Fundamentals of focal series inline electron holography

Fundamentals of Focal Series Inline Electron Holography

A. Lubk, K. Vogel, D. Wolf, J. Krehl, and F. Röder

Triebenberg Laboratory, Institute of Solid State Physics,
Technische Universität Dresden, 01062 Dresden, Germany

L. Clark, G. Guzzinati, and J. Verbeeck

EMAT, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

The loss of the quantum phase information in the measurement process constitutes a fundamental limitation to transmission electron microscopy as the electron wave's phase often contains valuable information about the studied specimen. Phase retrieval from focal series is a holographic technique that seeks to recover the lost information from a set of images recorded at different defoci. In spite of its widespread use, in particular for wave reconstruction at atomic resolution, a number of fundamental properties, e.g., regarding the conditions for a unique wave function reconstruction, the magnitude of the reconstruction error and the influence of inconsistencies or incomplete data are not well-understood as of today. Here, we elaborate on the fundamentals of the technique, making extensive use of the tomographic representation of a focal series as tilt series in phase space. Using this perspective we discuss, among others, requirements on the focal series for a unique reconstruction, such a the focus interval ranging from the far field at underfocus to the far field at overfocus or the focus step size. We reveal that the prominent Gerchberg-Saxton iterative projection algorithm corresponds to a numerical integration of the quantum Hamilton-Jacobi equation in the small focus step limit. Moreover, we show that the topology of the starting guess divides the solution space of the Gerchberg-Saxton algorithm into equivalence classes, which mitigates the impact of the incompleteness of typical focal series data. To facilitate a focal series reconstruction meeting the above theoretical requirements such as the long range focus interval, we develop a dedicated calibration procedure facilitating the determination of unknown electron optical parameters such as the focal length of the principal imaging lens or the position of object and image planes. The findings are demonstrated with an example of a focal series reconstruction of an electron vortex beam.

Keywords: Electron Holography, Quantum State Tomography, Focal Series Phase Retrieval

CONTENTS

I. INTRODUCTION

Focal series wave reconstruction in the Transmission Electron Microscope (TEM) is a well-established holographic technique for mapping electric, magnetic and strain fields (Dietrich et al. 2014, Koch et al. 2010, Song et al. 2013) in solids with nanometer resolution and for studying atomic configurations at crystal defects or grain boundaries (Allen et al. 2004, Koch 2014, Thust et al. 1996). More recently, also the structure of phase vortices at caustics has been explored using the technique (Petersen et al. 2013). The schematic optical setup, depicted in Figure 1, is an extension of Gabor's original inline holography (Gabor 1948), which used only one out-of-focus image for the reconstruction of the wave function. Focal series reconstruction does not require an undisturbed reference wave and a biprism like off-axis holography, hence may be carried out at practically every TEM. Furthermore, it may be conducted under relaxed partial coherence provided that the latter is well-behaved and well-known in advance (Koch 2008). Moreover, focal series holography possesses an intriguing connection to quantum state tomography (Lubk and Röder 2015), a technique that is successfully employed to study mixed (i.e., incoherent) quantum states of matter (e.g., atoms) and light (Breitenbach et al. 1997, Schleich 2001, Smithey et al. 1993).

These advantages are opposed by ambiguities in the reconstructed wave function (Fienup and Wackerman 1986, Luke et al. 2002), e.g., due to inconsistent and incomplete focal series data. Under experimental conditions

every focal series is inconsistent due to the presence of partial coherence (e.g., from the electron gun, inelastic interaction and thermal diffuse scattering (Rother et al. 2009)) or shot and detector noise (Niermann et al. 2012) as well as geometric and chromatic aberrations depending on the defocus. Similarly, every focal series is incomplete because of a limited number and range of defocus values, typically limited to the near field regime, and the restriction to isotropic foci, whereas astigmatic foci are required for an unambiguous reconstruction of a wave function (Lubk and Röder 2015). For instance, the problematic reconstruction of low spatial frequencies (Niermann and Lehmann 2016, Ophus and Ewalds 2012, Thust et al. 1996) can be traced back to missing focal series data in the far field.

The design of reconstruction algorithms behaving well in the presence of inconsistent and incomplete data including their characterization in terms of reconstruction errors is highly non-trivial and subject to large and ongoing efforts. Moreover, the accurate calibration of crucial experimental parameters such as the precise defocus values or aberrations and distortions in the series remain challenging. Thus, in spite of the relative experimental convenience of this technique, the majority of quantitative electric, magnetic and strain field studies are conducted by means of off-axis electron holography (see (Dunin-Borkowski et al. 2004, Kasama et al. 2011, Lichte et al. 2013, McCartney et al. 2010, Pozzi et al. 2014, Tonomura 1987, Völkl et al. 1999) and references therein), which has a less problematic linear and unique reconstruction procedure including well-defined error estimates (Lenz 1988, Röder et al. 2014).

In the following, we elaborate on analyzing focal series reconstruction from the perspective of quantum state tomography and use the obtained results to increase the scope of the technique in terms of convergence and uniqueness in particular for low spatial frequencies. Moreover, we are able to explain a number of previous results by exploiting the phase space analogy, and open pathways to further improvements.

The history of focal series reconstructions goes back to the first half of the 20th century, when Pauli raised the question, whether the amplitude of a complex (wave) function and of its Fourier transform completely define the underlying function (Pauli 1933). This phase retrieval problem can be considered as one particular instance of a reconstruction from a focal series consisting only of one in-focus and one far-field image. In the wake of finding whole classes of functions (up to constant phase factors) sharing the same amplitude both in position space and Fourier space (Jaming 2014, Luke et al. 2002), the prospects of such reconstructions have been considered rather pessimistic. For instance, if the Fourier transforms are symmetric $\mathcal{F}\{\Psi\}(k) = \mathcal{F}\{\Psi\}(-k)$, the corresponding amplitudes are invariant under complex
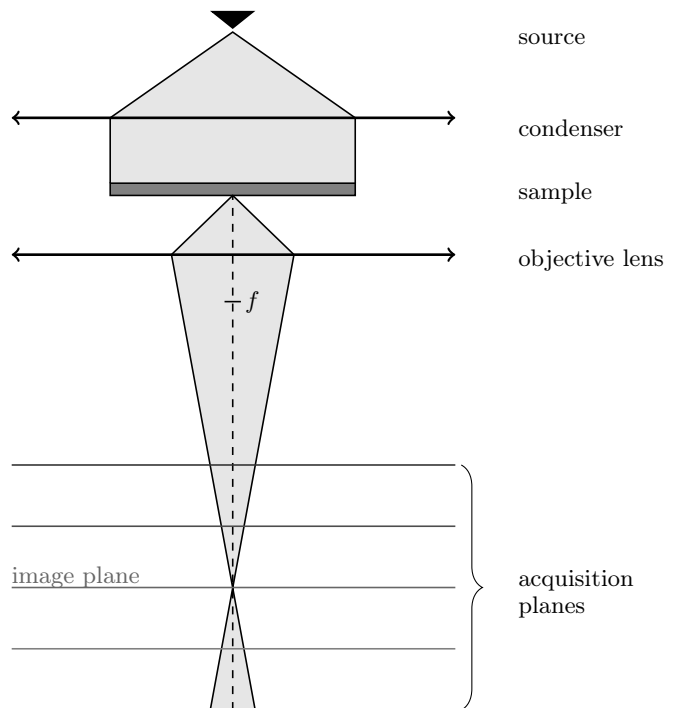

Figure 1. Principle optical setup of focal series inline holography. Accordingly, the defocus interval is ideally bounded by the far field located below (i.e., in focal plane f) and above the image plane.

conjugation, i.e.,

$$|\Psi(x)| = |\Psi^*(x)| \tag{1}$$
$$|\mathcal{F}\{\Psi\}(k)| = |\mathcal{F}\{\Psi^*\}(k)| \ .$$

In a seminal paper Gerchberg and Saxton, however, proposed a simple iterative projection algorithm, which produces surprisingly sensible wave reconstruction from two-dimensional image and diffraction intensities (Gerchberg and Saxton 1972). Subsequently, modifications and extension of the original Gerchberg-Saxton algorithm (Boucher 1980, Cederquist et al. 1989, Combettes and Trussell 1990, Dong et al. 1997, Fienup 1982, Gonsalves 1976, Levi and Stark 1984, Misell 1973, Ohneda et al. 2001, Quatieri and Oppenheim 1981, Stark and Sezan 1994, Takajo et al. 2002, Yang et al. 1994, Zou and Unbehauen 1997) have been successfully applied in a large number of different fields ranging from measuring the aberrations of telescopes (Baba and Mutoh 2001, Fienup et al. 1993, Lyon et al. 1997) to phase retrieval in electron microscopy (Allen et al. 2004, Allen and Oxley 2001, Koch 2008, 2014). It is interesting to note that in particular the unfortunate initial manufacturing error of the Hubble space telescope's optics (Burrows et al. 1991) and its subsequent characterization (Burrows 1991, Fienup et al. 1993, Lyon et al. 1997, 1991) lead to a boost in research on wave front reconstruction algorithms and their properties, which exhibits a remarkable parallel to Gabor's invention of holography that had been stimulated

by the search for aberration corrected electron optics in the TEM (Gabor 1948).

Fienup established the connection between the Gerchberg-Saxton algorithm (Gerchberg and Saxton 1972) and line-search methods. Indeed, the Gerchberg-Saxton algorithm is closely related to the discipline of convex optimization (Bauschke et al. 2002, Levi and Stark 1984), where similar algorithms, referred to as projection on convex sets or alternating projections, are employed. Within the convex setting, these algorithms are well-behaved and well-understood in their convergence behaviour. However, the Gerchberg-Saxton algorithm operates on non-convex sets, as will be discussed below. Thus, there remain a number of properties, which remain not well understood till today, in spite of the large progress made in the field of focal series reconstructions. For instance, the theoretical results on the convergence of convex projection algorithms do not apply to the problem of phase retrieval from focal series (Combettes 1996, Luke et al. 2002, Seldin and Fienup 1990), resulting in stalling iterations or reconstruction artifacts. Therefore, a large number of studies has been concerned with the investigation of constraints on the reconstructed function space narrowing the number of potential solutions (e.g,, pertaining to a different starting guess in the Gerchberg-Saxton algorithm), such as limited supports in position or Fourier space (Hayes et al. 1980, Marchesini et al. 2015) or the analyticity of the wave function (Huiser et al. 1976, Huiser and Ferwerda 1976, Huiser et al. 1977). Important aspects of these studies are whether the studied constraints are applicable to the experimental situation (e.g., taking into account inconsistent (noisy) data) and how these constraints can be implemented in actual reconstruction algorithms.

A particularly useful perspective is provided by lifting the phase retrieval problem into quantum mechanical phase space, where it corresponds to a matrix completion problem in the discrete setting (Candès et al. 2013), because the sets possessing the same image intensity are convex in phase space. Moreover, it is well understood that a focal series corresponds to a tilt series in quantum mechanical phase space (Lubk and Röder 2015, Raymer et al. 1994). Thus, the focal series reconstruction translates to a tomographic reconstruction of the underlying phase space distribution. In the following we explore this analogy and its manifold consequences for the Gerchberg-Saxton type reconstruction algorithms. Most importantly, we will show that the focal range needs to extend over both near and far field to facilitate a unique reconstruction. We implement such a long-range focal series reconstruction using a variant of the Landweber iteration, well-known from (quantum state) tomography (Natterer and Wübbeling 2001). The significance of the large focal range for a unique reconstruction is also proven by showing that the Gerchberg-Saxton algorithm is closely related to the numerical integration of the quantum Hamilton-Jacobi equation. Moreover, we discuss several boundary conditions narrowing the under-determined focal series reconstruction from a limited number of isotropic defocused images, such as topology or regularity constraints.

This article is organized as follows. We begin with a brief discussion of imaging with a thin lens, both wave optically as well as in the phase space setting, thereby providing the basis for the subsequent focal series reconstructions. We proceed with discussing some crucial experimental aspects pertaining to the recording of a long range focal series. Subsequently, we elaborate on the Gerchberg-Saxton reconstruction algorithm. Finally, we perform a case study using a higher-order electron vorex beam, not in the least because this topic has garnered much interest and development in the recent years. Such beams possess a non-trivial topology by design, and are therefore nicely suited to discuss the crucial impact of (implicit) topology constraints. The phase retrieval of such waves has proved elusive because of missing undisturbed plane references for off-axis holography or non-trivial boundary conditions in direct reconstruction techniques such as the Transport of Intensity (TIE) phase retrieval (Lubk et al. 2013b, Teague 1983).

II. THIN LENS IMAGING

We begin with introducing one particular representation of quantum mechanical phase space - the Wigner function, which is computed from a wave function Ψ (a state vector) according to the following definition (e.g. (Schleich 2001))

$$W(r,k) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi^* \left(r - \frac{1}{2}r' \right) \Psi \left(r + \frac{1}{2}r' \right) e^{-ikr'} \mathrm{d}r' \tag{2}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi^* \left(k - \frac{1}{2}k' \right) \Psi \left(k + \frac{1}{2}k' \right) e^{ik'r} \mathrm{d}k',$$

where the normalization factor $1/(2\pi)$ ensures

$$\int_{-\infty}^{\infty} \mathrm{d}r \int_{-\infty}^{\infty} \mathrm{d}k W(r,k) = 1. \tag{3}$$

It is important to note that the above definition is just a special case of the more general

$$W(r,k) := \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}r' \rho \left(r + \frac{1}{2}r', r - \frac{1}{2}r' \right) e^{-ikr'}, \tag{4}$$

where any mixed quantum state (i.e., incoherent superposition of pure state wave functions) represented by some density matrix

$$\rho \left(r + \frac{1}{2}r', r - \frac{1}{2}r' \right) = \sum_n \Psi_n^* \left(r - \frac{1}{2}r' \right) \Psi_n \left(r + \frac{1}{2}r' \right) \tag{5}$$

is allowed. The parameter space (r,k) of the Wigner function defines a quantum mechanical phase space, with
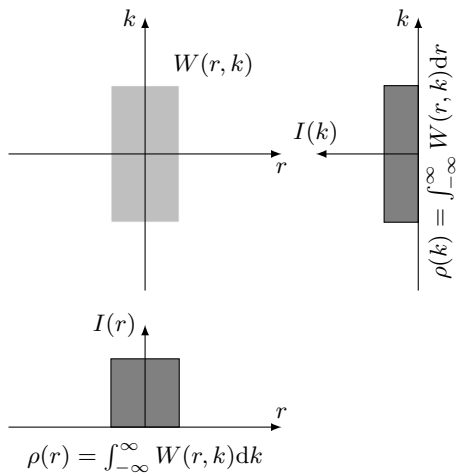
Figure 2. A Wigner function and its marginal densities.

coordinate k being the momentum coordinate. The two dimensional version required for electron optics follows in a straightforward manner by replacing the phase space coordinates r and k with 2D vectors $\mathbf{r}$ and $\mathbf{k}$.

