).
- [87] A. Rana, J. Zhang, M. Pham, A. Yuan, Y.-H. Lo, H. Jiang, S. J. Osher, and J. Miao, Potential of Attosecond Coherent Diffractive Imaging, *Phys. Rev. Lett.* **125**, 086101 (2020).
- [88] P. Thibault and A. Menzel, Reconstructing state mixtures from diffraction measurements, *Nature (London)* **494**, 68 (2013).
- [89] Z. Chen, M. Odstr  il, Y. Jiang, Y. Han, M.-H. Chiu, L.-J. Li, and D. A. Muller, Mixed-state electron ptychography enables sub-angstrom resolution imaging with picometer precision at low dose, *Nat. Commun.* **11**, 2994 (2020).
- [90] Z. Wen, C. Yang, X. Liu, and S. Marchesini, Alternating direction methods for classical and ptychographic phase retrieval, *Inverse Probl.* **28**, 115010 (2012).