The definition (2) implies that the Wigner function is strictly real. Projecting the Wigner function along k

yields a strictly positive quantity, which can be identified as the quantum mechanical density in position space

$$\rho\left(r\right) = \int_{-\infty}^{\infty} W\left(r,k\right) \mathrm{d}k\,. \tag{6}$$

The same holds for Fourier space, i.e.,

$$\rho\left(k\right) = \int_{-\infty}^{\infty} W\left(r,k\right) \mathrm{d}r\,. \tag{7}$$

The latter projection along r could be alternatively understood as a rotation of phase space around 90° with a subsequent projection along the vertical axis (Figure 2). The generalization of this concept to arbitrary tilt angles leads to the notion of fractional Fourier transforms (e.g., (Almeida 1994, Ozaktas and Mendlovic 1995)), which is closely related to the free space propagation in the paraxial regime (Fresnel propagation) discussed subsequently.

In the following, we only require the behaviour of the Wigner function under free space propagation. Within the paraxial regime it may be obtained by applying the Fresnel propagator to the wave function (propagation length z, wave number k_0) and inserting the propagated wave function in the above definition

$$
\begin{aligned}
W_{\delta z}(\mathbf{r},\mathbf{k}) &= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \mathrm{d}^2 k' e^{i\frac{\delta z}{2k_0}\left(\mathbf{k}-\frac{1}{2}\mathbf{k}'\right)^2} \Psi^*\left(\mathbf{k}-\frac{1}{2}\mathbf{k}'\right) \Psi\left(\mathbf{k}+\frac{1}{2}\mathbf{k}'\right) e^{-i\frac{\delta z}{2k_0}\left(\mathbf{k}+\frac{1}{2}\mathbf{k}'\right)^2} e^{i\mathbf{k}'\mathbf{r}} \\
&= \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \mathrm{d}^2 k' \Psi^*\left(\mathbf{k}-\frac{1}{2}\mathbf{k}'\right) \Psi\left(k+\frac{1}{2}\mathbf{k}'\right) e^{i\mathbf{k}'\left(\mathbf{r}-\frac{\delta z}{k_0}\mathbf{k}\right)} \\
&= W_0(\mathbf{r}-\frac{\delta z}{k_0}\mathbf{k},\mathbf{k})\,.
\end{aligned}
\tag{8}
$$

This result, also holding for mixed states, exhibits a remarkable simplicity. The paraxially propagated Wigner function is merely a sheared version of its initial state W_0, which is exactly the behaviour of a classical phase space ensemble of particles with position $\mathbf{r}$ and momentum $\mathbf{k}$. In other words, the quantum nature of the system is solely encoded in the underlying Wigner quasiprobability distribution.

It is now crucial that one may transform this shear into a rotation by suitably shearing (along $\mathbf{k}$) and rescaling the propagated Wigner function. Note that these transformations do not affect the projections along $\mathbf{k}$, i.e., the recorded intensities. Indeed, a focal series is nothing else then a scaled series of intensities of freely propagated waves (see below), and we may conclude that a focal series reconstruction corresponds to a tomographic reconstruction of the Wigner function, referred to as quantum state tomography (Raymer et al. 1994) (Figure 3). Because the Wigner function describes both pure quantum states (i.e., wave function) and mixed quantum states,

this is not just a mere reformulation of focal series phase retrieval of wave functions in phase space. Indeed, quantum state tomography may be employed to reconstruct quantum states of any coherence and the restriction to wave functions (pure states) is given by constraining the phase space volume of the quantum state according to

$$4\pi^2 \iint_{-\infty}^{\infty} \mathrm{d}^2 r \mathrm{d}^2 k W^2\left(\mathbf{r},\mathbf{k}\right) = 1\,. \tag{9}$$

Because the latter integral constraint only marginally reduces the infinitely-dimensional space of quantum states by one dimension, important properties of phase space tomography, e.g., pertaining to the required focal range, may be directly transferred to focal series reconstruction of wave functions, where they are much less obvious. Most importantly, the 180° tilt range required for a unique phase space reconstruction translates into a focus interval ranging from the far field at underfocus via the in-focus plane to the far field at overfocus. It is furthermore readily verified in phase space that both far fields
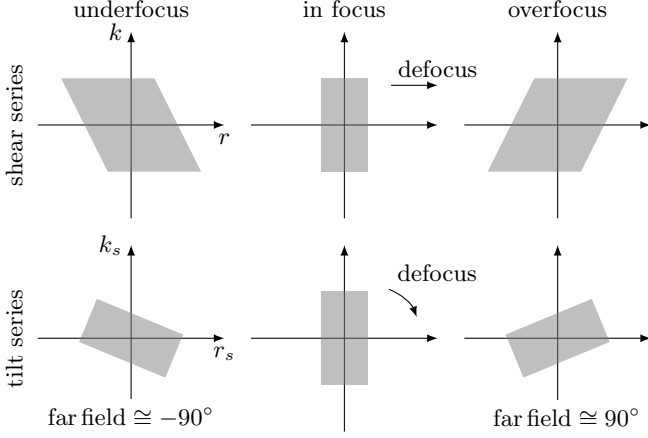
underfocus   in focus   overfocus

Figure 3. Inline holography as quantum state reconstruction. The shear series of the defocused Wigner function is shown in the first row. The corresponding tilt series obtained by suitably scaling the phase space is depicted in the second row.

(at over- and underfocus) may be identified by inverting the spatial coordinates.

As we are dealing with a four-dimensional phase space

in general, the projection geometry is more complicated than the simple single-tilt-axis geometry (where the set of projections is referred to as Radon transform) typically considered in tomography (including two-dimensional phase space tomography). In particular, each individual defocused image is obtained by projecting along a 2D plane in phase space, tilted by the defocus. Such a projection transformation represents a generalization of the 2D Radon transformation, referred to as K-plane transform K_2^4 (e.g., (Keinert 1989)), if the projections are performed over all 2D planes in 4D phase space. To realize arbitrary shears in 4D phase space electron optically, line defoci (where one direction remains in focus) with varying orientation have to be realized.

We will now discuss the scaling of the focal series, i.e., the relation between free space propagation and the experimental images in the series, for imaging with a thin lens. The thick lens case, required further below, is treated in Appendix A. Following Abbe, the object, back focal plane and image plane of a single lens are connected in an ideal (aberration-free) system wave optically by means of a Fourier and an inverse Fourier transformation, respectively. This result may be derived by propagating a wave given at some object distance z_{obj} through a thin lens with focal length f (see Figure 4)

$$
\begin{aligned}
\Psi(\mathbf{r}, z) &= \frac{1}{8\pi^3} \iiint_{-\infty}^{\infty} \mathrm{d}^2 r' \mathrm{d}^2 k' \mathrm{d}^2 k \, \tilde{\Psi}(\mathbf{k}', z_{\mathrm{obj}}) \, e^{-i\frac{z_{\mathrm{obj}}}{2k_0}\mathbf{k}'^2} e^{i(\mathbf{k}'-\mathbf{k})\mathbf{r}'} e^{-i\frac{k_0}{2f}\mathbf{r}'^2} e^{-i\frac{z}{2k_0}\mathbf{k}^2} e^{i\mathbf{k}\mathbf{r}} \quad (10)\\
&= \frac{f}{4\pi^2 i k_0} \iint_{-\infty}^{\infty} \mathrm{d}^2 k \mathrm{d}^2 k' \, \tilde{\Psi}(\mathbf{k}', z_{\mathrm{obj}}) \, e^{-i\frac{z_{\mathrm{obj}}}{2k_0}\mathbf{k}'^2} e^{i\frac{f}{2k_0}(\mathbf{k}-\mathbf{k}')^2} e^{-i\frac{z}{2k_0}\mathbf{k}^2} e^{i\mathbf{k}\mathbf{r}}\\
&= \frac{f}{2\pi(f-z)} e^{-i\frac{k_0}{2(f-z)}\mathbf{r}^2} \int_{-\infty}^{\infty} \mathrm{d}^2 k' e^{\frac{i}{2k_0} \underbrace{\frac{z_{\mathrm{obj}}z - f(z_{\mathrm{obj}}+z)}{f-z}}_{\delta z(z)} \mathbf{k}'^2} \tilde{\Psi}(\mathbf{k}', z_{\mathrm{obj}}) e^{i\underbrace{\frac{f}{f-z}}_{M^{-1}}\mathbf{r}\mathbf{k}'}\\
&= \frac{ik_0}{M\delta z} e^{-i\frac{k_0}{2M}\left(\frac{1}{f}+\frac{1}{\delta z}\right)\mathbf{r}^2} \mathcal{F}^{-1}\left\{ \Psi(\mathbf{r}, 0) e^{-i\frac{k_0}{2\delta z}\mathbf{r}^2} \right\} \left(\frac{k_0}{Mz}\mathbf{r} \right).
\end{aligned}
$$

In the second transformation we defined an effective defocus δz, i.e., the corresponding propagation length in free space, and a magnification M, describing the rescaling with respect to a freely propagated wave. Inserting the wave function propagating through a thin lens into the definition of the Wigner function (2) and taking into account (8), we obtain a succession of shears along the position and the momentum coordinate in phase space as depicted in Figure 4. Again, this shear series may be transformed into a tilt series by suitably rescaling phase space.

It follows that the diffraction pattern $\rho = \left| \tilde{\Psi} \right|^2$ occurs in the back focal plane ($z = f$) because (starting from

the second line in Eq. (10))

$$
\begin{aligned}
\Psi(\mathbf{r}, f) &= \frac{f}{i4\pi^2 k_0} \int_{-\infty}^{\infty} \mathrm{d}^2 k' e^{i\frac{f-z_{\mathrm{obj}}}{2k_0}\mathbf{k}'^2} \tilde{\Psi}(\mathbf{k}', z_{\mathrm{obj}}) \quad (11)\\
&\quad \times \left(\int_{-\infty}^{\infty} \mathrm{d}^2 k \, e^{i\left(\mathbf{r}-\frac{f}{k_0}\mathbf{k}'\right)\mathbf{k}} \right)\\
&= \frac{f}{ik_0} \int_{-\infty}^{\infty} \mathrm{d}^2 k' \delta\left(\mathbf{r} - \frac{f}{k_0}\mathbf{k}' \right) e^{i\frac{f-z_{\mathrm{obj}}}{2k_0}\mathbf{k}'^2} \tilde{\Psi}(\mathbf{k}', z_{\mathrm{obj}})\\
&= \frac{1}{i} e^{i\frac{k_0(f-z_{\mathrm{obj}})}{2f^2}\mathbf{r}^2} \tilde{\Psi}\left(\frac{k_0}{f}\mathbf{r}, z_{\mathrm{obj}} \right).
\end{aligned}
$$

Eq. (10) indicates three ways to change the effective defocus δz (facilitating a focal series) by varying z_{obj}, f or z_{img}, respectively. However, the object plane (i.e., the specimen or some conjugated plane) and the image plane (i.e., the detector plane or some conjugated plane) remain

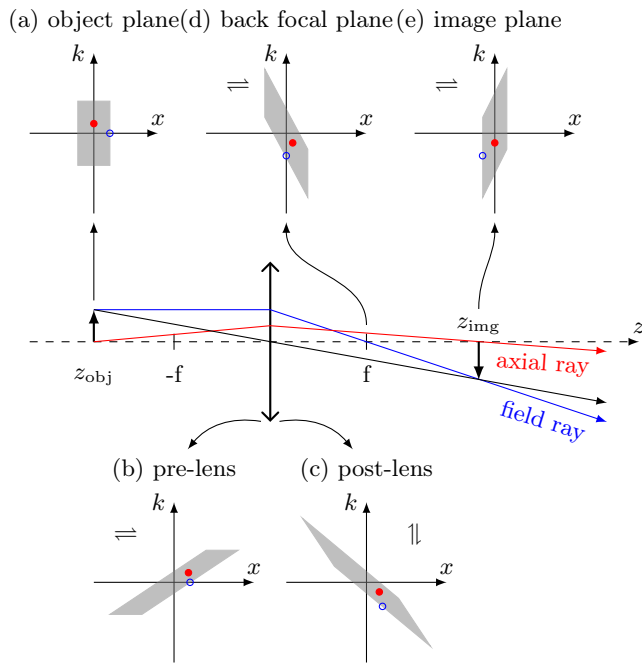(a) object plane (d) back focal plane (e) image plane

Figure 4. Imaging at a thin lens in phase space. Accordingly, the initial rectangular phase space distribution in the object plane (a) passes through several (affine) shear transformations (b)-(d) yielding a 180° rotated and, along the momentum coordinate, sheared version in the image plane (e).

fixed in a typical TEM. Thus, using a single lens, focal series are usually recorded by varying the lens excitation, hence f. It is important to note that by varying the focal length f not only in the effective defocus δz but also in the magnification M changes. This rescaling has to be taken into account, when reconstructing the wave function from an experimental focal series. An accurate calibration of the focal series in terms of defocus and magnification are absolutely crucial therefore. This is a serious challenge and we present a dedicated calibration procedure adapted to the above focal series recorded with the diffraction lens later on.

III. EXPERIMENTAL IMPLEMENTATION

As of today, the recording of a focal series ranging from the far field at underfocus via the focal plane to the far field at overfocus that is free from additional modifications such as magnification changes, rotations, distortions or incoherent aberrations represents a serious challenge. Several authors noted the importance of taking into account these imaging errors in the focal series reconstruction (Koch 2014, Meyer et al. 2004, 2002).

To mitigate some of the noted obstacles, we focus on medium resolution focal series in the following. This allows us to use a low magnification mode, where the focal series is facilitated by changing the excitation of the diffraction lens (Figure 5). In a first step we will treat
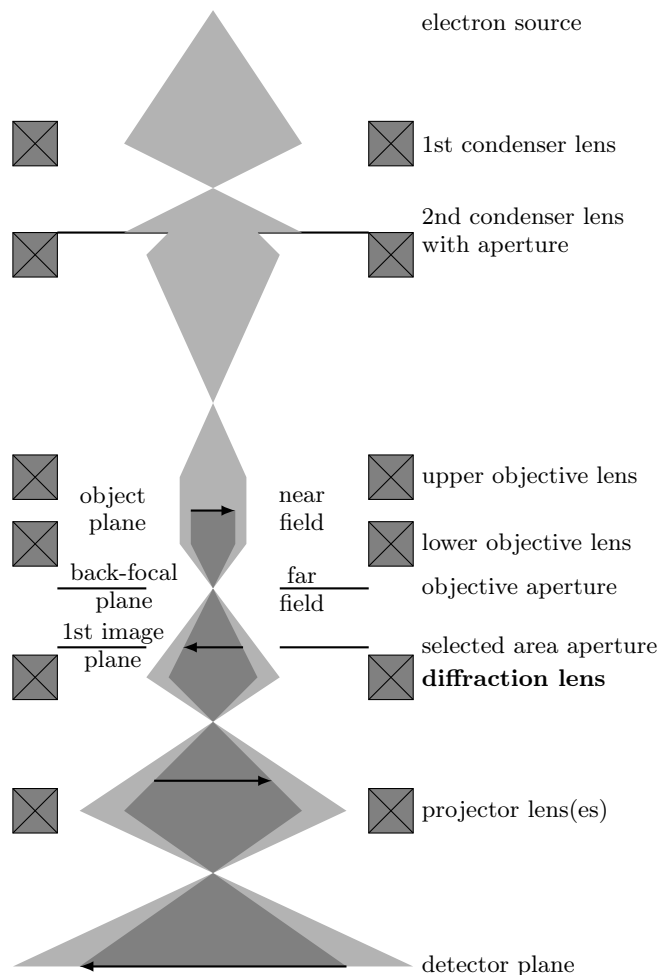

Figure 5. Conventional bright field TEM (CTEM) scheme with the electron source located in the far field (Fourier plane) of the object plane, which in turn is conjugated to the detector plane. Conjugated object planes are indicated by arrows. The medium resolution focal series considered here was recorded by varying the diffraction lens (highlighted) within the low magnification mode of the TEM.

the diffraction lens as a weak (thin) magnetic lens, which greatly simplifies the optics and the derivation of the pertinent equations linking the change of the lens excitation at the microscope with an actual defocus. Based on that, we discuss the crucial calibration of the focal series eventually also considering thick lenses. We illustrate the various steps at the example of a Philips CM200 TEM operated at 200 kV.

A. Weak Magnetic Lens Approximation

In order to calibrate the focal length as well as the object and image plane coordinates, we shall start with the weak lens approximation before discussing a more realistic optical model. For a thin lens the focal lengths and focal planes coincide and the principal planes fall

into the symmetry plane of the lens. The focal length of a thin round magnetic lens is given by the integral of the squared magnetic induction along the optical axis (Hawkes 2013)

$$\frac{1}{f} = \frac{e}{8m_e U^*} \int_{-\infty}^{\infty} B_z^2(z)\, dz \qquad (12)$$

with the relativistically corrected acceleration voltage given by

$$U^* = U\left(1 + \frac{eU}{m_e c^2}\right). \qquad (13)$$

Similarly, the image rotation φ is given by the integral of the induction

$$\varphi = \frac{e}{2m_e v} \int_{-\infty}^{\infty} B_z\, dz \qquad (14)$$

with the electron velocity denoted by v.

To evaluate the above integrals, we need to know the magnetic induction inside of the diffraction lens. In the absence of accurate magnetic field configurations from the manufacturer, we employ the GRAY model ($\alpha = 2.636$) for the magnetic induction in the pole piece (diameter D, gap S)

$$B_z(z) = \frac{\mu_0 I}{2S}\left(\tanh\left(\frac{\alpha S}{2D}\left(1 + 2\frac{z}{S}\right)\right)\right. \qquad (15)$$
$$\left. + \tanh\left(\frac{\alpha S}{2D}\left(1 - 2\frac{z}{S}\right)\right)\right),$$

which provides a rather accurate approximation for weakly excited round magnetic lenses (Hawkes 2013). In the GRAY model the inverse focal length is proportional to the squared lens excitation (lens current I in Ampere-turns) with the proportionality constant C_f

$$\frac{1}{f} = \frac{e\mu_0^2}{8m_e U^{*2} S}\left(\frac{1 + \tanh^2\left(\frac{\alpha S}{2D}\right)}{2\tanh\left(\frac{\alpha S}{2D}\right)} - \frac{D}{\alpha S}\right)I^2 = C_f I^2. \qquad (16)$$

Thus, we only have to determine C_f in our calibration procedure in order to know the focal length f as a function of the lens excitation. The image rotation on the other hand is linearly proportional to I according to

$$\varphi = \frac{e\mu_0 I}{2m_e v}. \qquad (17)$$

We use this dependency to calibrate the excitation strength displayed by the TEM control interface in terms of Ampere-turns by measuring the image rotation at different excitation strengths and solving Eq. (17) for I. Figure 6 exhibits the linear increase of the image rotation with the lens current in the diffraction lens at our Philips CM200 TEM operated at 200 kV. Accordingly, at 100% excitation strength the current amounts to 4904 Ampere-turns.
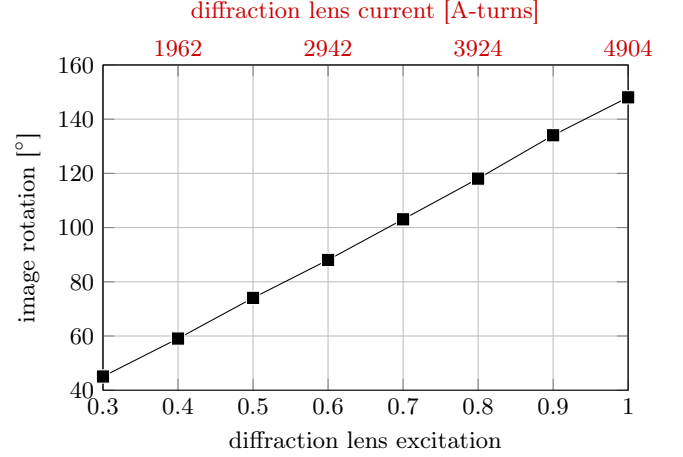

Figure 6. Image rotation as a function of the diffraction lens excitation. Note the linear increase in agreement with Eq. (17).

B. Calibration of Focal Series

The principal part of the calibration procedure consists of three measurements. First, we record two in-focus images of a Selected Area (SA) aperture (radius 50 μm) and a copper grid in the object plane (objective lens switched off) of known size and determine their respective magnifications $M_{1,2}$. These magnifications are composed of the post-magnification of the projector lenses and the actual magnification of the Dif lens according to

$$M_{1,2} = M_{\text{proj}} M_{\text{dif}}^{(1,2)} \qquad (18)$$
$$= M_{\text{proj}} \frac{f_{1,2} - z_{\text{img}}}{f_{1,2}}$$
$$= M_{\text{proj}}\left(1 - C_f I_{1,2}^2 z_{\text{img}}\right),$$

which can be solved for

$$C_f z_{\text{img}} = \frac{M_1 - M_2}{M_1 I_2^2 - M_2 I_1^2} \qquad (19)$$

and the post-magnification of the projector lenses

$$M_{\text{proj}} = \frac{M_{1,2}}{1 - C_f I_{1,2}^2 z_{\text{img}}}. \qquad (20)$$

To calibrate the remaining unknown optical parameters, namely C_f, z_{img} and z_{obj}, we record a reference focal series of a small grating (lattice constant $d = 100$ nm, objective lens switched off) because the corresponding large far field permits the following approximation for a large range of lens excitations

$$\Psi(\mathbf{r}, z) = \frac{1}{2\pi M} e^{-i\frac{k_0}{2(f-z)}\mathbf{r}^2} \qquad (21)$$
$$\times \int d^2 k' e^{\frac{i}{2k_0}\delta z \mathbf{k'}^2} \tilde{\Psi}(\mathbf{k'}, z_{\text{obj}}) e^{iM^{-1}\mathbf{r}\mathbf{k'}}$$
$$\overset{\text{SPA}}{\sim} \tilde{\Psi}\left(\frac{k_0}{M\delta z}\mathbf{r}, z_{\text{obj}}\right).$$
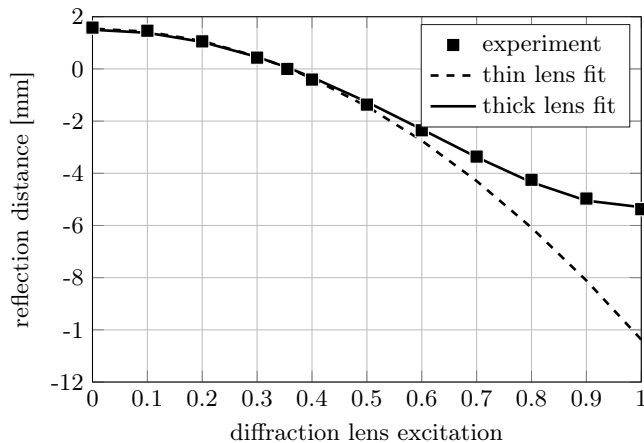
Figure 7. Distance between systematic reflections as a function of the lens excitation

In the last line we used the stationary phase approximation (SPA) assuming an effective defocus large enough to be in the far field of the grating. Consequently, tracking the difference $s = r_n - r_{n-1}$ between adjacent diffraction orders $(q = 2\pi n / d)$

$$s = M_{\text{proj}} \frac{\lambda}{d} \left(\frac{z_{\text{obj}} z_{\text{img}} - f \left(z_{\text{img}} + z_{\text{obj}} \right)}{f} \right) \qquad (22)$$

$$= M_{\text{proj}} \frac{\lambda}{d} \left(z_{\text{obj}} z_{\text{img}} C_f I^2 - z_{\text{img}} - z_{\text{obj}} \right)$$

$$= M_{\text{proj}} \frac{\lambda}{d} \left(z_{\text{obj}} \frac{M_1 - M_2}{M_1 I_2^2 - M_2 I_1^2} I^2 - z_{\text{img}} - z_{\text{obj}} \right)$$

permits the fitting of the unknown parameters z_{obj} and z_{img}. Figure 7 shows the dependency of the distance s to the lens excitation I. Accordingly, we obtain a good agreement with the predicted quadratic behaviour for weak lens excitation only. Combining (22), (19) and (20) we can compute C_f and finally the effective defocus

$$\delta z \left(I \right) = \frac{z_{\text{obj}} z_{\text{img}} - f \left(I \right) \left(z_{\text{obj}} + z_{\text{img}} \right)}{f \left(I \right) - z_{\text{img}}} \qquad (23)$$

and magnification

$$M \left(I \right) = \frac{f \left(I \right) - z_{\text{img}}}{f \left(I \right)} \qquad (24)$$

within a long-range focal series recorded by varying the excitation of the diffraction lens.

Figure 7 exhibits that the propagation of the wave *within* the magnetic field of the lens, i.e., thick lens effects, cannot be neglected anymore for higher excitations. In Appendix A we thus consider the propagation through a thick lens modeled as a sum of two refracting surfaces at a distance

$$D \left(I \right) = C_D I \qquad (25)$$

depending linearly on the lens current, and derive the corresponding quantities measured in the above calibration

parameter	value
z_O	$18.3\,cm$
z_{SA}	$3.6\,cm$
z_{img}	$47.3\,cm$
C_D	$7.9\,\mu m/A$
C_f	$2.8\ (mA)^{-2}/m$
M_{proj}	88

Table I. Table of optical parameters of the diffraction lens in a Philips CM200 TEM operated in a low magnification mode.
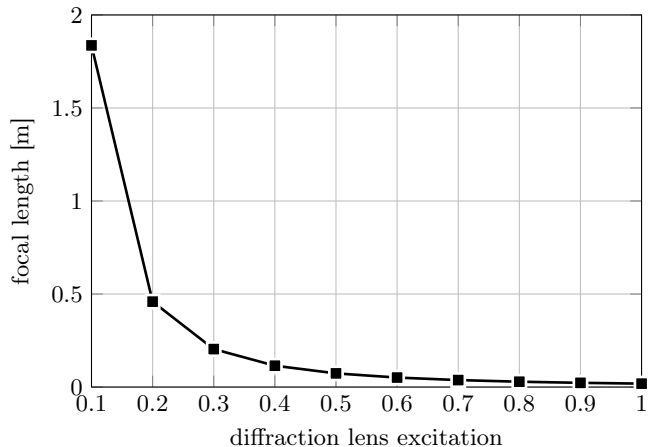

Figure 8. Focal length calibrated from the calibration experiments using the thick lens model for the fitting of the optical parameters.

experiments (see Tab. I). Such a thick lens model yields very good agreement for the distance between the diffraction orders as a function of the lens excitation (see Figure 7). The calibrated optical parameters of our Philips CM200 TEM are in the expected range although estimations of the distance between diffraction lens and object plane suggest an error in the range of 20%. The latter translates to a deviation between the predicted and the experimentally observed position of the in-focus plane in the actual focal series, which has to be taken into account the when discussing the reconstruction results.

We conclude that the characterization of the optical system and hence the parameters of the focal series obtained from the above procedure necessarily represent an approximation because of several assumptions, which might be violated to some extent in reality. For instance, the quadratic dependency of the reciprocal focus and the linear dependency of the thickness of the lens from the lens current (16,25) are expected to break down at higher excitations. Similarly, very weak excitations may be affected by hysteresis effects, disturbing the proportionality between I^2 and f^{-1} as the lens is operated outside of its typical regime. More precise models, however, would require ray tracing simulations within realistic magnetic fields based on the exact geometry of the lenses.
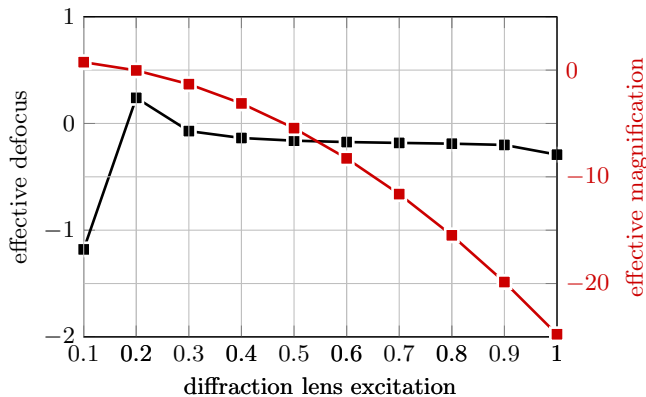
Figure 9. Effective defocus and magnification calibrated from the calibration experiments using the thick lens model for the fitting of the optical parameters.

To account for the inaccurate calibration, we therefore implement a self-consistent adaption of the effective magnification in the focal series reconstruction, similarly to (Koch 2014). Note, however, that such a procedure might also modify the reconstructed wave function and should be only employed to correct for calibration errors. Alternatively, it is possible to fit the effective defocus and magnification by comparing a reference series of some well-defined beam shape (e.g., plane wave through a round aperture) to a simulation.

C. Numerical Propagation Algorithms

When performing a focal series reconstruction from data covering both the near and the far field the two regimes must be treated separately numerically to avoid a spread of the image beyond the simulated domain, while maintaining both a sufficient accuracy and speed in the numerical Fresnel propagation. In the near field regime, the numerical Fresnel propagation is performed in Fourier space exploiting both the small slope of the quadratic Fresnel phase (i.e., slow spreading) and the speed of the Fast Fourier transform:

$$\Psi\left(\mathbf{r}, z\right) = \mathcal{F}^{-1}\left\{e^{i\frac{z}{2k_0}\mathbf{k}^2}\mathcal{F}\left\{\Psi\left(\mathbf{r}, 0\right)\right\}\right\}. \quad (26)$$

When entering the far field regime (characterized by the Fresnel number $F = a^2 / \left(\lambda z\right) \ll 1$ with a the extent of the wave function), the reconstruction algorithm directly evaluates the convolution with the position space Fresnel propagator according to (García et al. 1996)

$$\Psi\left(\mathbf{r}, z\right) = \frac{1}{i\lambda z}e^{i\frac{k_0}{2z}\mathbf{r}^2}\int_{-\infty}^{\infty}\Psi\left(\mathbf{r}', 0\right)e^{i\frac{k_0}{2z}\mathbf{r}'^2}e^{-i\frac{k_0}{z}\mathbf{r}\mathbf{r}'}d^2r' \quad (27)$$

$$= \frac{2\pi}{i\lambda z}e^{i\frac{k_0}{2z}\mathbf{r}^2}\mathcal{F}\left\{\Psi\left(\mathbf{r}, 0\right)e^{i\frac{k_0}{2z}\mathbf{r}^2}\right\}\left(\frac{k_0}{z}\mathbf{r}\right).$$

By separating the different $\mathbf{r}$ and $\mathbf{r}'$ dependencies we can use Fast Fourier transformations for efficiently computing the above convolution. Note, however, that the latter

implicitly rescales the defocused wave function by k_0/z, counteracting the spread upon free propagation. As the rescaling is generally different from the spread of the wave function, the sampling has to be chosen carefully to ensure an adequate scaling. In order to facilitate a matching scale between experimentally recorded images and simulated ones, both the near field and the far field wave functions computed by means of (26) and (27) need to be rescaled with M to fit to the experimental intensity recorded in the corresponding plane.

IV. LONG RANGE FOCAL SERIES RECONSTRUCTION

In the following, we consider the focal series wave reconstruction as a quantum state tomography restricted to pure states (i.e., wave functions). Note, however, that a certain violation of the pure state assumption, e.g., due to inelastic scattering at the sample, the partial coherence of the electron emitter, or the point spread of the detector (Niermann et al. 2012), can never be completely avoided, and must therefore be taken into account therefore (Martin et al. 2006).

In a hardware-corrected TEM or at medium (nanometer) resolution the defocus remains as the main source of aberrations. Thereby, we assume that the effects of temporal incoherence may be sufficiently suppressed due to chromatic aberration correction and energy filtering. In order to compute the impact of a finite source size on the Wigner function in the presence of defocus only, it suffices to incoherently sum the underlying coherent Wigner function over the transverse momentum distribution of the source $\rho_S\left(\mathbf{k}\right)(\cong$ effective source size) to obtain the corresponding mixed quantum state

$$W(\mathbf{r}, \mathbf{k}) = \int_{-\infty}^{\infty}W^{(\text{coh})}(\mathbf{r}, \mathbf{k} - \mathbf{k}')\rho_S\left(\mathbf{k}'\right)d^2k', \quad (28)$$

which is defocused in a second step

$$W(\mathbf{r}, \mathbf{k}, z) = W(\mathbf{r} - \frac{z}{k_0}\mathbf{k}, \mathbf{k}, 0) \quad (29)$$

$$= \int_{-\infty}^{\infty}W^{(\text{coh})}(\mathbf{r} - \frac{z}{k_0}\mathbf{k}, \mathbf{k} - \mathbf{k}', 0)\rho_S\left(\mathbf{k}'\right)d^2k'.$$

This convolution commutes with the projection of the Wigner function along the momentum coordinate (i.e., the measurement) according to

$$\rho\left(\mathbf{r}, z\right) = \int_{-\infty}^{\infty}W(\mathbf{r}, \mathbf{k}, z)d^2k \quad (30)$$

$$= \iint_{-\infty}^{\infty}W^{(\text{coh})}(\mathbf{r} - \frac{z}{k_0}\mathbf{k}, \mathbf{k} - \mathbf{k}', 0)\rho_S\left(\mathbf{k}'\right)d^2kd^2k'$$

$$= \int_{-\infty}^{\infty}\rho^{(\text{coh})}(\mathbf{r} - \frac{z}{k_0}\mathbf{k}', 0)\rho_S\left(\mathbf{k}'\right)d^2k',$$

hence can be inverted (deconvoluted) to a certain extent directly in the recorded image within the limits determined by noise (Koch 2008, Lubk and Röder 2015). From
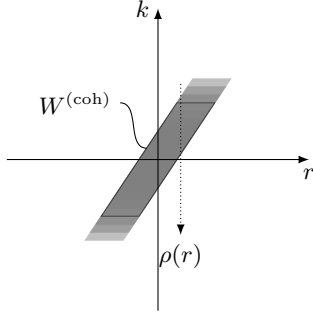
Figure 10. Impact of partial spatial coherence in combination with defocus on the average mixed quantum state. Spatial incoherence in combination with a defocus leads to a convolution in phase space, which may be deconvolved from the recorded image therefore.

a phase space perspective, such a deconvolution corresponds to a purification of the quantum state, as the partial trace over a subspace of the total configuration space (i.e., the incoherent emitter surface) is reverted.

Because of its particular link to fundamental laws of quantum mechanics and its preeminent practical significance, we focus in the following on the archetypical reconstruction algorithm in the field, the Gerchberg-Saxton algorithm. The latter consists of propagating a trial wave function into the various experimental focal planes, where its modulus is iteratively updated to the experimental value. In the limit of an infinitesimally small focal step size, we may derive the following analytical description of this iterative procedure. We start be writing down the

paraxial Klein-Gordon equation in field-free space

$$i\partial_z \Psi(\mathbf{r}, z) = -\frac{1}{2k_0} \Delta \Psi(\mathbf{r}, z), \quad (31)$$

which adequately describes the propagation of an electron beam of wave vector k_0 along the optical axis z in the TEM. The structure of this equation corresponds to a two-dimensional time-dependent Schrödinger equation with z taking the place of the time coordinate. In full analogy to the time-dependent Schrödinger equation one can rewrite (31) as a coupled system of two differential equations for the wave's probability density $\rho = \Psi\Psi^* = A^2$ and the phase φ

$$\frac{\partial \rho(\mathbf{r}, z)}{\partial z} = -\frac{1}{k_0} \nabla \underbrace{(\rho(\mathbf{r}, z) \nabla \varphi(\mathbf{r}, z))}_{\mathbf{j}} \quad (32)$$

$$\frac{\partial \varphi(\mathbf{r}, z)}{\partial z} = -\frac{1}{2k_0} (\nabla \varphi(\mathbf{r}, z))^2 + \underbrace{\frac{1}{2k_0} \frac{\triangle A(\mathbf{r}, z)}{A(\mathbf{r}, z)}}_{\text{quantum potential}}. \quad (33)$$

The first line equates the change of the probability density (or intensity) along z with the lateral divergence of the probability current $\mathbf{j}$. Consequently, it represents the continuity equation for the paraxial regime. In the context of the corresponding holographic technique (Lubk et al. 2013b, Teague 1983), Eq. (32) is also referred to as Transport of Intensity Equation (TIE). The second equation is the quantum version of the paraxial Hamilton-Jacobi equation (QHJE) differing from the classical one by the so-called quantum potential.

If we now start the Gerchberg-Saxton iteration with a deliberately chosen guess for the phase $\Psi_0 (z = 0) = \sqrt{\rho(\mathbf{r}, 0)} \exp(i\varphi_0(\mathbf{r}, 0))$ at the initial plane $z = 0$, the propagation over an infinitesimal small defocus step δz at some point $z = n\delta z$ along the optical axis reads

$$\Psi_n(\mathbf{r}, z + \delta z) = \left(1 + \frac{i\delta z}{2k_0} \Delta\right) \Psi_n(\mathbf{r}, z) \quad (34)$$

$$= \left(1 - \frac{\delta z}{2k_0} \left(\frac{\nabla A(\mathbf{r}, z)\nabla \varphi_n(\mathbf{r}, z)}{A(\mathbf{r}, z)} - i\frac{\Delta A(\mathbf{r}, z)}{A(\mathbf{r}, z)} + i (\nabla \varphi_n(\mathbf{r}, z))^2\right)\right) A(\mathbf{r}, z) e^{i\varphi_n(\mathbf{r}, z)}$$

$$= e^{-\frac{\delta z}{2k_0}\left(\frac{\nabla A(\mathbf{r}, z)\nabla \varphi_n(\mathbf{r}, z)}{A(\mathbf{r}, z)} - i\frac{\Delta A(\mathbf{r}, z)}{A(\mathbf{r}, z)} + i(\nabla \varphi_n(\mathbf{r}, z))^2\right)} A(\mathbf{r}, z) e^{i\varphi_n(\mathbf{r}, z)}.$$

In the last line, we used the smallness of the focal step δz to write the expression appearing in the bracket of the second line as an exponential. Note that this transformation is only possible if the amplitude $A \neq 0$. This reflects the fact that a large relative change of the wave function may occur upon small propagation steps at points, where the amplitude is zero. The Gerchberg-Saxton algorithm proceeds by replacing the amplitude in the plane $z + \delta z$ with the experimental amplitude

$A(\mathbf{r}, z + dz) = \sqrt{\rho(\mathbf{r}, z + \delta z)}$ in that plane

$$\Psi_{n+1}(\mathbf{r}, z + \delta z) = e^{-i\frac{\delta z}{2k_0}\left(\frac{\Delta A(\mathbf{r}, z)}{A(\mathbf{r}, z)} - (\nabla \varphi_n(\mathbf{r}, z))^2\right)} \quad (35)$$

$$\times A(\mathbf{r}, z + \delta z) e^{i\varphi_n(\mathbf{r}, z)}.$$

Consequently, the iterated phase reads

$$\varphi_{n+1}(\mathbf{r}, z + \delta z) = \varphi_n(\mathbf{r}, z) \tag{36}$$
$$- \frac{\delta z}{2k_0} \left(\frac{\Delta A(\mathbf{r}, z)}{A(\mathbf{r}, z)} - (\nabla \varphi_n(\mathbf{r}, z))^2 \right)$$
$$= \varphi_n(\mathbf{r}, z) + \frac{\partial \varphi_n(\mathbf{r}, z)}{\partial z} \delta z ,$$

where we inserted (33) in the last line.

The last equation reveals that the ordered Gerchberg-Saxton iteration with small step size corresponds to the (numerical) integration of the quantum Hamilton-Jacobi equation starting from some deliberately chosen initial wave function. Thus, the Gerchberg-Saxton algorithm may be considered as conjugate to the phase retrieval based on solving the Transport of Intensity Equation as the coupled system of TIE and QHJE is equivalent to the paraxial Klein-Gordon equation (31). Noting that any starting guess φ_0 yields a solution to the above integration, reconstructing the correct wave function underlying the observed intensities in the various focal planes corresponds to single out the correct starting guess. The latter may be accomplished by starting the integration at the far field at underfocus and integrating until reaching the far field at overfocus, because only the correct solution(s) may be identified in the far field at overfocus with the corresponding starting guess at underfocus after inversion of spatial coordinates. In other words, the far field data is necessary to identify the self-consistent solution(s) (and the pertaining periodic Bohm trajectories forming the characteristics of the quantum Hamilton-Jacobi equation).

Relating the Gerchberg-Saxton algorithm with the numerical integration of the (quantum) Hamilton-Jacobi equation allows us to use the comprehensive Hamilton-Jacobi theory for characterizing the algorithm. Within the scope of this work, however, we are content with finding an alternative justification for the long focal range and leave questions, such as how many self-consistent solutions, i.e., how many wave function possess the same intensities in a long-range isotropic focal series to future work. Nevertheless, we will encounter an example below, where at least two different wave functions (i.e., not differing by a mere phase offset only) may not be distinguished from their intensities in isotropic focal planes.

In practice, the focal step employed in a focal series is too large to allow for a linear approximation of the propagation as in (34). Therefore, the full (Fresnel) propagator has to be employed to propagate the wave function from focal plane to focal plane. Under these circumstances, the direct interpretation of the algorithm as numerical integrator of the quantum Hamilton Jacobi equation is violated to a certain extent and one may not prove in a rigorous way anymore that the true solution can be obtained by exploring the space of starting conditions. In fact, the larger defocus steps cause the Gerchberg-Saxton iteration to deviate from the QHJ solution pertaining to one particular starting wave by meandering in a complicated way through QHJE solutions corresponding to different starting waves while iterating. To get more insight into this convergence behaviour we first note the following bound for the distance between iterated wave functions

$$\|\Psi_n(\mathbf{r}, z) - \Psi_{n-1}(\mathbf{r}, z)\|^2 = \left\| A_n(\mathbf{r}, z) e^{i\varphi_n(\mathbf{r}, z)} - A(\mathbf{r}, z) e^{i\varphi_n(\mathbf{r}, z)} \right\|^2 \tag{37}$$
$$= \int A_n^2(\mathbf{r}, z) + A^2(\mathbf{r}, z) - 2A_2(\mathbf{r}, z) A(\mathbf{r}, z) \, \mathrm{d}^2 r$$
$$\leq \int A_n^2(\mathbf{r}, z) + A^2(\mathbf{r}, z) - 2A_2(\mathbf{r}, z) A(\mathbf{r}, z) \cos(\varphi_n(\mathbf{r}, z) - \varphi(\mathbf{r}, z)) \, \mathrm{d}^2 r$$
$$= \left\| A_n(\mathbf{r}, z) e^{i\varphi_n(\mathbf{r}, z)} - A(\mathbf{r}, z) e^{i\varphi(\mathbf{r}, z)} \right\|^2 ,$$

which shows that one iteration step corresponds to an orthogonal projection on the closest representative from a set of wave functions (with arbitrary φ) sharing the same modulus A with the experimental value in that plane (Fienup 1982). Consequently, the Gerchberg-Saxton algorithm can be considered as an iterative projection algorithm as used in convex optimization, with the crucial difference between both settings being the non-convexity of the set of wave functions sharing the same modulus in some optical plane (see Figure 11). This feature prevents the transfer of existence and convergence theorems from the convex setting to the typical long-focal step

Gerchberg-Saxton iteration (cf. Figure 12 for a visual representation of the problem). Accordingly, the algorithm may fail to converge even if a consistent solution to the focal series exists (which is typically not the case due to noise, partial coherence, etc.). Note furthermore that the solution depends on the initial guess in the underdetermined case (corresponding to incomplete focal series typically recorded in the TEM) even in the convex setting. These two characteristic properties again show the susceptibility of the Gerchberg-Saxton algorithm with respect to the starting guess, which has to be taken into account in the focal series reconstruction.
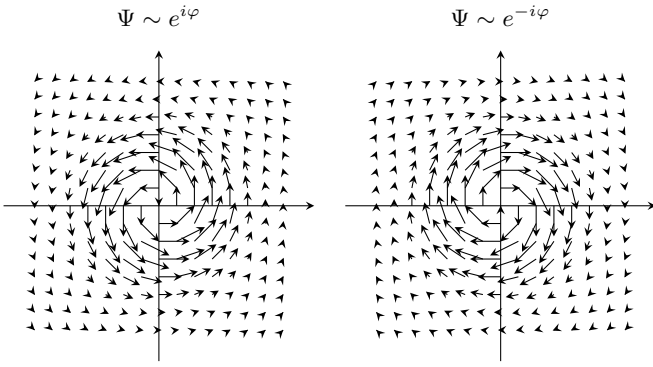
Figure 11. Polya plots of two vortex wave functions with the same amplitude $A = |\Psi|$, where each vector corresponds to a complex number in the complex plane. The complete set of wave functions with the same amplitude is obtained by arbitrarily changing the orientation of the vectors. Accordingly, linear combinations of wave functions from this set, $\Psi = \sum_i c_i \Psi_i$ with $\sum c_i = 1$, generally do not belong to this set. Consequently, the set of wave functions with the same modulus is not convex.
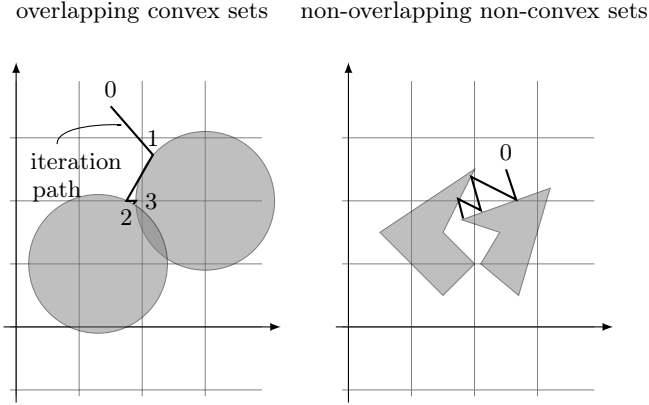

Figure 12. Projection on overlapping convex versus non-overlapping (i.e., inconsistent) non-convex sets. As long as the convex sets are overlapping the iterative projection converges to the intersection of the sets closest to the initial guess of the iteration in the here depicted under-determined case. In the non-convex setting the iteration may get trapped at points of close distance between the sets, not reaching other (and possibly better) solutions.

In spite of the above noted complicated convergence behaviour, the large number of sensible solutions, which have been obtained previously, shows that increased focal step sizes do not necessarily invalidate the reconstruction algorithm as such. Indeed, various strategies and constraints, mainly derived from the phase space perspective, are suited to stabilize the algorithm to a certain extend as will be discussed below.

(A) First of all, we note that the phase space perspective is intimately connected to matrix completion strategies (Candès et al. 2013), which have been successfully employed to transfer projections on non-convex sets

to underdetermined projections on convex sets, recently. The crucial point here is that the set of phase space distributions possessing the same projection in some direction is convex, which allows us to lift the original Gerchberg-Saxton projection on non-convex sets of complex wave functions to convex sets of phase space distributions pertaining to the same modulus in some plane. However, the mixed state formulation of through focus reconstructions (i.e., quantum state tomography) discussed above revealed that a set of line foci ranging from the far field at underfocus via the in-focus plane to the far field at overfocus and comprising all possible orientations of the line focus is necessary to obtain a unique quantum state reconstruction. Presently used stigmators do not permit the acquisition of a line focal series from the near to the far field, however. Thus, focal series typically comprise only isotropically defocused images, although it has been noted that the use of line foci in the near field reduces the ambiguity of the reconstruction (Henderson et al. 2009, Petersen and Keast 2007). Furthermore, in spite of significant progress being achieved in the fields of image registration (Meyer et al. 2002, Saxton 1994), aberration correction and aberration assessment (Meyer et al. 2002), it remains a formidable challenge to record and register a long range isotropic defocus series free from spurious aberrations, distortions, rotations and magnification changes. To mitigate these issues focal series are typically recorded in or close to the near field of the object taking into account that the limited focal range limits the focal series reconstruction of low spatial frequencies (Haigh et al. 2013, Niermann and Lehmann 2016). More recently, also non-linear defocus variations extending further into the far field have been employed successfully in various inline holography studies (Haigh et al. 2013, Koch 2014, Song et al. 2013) to increase the reconstructed spatial frequency band. Because equal tilt intervals in phase space correspond to non-equally spaced focal steps a non-linear sampling of the defocus

$$
\begin{aligned}
\delta z\left(\alpha\right) & = \frac{k_0}{k_\sigma^2}\left(\tan\left(\alpha + \delta\alpha\right) - \tan\alpha\right) \qquad (38) \\
& \overset{\delta\alpha \ll 1}{\approx} \frac{2k_0}{k_\sigma^2}\frac{1}{\cos\left(2\alpha\right) + 1}\delta\alpha
\end{aligned}
$$

decreasing towards the far field is optimal in these studies if seeking a minimal number of images in the focal series. Here k_σ corresponds to the characteristic width of the wave function in Fourier space.

(B) To fulfill the support theorem of tomography (Helgason 2011), the complete electron wave to be reconstructed should be contained within the field of view throughout the entire tilt series. In practice, however, focal series reconstructions, in particular at the atomic resolution regime, are frequently carried out for electron beams being larger than the recorded field of view, hence violating the tomographic support theorem. It has been noted that the corresponding artifacts in the reconstruction may be mitigated to some extend by numerically padding the intensity images with zeros, i.e., artificially

introducing a boundary to the beam (Lin et al. 2006, Ophus and Ewalds 2012).

(C) The alternating projection on convex sets in the phase space setting corresponds to the Kaczmarz (Kaczmarz 1937) (or Algebraic Reconstruction Technique (Gordon et al. 1970)) algorithm in the parlance of tomography (Wei 2015). This analogy also suggests that a non-convergence of the Gerchberg-Saxton algorithm due to ubiquitous inconsistencies in the recorded data, may be mitigated by combining all Kaczmarz iterations within one cycle, i.e, by projecting on all focal planes simultaneously. This strategy is an adaption of the Landweber algorithm or Sequential Iterative Reconstruction Technique (SIRT) well-known from tomography (Natterer and Wübbeling 2001) and forms the basis behind Allen's improved Gerchberg-Saxton algorithm (Allen et al. 2004). Figure 13 shows a modified version of the latter, including the previously mentioned iteration over the starting guess and some additional processing step, which is used in the course of this work. Various TEM studies employed similar variants of the Gerchberg-Saxton algorithm and a commercial version of Allen's IFWR algorithm is available (Allen et al. 2004, Allen and Oxley 2001, Koch 2008, 2014).

The above lines show that typically recorded focal series are not sufficient to guarantee a unique and error-free reconstruction, which requires additional care in the interpretation of the results, in particular because error estimates are less straight-forward compared to off-axis holography. Therefore, the application of additional assumptions restricting the set of possible quantum states can be very useful to further improve the quality of the focal series reconstructions. Which assumptions can be used and how they may be incorporated into efficient focal series reconstruction algorithms represents an active area of research, bearing close analogies to the regularization strategies discussed within the context of tomography and compressive sensing in general. In the following a short overview of possible strategies is given.

We first note that each step of the Gerchberg-Saxton algorithm does not modify the topology of a wave function, which is given by the total winding number

$$w = \frac{1}{2\pi} \oint_{\partial D} \mathrm{d}\mathbf{s} \cdot \nabla \varphi \qquad (39)$$

computed on the outer boundary ∂D of the electron beam because only the amplitude is updated in the differently defocused planes. Thus, the topology of the starting guess for the wave function is preserved throughout the iteration (Martin and Allen 2007) and it is important to endow the starting guess with the correct topology to ensure convergence to the true solution. Unless working with electron vortex beams (Lubk et al. 2013b, McMorran et al. 2011, Verbeeck et al. 2010) the wave functions making up a conventional TEM beam posses a trivial topology (winding number equals zero), because elastic scattering at the object does not change the topology of the initial wave function. Consequently, a wave of trivial
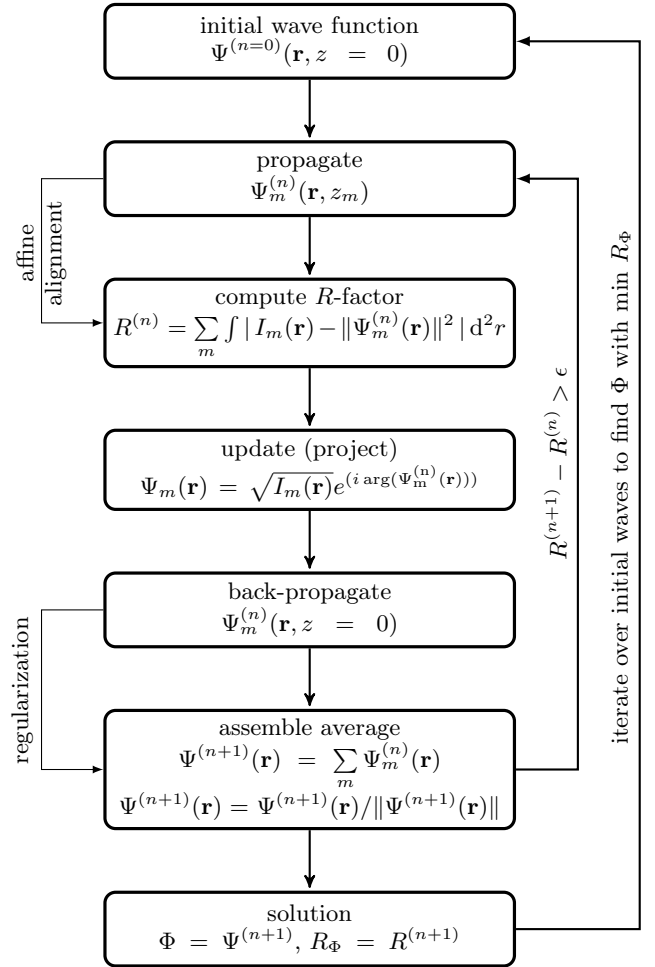

Figure 13. Modified Gerchberg-Saxton algorithm, where the wave functions Ψ_m in the various focal planes (index m) are computed simultaneously by propagating the actual iterated wave function (iteration index n) from some predefined (focal) plane (e.g., $z = 0$). The affine alignment of the recorded intensities with respect to the simulated wave functions may optionally be inserted to correct for spurious image rotations, magnification changes, distortions, etc. The regularization before computing the updated wave function may be applied to impose additional constraints, e.g., on the smoothness (regularity) of the wave. The iteration stops if some convergence criteria (e.g., R-factor limit ϵ or maximal iteration number) is reached.

topology has to be used as starting guess under these typical circumstances. This implicit restriction of the solution space present in almost all focal series reconstruction from TEM images reported to date of this publication is one of the main reasons for reasonable reconstructions from restricted focal ranges. In case of an unknown wave topology, it may also be possible to single out the correct one by comparing the R-factors of the reconstructions pertaining to different topologies. Note, however, that such a test requires well-converged and unique solutions for the different topologies, in order to ensure that the inconsistencies due to an erroneous topology are not over-

shadowed by partial coherence, alignment issues, noise, etc.

Additional assumptions may pertain to a possible sparsity of the wave function in some basis, possible smoothness (i.e., regularity) restrictions (Allen et al. 2004, Parvizi et al. 2016), support constraints of the wave function (Latychevskaia et al. 2010) or positivity of the phase shift. If inconsistent with the underlying wave function, these constraints introduce additional regularization errors (e.g., adding large gradients in the wave), which need to be balanced against the dampened reconstruction error.

V. CASE STUDY

In the following we perform an exemplary focal series reconstruction of a higher order vortex beam, to study the impact of the above principles, such as including the far field images or predefining the topology of the starting guess. The electron vortex beam has been created by inserting a forked hologram grating (Grillo et al. 2014, McMorran et al. 2011, Verbeeck et al. 2010) (diameter= $10\,\mu m$) into the condenser aperture of a FEI Titan^3 TEM operated at 300 kV. Such masks produce a set of vortex beams of increasing order corresponding to the various diffraction orders of the grating (Hecht 1987). Here, we cut out one of the $|w| = 3$ sidebands with a square aperture (side length= $2\,\mu m$) in the selected area plane. The focal series was recorded by varying the excitation of the diffraction lens (see (Clark et al. 2016) for details). A total number of 20 defocused images was recorded by changing the excitation of the diffraction lens from 32% to 60%. The calibration of the effective propagation length and magnification was conducted with the above methods. To further improve the accuracy of the calibration, we employed a comparison with a simulated reference series (see (Clark et al. 2016) for details). In spite of using a large range of diffraction lens excitations, the defocus values ranged from 0 (in-focus) to 0.14 m in the far field at overfocus only. Thus, 90° of the associated phase space tilt series are missing, which translates into considerable missing information and needs to be taken into account when evaluating the results.

To register the images, we aligned their center of mass and image rotation with a non-linear fitting procedure prior to the focal series reconstruction. Figure 14 shows a subset of registered images from the series used for the actual reconstruction. Subsequently, we carried out an initial reconstruction with the help of the algorithm depicted in Figure 13 using a starting wave with $w = 3$ topology, where we performed an image registration step after each 100 steps to remove spurious magnification and rotation errors in the experimental data. Figure 14 exhibits the convergence behaviour of that run. The resultant series was subsequently used in the following reconstructions, where 500 iterations proved sufficient to ensure a stable solution, when incorporating a small defocus into the starting guess.

In the middle row of Figure 14, we compare reconstructions pertaining to different starting waves with the help of their R-factor (see Figure 13). Accordingly, the solutions fall into different equivalence classes distinguished by the topology of the starting wave. This behaviour is due to the preservation of topology within each Gerchberg-Saxton iteration, as noted previously. In full agreement with the imprinted topology due to the fork aperture, the $w = 3$ reconstruction yields the best match between experiment and reconstruction. The phase vortices appear separated in the reconstruction, which is a consequence of small manufacturing faults of the fork aperture and a small misalignment of the square aperture, which slightly breaks the 4-fold symmetry (Lubk et al. 2013a). As a consequence of the only slightly broken four-fold symmetry, however, the $w = -3$ solution may not be ruled out definitely. To clearly distinguish between both solutions an astigmatic focal series would be required.

In the following we use the correct $w = 3$ topology to elaborate on various characteristics of the focal series reconstructions. The above reconstructions have been obtained without constraining the regularity (i.e., band limit). Figure 15 shows two reconstructions with increasing band limits imposed by multiplying a Gaussian low pass in each iteration cycle, where the left column in Figure 15 corresponds to the right column in Figure 14. In particular, if no regularity constraint is applied (maximal spatial resolution), one observes several vortex-antivortex pairs at the aperture corners. Their appearance is a hallmark of the above noted instability of the Gerchberg-Saxton iteration at amplitude zeros, which get more pronounced at higher resolutions progressively resolving the interference phenomena at the aperture edges. Implementing a regularity constraint allows us to smooth out these interference effects, hence minimizing the number of vortices. This effect of the low-pass filter resembles that of partial coherence smearing out the zeros in the recorded intensity (Allen et al. 2004). Reducing the band limit beyond a certain limit eventually severely degrades the reconstruction quality at the aperture edges, hence increasing the R-factor. The observed variation in the actual vortex positions indicates susceptibility of the vortices to the iteration path and small changes in the reconstructed amplitude.

Figure 16 displays reconstructions obtained from a variation of different starting conditions. Again, the left column displays the best fitting solution from Figure 14 for comparison. The other two solutions exhibit the susceptibility with respect to the initial defocus of the starting wave, eventually leading to an expulsion of vortex lines out from the ring-shaped intensity maximum, which may not be reversed, even if extending the number of iterations. On top of that, additional vortex-antivortex pairs appear. These observations highlights the importance of singling out a suitable starting wave in the focal series reconstruction. In the absence of previous knowledge one
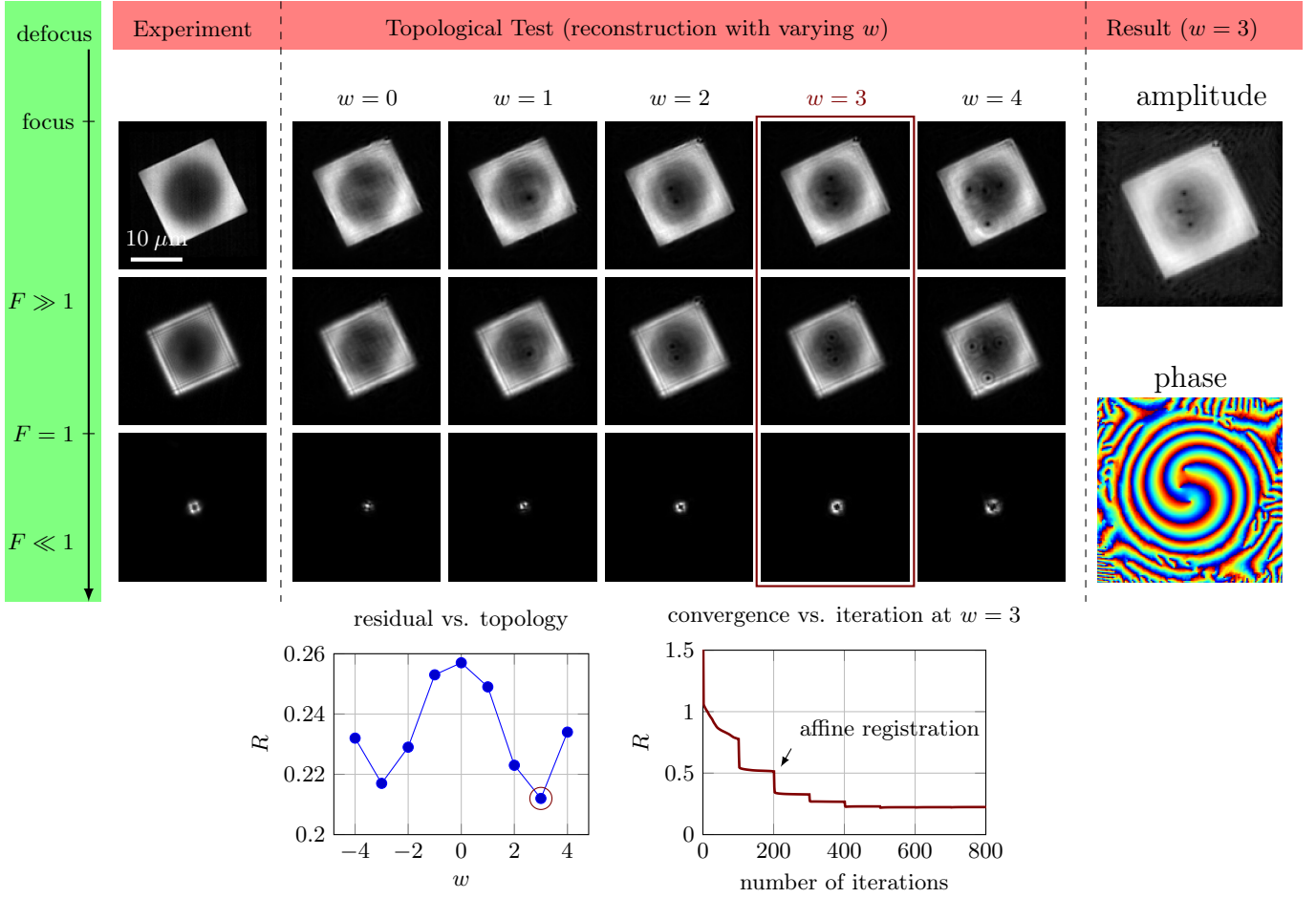
Figure 14. Focal series reconstruction of the vortex beam. Left column: Three recorded densities from the defocus series. Central column: Solutions to $w = 0, 1, 2, 3, 4$ vortices with $w = 3$ fitting best to the experimental data according to R-factor plotted for the different topologies (highlighted by small red circle). The convergence of the $w = 3$ solution exhibits large improvements after the first few affine image registration steps correcting for spurious misalignments of the magnification and rotation. Right column: Amplitude and phase in the aperture plane (selected area plane) pertaining to the $w = 3$ solution.
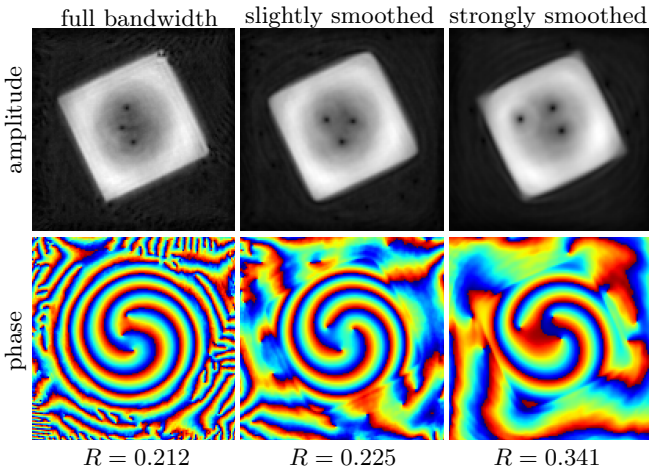

Figure 15. Focal series reconstruction of the vortex beam using different regularity constraints (band limits).

may employ computationally demanding brute force testing or more advanced strategies, such as employing TIE reconstructions from a set of slightly defocused images close to the starting plane of the iteration.

Figure 17 displays reconstructions obtained from truncated focal series, where an increasing number of images from the far field has been removed. As usual, the left column contains the original solution for comparison. Upon removing far field images from the focal series data, one observes almost no degradation of the quality of the reconstruction, in spite of the previously mentioned problematic reconstruction of small spatial frequencies in case of restricted defocus intervals. This result demonstrates that the conservation of topology implicitly restricts the solution space of the Gerchberg-Saxton algorithm even in the presence of an incomplete isotropic focal series. However, the missing far field information eventually becomes obvious when comparing the R-factors pertaining to different topology classes (Figure 17). Accordingly, the different topologies may not be distinguished by their R-
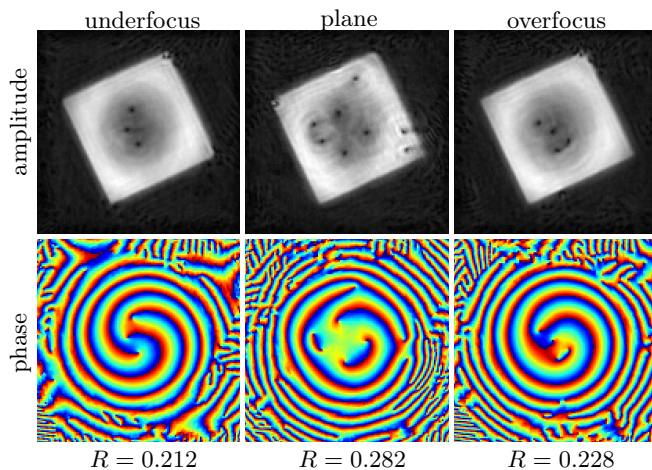
Figure 16. Reconstructions of the vortex beam using different starting conditions, namely a small defocus, no defocus and a small overfocus of the initial wave function.
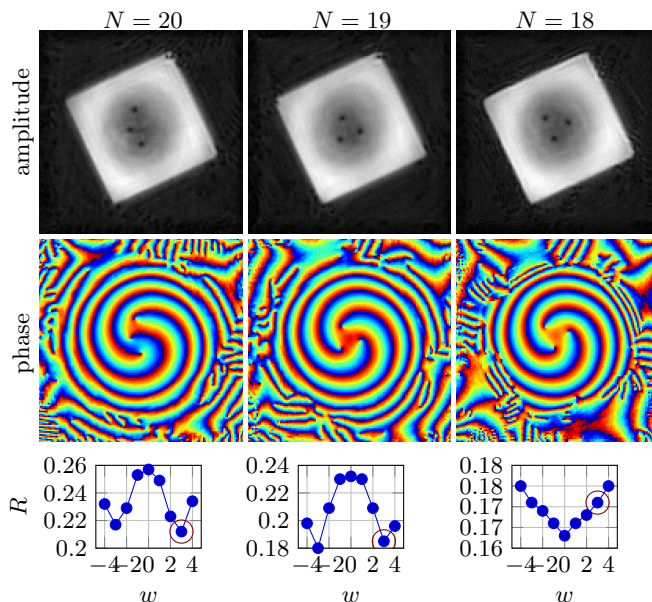

Figure 17. Reconstructions of the vortex beam removing the number of images in the focal series starting at the far field.

factors anymore, if the last two mostly defocused images have been removed in our case.

VI. SUMMARY AND OUTLOOK

In summary, we elaborated on the theoretical foundations as well as the algorithmic and experimental implementation of focal series inline holography. Employing the correspondence between focal series phase retrieval and tomographic reconstruction of the corresponding Wigner function, we showed that a complete reconstruction may only be obtained from a long-range focal series ranging from the far field at underfocus via the in-focus image to the far-field at overfocus. Moreover, the electron beam has to be completely contained within the field of view throughout the focal series. Typically, the focal ranges employed in focal series reconstructions in the TEM do not comply with these principles. In combination with the ubiquitous experimental noise and the violation of the pure state assumption due to partial coherence, reconstruction errors due to incompleteness (in particular at low spatial frequencies) and inconsistencies (in particular at large spatial frequencies) must, therefore, be taken into account.

To reduce reconstruction errors introduced by inaccurate focal values and magnifications, we developed a calibration scheme for the diffraction lens (used here to facilitate the focal series), which may be adapted to other lenses in a straight-forward manner. To improve the calibration and hence the quality of the reconstruction, more accurate wave propagation schemes through the extended magnetic fields of the lenses are required, however. In combination with more open information policies of the microscope manufacturers that could eventually facilitate the precise modeling and calibration of a whole TEM's optics including lenses, deflectors and apertures. Such a virtual TEM would provide the means for a rapid design of novel optical setups (e.g., split illumination) or the automated control of the TEM (fast focal series acquisition). In the particular case of the focal series reconstruction, it would be possible to identify optimal combinations of multiple lens excitations for a complete far field at underfocus - in focus - far field at overfocus focal series, presently not available within the TEM.

Subsequently, the archetypal Gerchberg-Saxton iterative projection algorithm has been identified as a numerical integration of the quantum Hamilton-Jacobi equation in the limit of small focus increments. It is readily derived that the dependency of the integration on the starting wave transfers to the Gerchberg-Saxton algorithm. Again, the far-field data are crucial to single out the correct starting wave in the iteration because incorrect starting waves in the far-field at underfocus are not matched by the corresponding quantum Hamilton-Jacobi solution at the far field at overfocus. More advanced implications following from the analogy between Gerchberg-Saxton and quantum Hamilton-Jacobi remain to be explored in the future.

We further showed that the topology of the starting guess is conserved throughout the Gerchberg-Saxton iteration. Hence, the solutions pertaining to different starting conditions may be separated into topology classes. This implicit constraint mitigates incompleteness problem, if knowing the topology of the solution prior to the reconstruction. This is usually the case as the typical topology of a TEM beam is the trivial one. Moreover, we showed that the different topology classes may be distinguished with the help of their reconstruction quality, if the focal series comprises sufficient images in the far field. The latter facilitates the topological characterization of

beams in singular electron optics, which can serve as a fingerprint for the orbital angular momentum of vortex beams, even if the overall quality of the reconstruction remains low (e.g., due to low signal-to-noise).

We conclude by noting that one of the biggest issues of the above focal series reconstructions was the presence of partial coherence, violating the pure wave function assumption used throughout. The implementation of a quantum state reconstruction from the focal series, i.e., quantum state tomography can resolve that problem facilitating the reconstruction of intrinsically incoherent beams such as resulting from inelastic scattering (e.g., at

plasmons).

Allen, L. J., McBride, W., O'Leary, N. L., and Oxley, M. P. (2004). Exit wave reconstruction at atomic resolution. *Ultramicroscopy*, 100:91–104.

Allen, L. J. and Oxley, M. P. (2001). Phase retrieval from series of images obtained by defocus variation. *Optics Communications*, 199:65.

Almeida, L. B. (1994). The fractional fourier transform and time-frequency representations. *IEEE Transactions on Signal Processing*, 42(11):3084–3091.

Baba, N. and Mutoh, K. (2001). Measurement of telescope aberrations through atmospheric turbulence by use of phase diversity. *Appl. Opt.*, 40(4):544–552.

Bauschke, H. H., Combettes, P. L., and Luke, D. R. (2002). Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization. *J. Opt. Soc. Am. A*, 19(7):1334–1345.

Boucher, R. H. (1980). Convergence of algorithms for phase retrieval from two intensity distributions. *Proc. SPIE*, 0231:130–141.

Breitenbach, G., Schiller, S., and Mlynek, J. (1997). Measurement of the quantum states of squeezed light. *Nature*, 387(6632):471–475.

Burrows, C. J. (1991). Hubble space telescope optics status. In *Proc. SPIE*, volume 1567, pages 284–293.

Burrows, C. J., Holtzman, J. A., Faber, S. M., Bely, P. Y., Hasan, H., Lynds, C. R., and Schroeder, D. (1991). The imaging performance of the hubble space telescope. *Astrophysical Journal*, 369:L21–L25.

Candès, E. J., Eldar, Y. C., Strohmer, T., and Voroninski, V. (2013). Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225.

Cederquist, J. N., Fienup, J. R., Wackerman, C. C., Robinson, S. R., and Kryskowski, D. (1989). Wave-front phase estimation from fourier intensity measurements. *J. Opt. Soc. Am. A*, 6(7):1020–1026.

Clark, L., Guzzinati, G., Béché, A., Lubk, A., and Verbeeck, J. (2016). Symmetry-constrained electron vortex propagation. *Phys. Rev. A*, 93:063840.

Combettes, P. L. (1996). The convex feasibility problem in image recovery. *Advances in Imaging and Electron Physics*, 95:155–270.

Combettes, P. L. and Trussell, H. J. (1990). Method of successive projections for finding a common point of sets in metric spaces. *Journal of Optimization Theory and Applications*, 67(3):487–507.

Dietrich, J., Abou-Ras, D., Schmidt, S. S., Rissom, T., Unold, T., Cojocaru-Mirédin, O., Niermann, T., Lehmann, M., Koch, C. T., and Boit, C. (2014). Origins of electrostatic potential wells at dislocations in polycrystalline Cu(In,Ga)Se2 thin films. *Journal of Applied Physics*, 115(10).

Dong, B.-Z., Zhang, Y., Gu, B.-Y., and Yang, G.-Z. (1997). Numerical investigation of phase retrieval in a fractional fourier transform. *J. Opt. Soc. Am. A*, 14(10):2709–2714.

Dunin-Borkowski, R., McCartney, M., and Smith, D. (2004). Electron holography of nanostructured materials. In *Encyclopedia of Nanoscience and Nanotechnology*, volume 3, pages 41–100. American Scientific Publishers.

Fienup, J. R. (1982). Phase retrieval algorithms - a comparison. *Appl. Optics*, 21:2758–2769.

Fienup, J. R., Marron, J. C., Schulz, T. J., and Seldin, J. H. (1993). Hubble space telescope characterized by using phase-retrieval algorithms. *Appl. Opt.*, 32(10):1747–1767.

Fienup, J. R. and Wackerman, C. C. (1986). Phase-retrieval stagnation problems and solutions. *J. Opt. Soc. Am. A*, 3(11):1897–1907.

Gabor, D. (1948). A new microscopic principle. *Nature*, 161:777.

García, J., Mas, D., and Dorsch, R. G. (1996). Fractional-fourier-transform calculation through the fast-fourier-transform algorithm. *Appl. Opt.*, 35(35):7013–7018.

Gerchberg, R. W. and Saxton, W. O. (1972). A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35(2):237–246.

Gonsalves, R. A. (1976). Phase retrieval from modulus data. *J. Opt. Soc. Am.*, 66(9):961–964.

Gordon, R., Bender, R., and Herman, G. T. (1970). Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and x-ray photography. *JoTB*, 29(3):471–481.

Grillo, V., Carlo Gazzadi, G., Karimi, E., Mafakheri, E., Boyd, R. W., and Frabboni, S. (2014). Highly efficient electron vortex beams generated by nanofabricated phase holograms. *Applied Physics Letters*, 104(4).

Haigh, S., Jiang, B., Alloyeau, D., Kisielowski, C., and Kirkland, A. (2013). Recording low and high spatial frequencies in exit wave reconstructions. *Ultramicroscopy*, 133:26–34.

Hawkes, P., editor (2013). *Magnetic Electron Lenses*. Topics in Current Physics. Springer Berlin Heidelberg.

Hayes, M., Lim, J., and Oppenheim, A. (1980). Signal reconstruction from phase or magnitude. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(6):672–680.

Hecht, E. (1987). *Optics.* Addison Wesley Publishing Company, 2nd. edition.

Helgason, S. (2011). *Integral Geometry and Radon Transforms.* Springer New York.

Henderson, C. A., Williams, G. J., Peele, A. G., Quiney, H. M., and Nugent, K. A. (2009). Astigmatic phase retrieval: an experimental demonstration. *Opt. Express*, 17(14):11905–11915.

Huiser, A. M. J., Drenth, A. J. J., and Ferwerda, H. A. (1976). On phase retrieval in electron microscopy from image and diffraction pattern. *Optik*, 45:303.

Huiser, A. M. J. and Ferwerda, H. A. (1976). On the problem of phase retrieval in electron microscopy from image and diffraction pattern. II: On the uniqueness and stability. *Optik*, 46:407.

Huiser, A. M. J., van Toorn, P., and Ferwerda, H. A. (1977). On the problem of phase retrieval in electron microscopy from image and diffraction pattern. III: The development of an algorithm. *Optik*, 47:1.

Jaming, P. (2014). Uniqueness results in an extension of pauli's phase retrieval problem. *Applied and Computational Harmonic Analysis*, 37(3):413–441.

Kaczmarz, S. (1937). Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l'Académie Polonaise des Sciences et des Lettres. Classe des Sciences Mathématiques et Naturelles. Série A, Sciences Mathématiques*, 35:355–357.

Kasama, T., Dunin-Borkowski, R. E., and Beleggia, M. (2011). Electron holography of magnetic materials. In *Holography - Different Fields of Application*. InTech.

Keinert, F. (1989). Inversion of k-plane transforms and applications in computer tomog. *SIAM Review*, 31(2):273–298.

Koch, C. T. (2008). A flux-preserving non-linear inline holography reconstruction algorithm for partially coherent electrons. *Ultramicroscopy*, 108(2):141–150.

Koch, C. T. (2014). Towards full-resolution inline electron holography. *Micron*, 63:69–75.

Koch, C. T., Özdöl, V. B., and van Aken, P. A. (2010). An efficient, simple, and precise way to map strain with nanometer resolution in semiconductor devices. *Applied Physics Letters*, 96(9):091901.

Latychevskaia, T., Formanek, P., Koch, C., and Lubk, A. (2010). Off-axis and inline electron holography: Experimental comparison. *Ultramicroscopy*, 110(5):472 – 482. Hannes Lichte 65th Birthday.

Lenz, F. (1988). Statistics of phase and contrast determination in electron holograms. *Optik*, 79:13–14.

Levi, A. and Stark, H. (1984). Image restoration by the method of generalized projections with application to restoration from magnitude. *J. Opt. Soc. Am. A*, 1(9):932–943.

Lichte, H., Börrnert, F., Lenk, A., Lubk, A., Röder, F., Sickmann, J., Sturm, S., Vogel, K., and Wolf, D. (2013). Electron holography for fields in solids: Problems and progress. *Ultramicroscopy*, 134(0):126–134.

Lin, F., Chen, F. R., Chen, Q., Tang, D., and Peng, L.-M. (2006). The wrap-around problem and optimal padding in the exit wave reconstruction using hrtem images. *Journal of Electron Microscopy*, 55(4):191–200.

Lubk, A., Clark, L., Guzzinati, G., and Verbeeck, J. (2013a). Topological analysis of paraxially scattered electron vortex beams. *Phys. Rev. A*, 87:033834.

Lubk, A., Guzzinati, G., Börrnert, F., and Verbeeck, J. (2013b). Transport of intensity phase retrieval of arbitrary wave fields including vortices. *Phys. Rev. Lett.*, 111:173902.

Lubk, A. and Röder, F. (2015). Phase-space foundations of electron holography. *Phys. Rev. A*, 92:033844.

Luke, D., Burke, J., and Lyon, R. (2002). Optical wavefront reconstruction: Theory and numerical methods. *SIAM Rev.*, 44(2):169–224.

Lyon, R. G., Dorband, J. E., and Hollis, J. M. (1997). Hubble space telescope faint object camera calculated point-spread functions. *Appl. Opt.*, 36(8):1752–1765.

Lyon, R. G., Miller, P. E., and Gruszczak, A. (1991). HST phase retrieval: a parameter estimation. In *Proc. SPIE*, volume 1567, pages 317–326.

Marchesini, S., Tu, Y.-C., and Wu, H.-T. (2015). Alternating projection, ptychographic imaging and phase synchronization. *Applied and Computational Harmonic Analysis*.

Martin, A. and Allen, L. (2007). Phase imaging from a diffraction pattern in the presence of vortices. *Optics Communications*, 277(2):288–294.

Martin, A. V., Chen, F.-R., Hsieh, W.-K., Kai, J.-J., Findley, S. D., and Allen, L. J. (2006). Spatial incoherence in phase retrieval based on focus variation. *Ultramicroscopy*, 106:914–924.

McCartney, M. R., Agarwal, N., Chung, S., Cullen, D. A., Han, M.-G., He, K., Li, L., Wang, H., Zhou, L., and Smith, D. J. (2010). Quantitative phase imaging of nanoscale electrostatic and magnetic fields using off-axis electron holography. *Ultramicroscopy*, 110(5):375–382.

McMorran, B. J., Agrawal, A., Anderson, I. M., Herzing, A. a., Lezec, H. J., McClelland, J. J., and Unguris, J. (2011). Electron vortex beams with high quanta of orbital angular momentum. *Science*, 331(6014):192–5.

Meyer, R., Kirkland, A., and Saxton, W. (2004). A new method for the determination of the wave aberration function for high-resolution TEM. 2. Measurement of the antisymmetric aberrations. *Ultramicroscopy*, 99:115–123.

Meyer, R. R., Kirkland, A. I., and Saxton, W. O. (2002). A new method for the determination of the wave aberration function for high resolution TEM. 1. Measurement of the symmetric aberrations. *Ultramicroscopy*, 92:89–109.

Misell, D. L. (1973). An examination of an iterative method for the solution of the phase problem in optics and electron optics: I. Test calculations. *Journal of Physics D: Applied Physics*, 6(18):2200.

Natterer, F. and Wübbeling, F. (2001). *Mathematical Methods in Image Reconstruction.* Society for Industrial and Applied Mathematics.

Niermann, T. and Lehmann, M. (2016). Holographic focal series: differences between inline and off-axis electron holography at atomic resolution. *Journal of Physics D: Applied Physics*, 49(19):194002.

Niermann, T., Lubk, A., and Röder, F. (2012). A new linear transfer theory and characterization method for image detectors. Part I: Theory. *Ultramicroscopy*, 115(0):68–77.

Ohneda, Y., Baba, N., Miura, N., and Sakurai, T. (2001). Multiresolution approach to image reconstruction with phase-diversity technique. *Optical Review*, 8(1):32–36.

Ophus, C. and Ewalds, T. (2012). Guidelines for quantitative reconstruction of complex exit waves in HRTEM. *Ultramicroscopy*, 113(0):88–95.

Ozaktas, H. M. and Mendlovic, D. (1995). Fractional fourier optics. *JOSA A*, 12(4):743–751.

Parvizi, A., Broek, W. V. d., and Koch, C. T. (2016). Recovering low spatial frequencies in wavefront sensing based on intensity measurements. *Advanced Structural and Chemical Imaging*, 2(1):1–9.

Pauli, W. (1933). Die allgemeinen Prinzipien der Wellenmechanik. In *Handbuch der Physik*, volume 24. Springer Verlag, Berlin.

Petersen, T. C. and Keast, V. J. (2007). Astigmatic intensity equation for electron microscopy based phase retrieval. *Ultramicroscopy*, 107:635–643.

Petersen, T. C., Weyland, M., Paganin, D. M., Simula, T. P., Eastwood, S. A., and Morgan, M. J. (2013). Electron vortex production and control using aberration induced diffraction catastrophes. *Phys. Rev. Lett.*, 110:033901.

Pozzi, G., Beleggia, M., Kasama, T., and Dunin-Borkowski, R. E. (2014). Interferometric methods for mapping static electric and magnetic fields. *Comptes Rendus Physique*, 15:126 – 139. Seeing and measuring with electrons: Transmission Electron Microscopy today and tomorrow.

Quatieri, T. and Oppenheim, A. (1981). Iterative techniques for minimum phase signal reconstruction from phase or magnitude. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6):1187–1193.

Raymer, M. G., Beck, M., and McAlister, D. (1994). Complex wave-field reconstruction using phase-space tomography. *Phys. Rev. Lett.*, 72:1137–1140.

Reimer, L. (1989). *Transmission Electron Microscopy*. Springer Verlag.

Röder, F., Lubk, A., Wolf, D., and Niermann, T. (2014). Noise estimation for off-axis electron holography. *Ultramicroscopy*, 144(0):32–42.

Rother, A., Gemming, T., and Lichte, Hannes, L. (2009). The statistics of the thermal motion of the atoms during imaging process in transmission electron microscopy and related techniques. *Ultramicroscopy*, 109(2):139–146.

Saxton, W. O. (1994). Accurate alignment of sets of images. *Journal of Microscopy*, 174(2):61–68.

Schleich, W. P. (2001). *Quantum Optics in Phase Space*. Wiley VCH, Berlin.

Seldin, J. H. and Fienup, J. R. (1990). Numerical investigation of the uniqueness of phase retrieval. *J. Opt. Soc. Am. A*, 7(3):412–427.

Smithey, D. T., Beck, M., Raymer, M. G., and Faridani, A. (1993). Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.*, 70:1244–1247.

Song, K., Shin, G.-Y., Kim, J. K., Oh, S. H., and Koch, C. T. (2013). Strain mapping of led devices by dark-field inline electron holography: Comparison between deterministic and iterative phase retrieval approaches. *Ultramicroscopy*, 127:119–125.

Stark, H. and Sezan, M. I. (1994). *Real-Time Optical Information Processing*, chapter Image processing using projection methods, pages 185–232. Academic Press, London, UK.

Takajo, H., Takahashi, T., Itoh, K., and Fujisaki, T. (2002). Reconstruction of an object from its Fourier modulus: Development of the combination algorithm composed of the hybrid input-output algorithm and its converging part. *Appl. Opt.*, 41(29):6143–6153.

Teague, M. (1983). Deterministic phase retrieval: A Green's function solution. *J. Opt. Soc. Am.*, 73:1434–1441.

Thust, A., Coene, W., de Beeck, M. O., and Dyck, D. V. (1996). Focal-series reconstruction in hrtem: simulation studies on non-periodic objects. *Ultramicroscopy*, 64:211–230.

Tonomura, A. (1987). Applications of electron holography. *Rev. Mod. Phys.*, 59:639.

Verbeeck, J., Tian, H., and Schattschneider, P. (2010). Production and application of electron vortex beams. *Nature*, 467(7313):301–304.

Völkl, E., Allard, L. F., and D.C.Joy, editors (1999). *Introduction to Electron Holography*. Kluwer Academic / Plenum Publishers.

Wei, K. (2015). Solving systems of phaseless equations via Kaczmarz methods: a proof of concept study. *Inverse Problems*, 31(12):125008.

Yang, G., Dong, B., Gu, B., Zhuang, J., and Ersoy, O. K. (1994). Gerchberg–saxton and Yang–Gu algorithms for phase retrieval in a nonunitary transform system: a comparison. *Appl. Opt.*, 33(2):209–218.

Zou, M.-Y. and Unbehauen, R. (1997). Methods for reconstruction of 2-D sequences from Fourier transform magnitude. *IEEE Transactions on Image Processing*, 6(2):222–233.

Appendix A: Calibration of a Thick Lens

In the following a symmetric thick lens is modeled as two identical refracting (i.e. phase shifting) surfaces at distance D. The focal length of each surface is $f = 2f_{\text{thin}}$ to ensure that the resulting thin lens in the limit $D \to 0$ has focal length f_{thin}. To establish the wave optical transfer through such a thick lens, we first propagate the wave from the object plane (located at distance z_{obj} to the plane of the first refracting plane) to that of the second refracting plane, employing the previously derived thin lens formula (10)

$$\Psi\left(\mathbf{r}, D\right) = \frac{f}{2\pi\left(f - D\right)} e^{-i\frac{k_0}{2(f-D)}\mathbf{r}^2} \int_{-\infty}^{\infty} \mathrm{d}^2 k' e^{\frac{i}{2k_0}\frac{z_{\text{obj}}D - f(z_1+D)}{f-D}\mathbf{k}'^2} \tilde{\Psi}\left(\mathbf{k}', z_1\right) e^{i\frac{f}{f-D}\mathbf{r}\mathbf{k}'} . \tag{A1}$$

The corresponding Fourier transform reads

$$\tilde{\Psi}\left(\mathbf{k}, D\right) = \frac{f}{4\pi^2\left(f - D\right)} \int_{-\infty}^{\infty} \mathrm{d}^2 r e^{-i\frac{k_0}{2(f-D)}\mathbf{r}^2} \int \mathrm{d}^2 k' e^{\frac{i}{2k_0}\frac{z_{\text{obj}}D - f\left(z_{\text{obj}}+D\right)}{f-D}\mathbf{k}'^2} \tilde{\Psi}\left(\mathbf{k}', z_1\right) e^{i\left(\frac{f}{f-D}\mathbf{k}' - \mathbf{k}\right)\mathbf{r}} \tag{A2}$$

$$= \frac{f}{2\pi i k_0} \int_{-\infty}^{\infty} \mathrm{d}^2 k' e^{i\frac{f-D}{2k_0}\left(\frac{f}{f-D}\mathbf{k}' - \mathbf{k}\right)^2} e^{\frac{i}{2k_0}\frac{z_{\text{obj}}D - f\left(z_{\text{obj}}+D\right)}{f-D}\mathbf{k}'^2} \tilde{\Psi}\left(\mathbf{k}', z_1\right)$$

$$= \frac{f}{2\pi i k_0} e^{i\frac{f-D}{2k_0}\mathbf{k}^2} \int_{-\infty}^{\infty} \mathrm{d}^2 k' e^{-i\frac{f}{k_0}\mathbf{k}'\mathbf{k}} e^{\frac{i}{2k_0}\frac{z_{\text{obj}}D - f\left(z_{\text{obj}}+D-f\right)}{f-D}\mathbf{k}'^2} \tilde{\Psi}\left(\mathbf{k}', z_1\right) ,$$

and we use that result to perform the second propagation via the second refracting surface to the (conjugate) recording plane

$$\Psi\left(\mathbf{r},z\right) = \frac{f}{2\pi\left(f-z\right)}e^{-i\frac{k_0}{2(f-z)}\mathbf{r}^2}\int_{-\infty}^{\infty}\mathrm{d}^2k\, e^{\frac{i}{2k_0}\frac{-fz}{f-z}\mathbf{k}^2}\,\tilde{\Psi}\left(\mathbf{k},d\right)e^{i\frac{f}{f-z}\mathbf{r}\mathbf{k}} \tag{A3}$$

$$= \frac{f^2}{4\pi^2 i k_0\left(f-z\right)}e^{-i\frac{k_0}{2(f-z)}\mathbf{r}^2}\int_{-\infty}^{\infty}\mathrm{d}^2k'\tilde{\Psi}\left(\mathbf{k}',z_1\right)e^{\frac{i}{2k_0}\frac{z_1 d - f(z_1 + D - f)}{f - D}\mathbf{k}'^2}\int_{-\infty}^{\infty}\mathrm{d}^2k\, e^{-\frac{i}{2k_0}\left(\frac{fz}{f-z}-f+D\right)\mathbf{k}^2}e^{-i\frac{f}{k_0}\mathbf{k}'\mathbf{k}}e^{i\frac{f}{f-z}\mathbf{r}\mathbf{k}}$$

$$= \frac{f^2}{2\pi\left(\frac{fz}{f-z}-f+D\right)\left(f-z\right)}e^{-i\frac{k_0}{2(f-z)}\mathbf{r}^2}\int_{-\infty}^{\infty}\mathrm{d}^2k'\tilde{\Psi}\left(\mathbf{k}',z_1\right)e^{\frac{i}{2k_0}\frac{z_1 D - f(z_1 + D - f)}{f - D}\mathbf{k}'^2}e^{\frac{ik_0}{2\left(\frac{fz}{f-z}-f+D\right)}\left(\frac{f}{k_0}\mathbf{k}'-\frac{f}{f-z}\mathbf{r}\right)^2}$$

$$= \frac{f^2}{2\pi\left(\frac{fz}{f-z}-f+D\right)\left(f-z\right)}e^{-i\frac{k_0}{2(f-z)}\left(1+\frac{f}{\left(\frac{fz}{f-z}-f+D\right)}\right)\mathbf{r}^2}$$

$$\times\int\mathrm{d}^2k'\tilde{\Psi}\left(\mathbf{k}',z_1\right)e^{\frac{i}{2k_0}\underbrace{\left(\frac{z_1 D - f\left(z_1 + D - f\right)}{f - D}+\frac{f^2}{\frac{fz}{f-z}-f+D}\right)}_{\delta z}\mathbf{k}'^2}e^{i\underbrace{\frac{f^2}{fz-(f-D)(f-z)}}_{M^{-1}}\mathbf{k}'\mathbf{r}}.$$

The obtained expression has a similar structure than the thin lens result in that an effective free-space propagation together with an effective rescaling may be identified.

The product of these two terms

$$M\delta z = \left(\frac{z_1 D - f\left(z_1 + D - f\right)}{f - D}+\frac{f^2}{\frac{fz}{f-z}-f+D}\right)\frac{fz-\left(f-D\right)\left(f-z\right)}{f^2} \tag{A4}$$

$$= \frac{D(f - z_1)(f - z) + f\left(f(z_1 + z) - 2z_1 z\right)}{f^2}$$

occurs in the stationary phase approximation of the wave function used for evaluating the focal series of the small grating for calibration. Taking into account that the associated focal length of one refracting plane reads $f = 2/\left(C_f I^2\right)$, for the spacing between two diffraction orders we obtain

$$s = M_{\text{proj}}\frac{\lambda}{d}\left(z + z_1 + D - \left(z_1 z + \frac{D}{2}\left(z_1 + z\right)\right)C_f I^2 + \frac{D}{4}z_1 z C_f^2 I^4\right). \tag{A5}$$

It can be readily verified that this result converges to the thin lens expression (22) in the limit $D \to 0$. Expressing this result in the image and object distance from the symmetry plane of the Dif lens $z_{\text{img}} = z + D/2$ and $z_{\text{obj}} = z_1 + D/2$, we obtain

$$s = M_{\text{proj}}\frac{\lambda}{d}\left(z_{\text{img}} + z_{\text{obj}} - \left(\left(z_{\text{obj}} - \frac{D}{2}\right)\left(z_{\text{img}} - \frac{D}{2}\right) + \frac{D}{2}\left(z_{\text{obj}} + z_{\text{img}} - D\right)\right)C_f I^2\right. \tag{A6}$$

$$\left. + \frac{D}{4}\left(z_{\text{obj}} - \frac{D}{2}\right)\left(z_{\text{img}} - \frac{D}{2}\right)C_f^2 I^4\right)$$

If assuming a linear proportionality between the distance of the two refracting planes and the lens current

$$D\left(I\right) = C_D I, \tag{A7}$$

such as valid for the Glaser model of moderately excited lenses (Reimer 1989), we obtain a fourth-order polynomial dependency of the distance between two adjacent diffraction orders and the lens current

$$s = M_{\text{proj}}\frac{\lambda}{d}\left(z_{\text{img}} + z_{\text{obj}} - \left(\left(z_{\text{obj}} - \frac{C_D I}{2}\right)\left(z_{\text{img}} - \frac{C_D I}{2}\right) + \frac{C_D I}{2}\left(z_{\text{obj}} + z_{\text{img}} - C_D I\right)\right)C_f I^2\right. \tag{A8}$$

$$\left. + \frac{C_D I}{4}\left(z_{\text{obj}} - \frac{C_D I}{2}\right)\left(z_{\text{img}} - \frac{C_D I}{2}\right)C_f^2 I^4\right).$$

We used this model to fit the unknown optical parameters of the thick lens, yielding a very good agreement with the experimental data (see Figure 7).

Last but not least, we note the thick lens expressions for the other two calibration experiments, i.e., the in-focus images of the object and SA plane objects of known size

$$M_{1,2} = M_{\text{proj}} M_{1,2}^{\text{dif}} \tag{A9}$$

$$= M_{\text{proj}} \frac{f_{1,2} - z_{\text{img}} + D/2}{f_{1,2}}$$

$$= M_{\text{proj}} \left(1 - C_f I_{1,2}^2 \left(z_{\text{img}} - \frac{C_D I_{1,2}}{2} \right) \right) .$$

Eliminating M^{proj} then yields

$$M_2 C_f I_1^2 \left(z_{\text{img}} - \frac{C_D I_1}{2} \right) - M_2 \tag{A10}$$
$$= M_1 C_f I_2^2 \left(z_{\text{img}} - \frac{C_D I_2}{2} \right) - M_1 ,$$

which can be solved for

$$C_f = \frac{M_2 - M_1}{M_2 I_1^2 \left(z_{\text{img}} - \frac{C_D I_1}{2} \right) - M_1 I_2^2 \left(z_{\text{img}} - \frac{C_D I_2}{2} \right)} \tag{A11}$$

and the post-magnification of the projector lenses

$$M_{\text{proj}} = \frac{M_{1,2}}{1 - C_f I_{1,2}^2 \left(z_{\text{img}} - \frac{C_D I_{1,2}}{2} \right)} . \tag{A12}$$

Accordingly, the three calibration experiments employed for the "thin" Diffraction lens are also sufficient for modeling the Diffraction lens as a thick lens. All optical parameters shown in the main text are obtained from this thick lens model. Possible sources for the remaining errors in the calibrated parameters are the experimental errors of the calibration experiments and the limited validity of the proportionality between the lens thickness and the current as well as the reciprocal focal length (refractive power) and the squared lens current.