

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Lemont, Illinois 60439

**Single-View Phase Retrieval of an Extended Sample by
Exploiting Edge Feature Detection and Sparsity¹**

Ashish Tripathi, Ian McNulty, Todd Munson, and Stefan M. Wild

Mathematics and Computer Science Division

Preprint ANL/MCS-P6028-0716

July 2016

Updates to this preprint may be found at
<http://www.mcs.anl.gov/publications>

¹This material was based upon work supported by the U.S. Department of Energy, Office of Science, Offices of Basic Energy Sciences and Advanced Scientific Computing Research under Contract No. DE-AC02-06CH11357.

Single-View Phase Retrieval of an Extended Sample by Exploiting Edge Detection and Sparsity

Ashish Tripathi,^{1*} Ian McNulty,² Todd Munson,¹ and Stefan M. Wild¹

¹Mathematics and Computer Science Division,

²Center for Nanoscale Materials;

Argonne National Laboratory, Lemont, IL 60439, USA

[*atripathi@mcs.anl.gov](mailto:atripathi@mcs.anl.gov)

Abstract: We propose a new approach to robustly retrieve the exit wave of an extended sample from its coherent diffraction pattern by exploiting sparsity of the sample's edges. This approach enables imaging of an extended sample with a single view, without ptychography. We introduce nonlinear optimization methods that promote sparsity, and we derive update rules to robustly recover the sample's exit wave. We test these methods on simulated samples by varying the sparsity of the edge-detected representation of the exit wave. Our tests illustrate the strengths and limitations of the proposed method in imaging extended samples.

© 2016 Optical Society of America

OCIS codes: (100.5070) Phase retrieval; (110.3010) Image reconstruction techniques; (100.3190) Inverse problems; (110.7440) X-ray imaging; (340.6720) Synchrotron radiation; (110.4280) Noise in imaging systems

References and links

1. D. Paganin, *Coherent X-ray Optics* (Oxford Press, 2006).
2. Y. S. G. Nashed, D. J. Vine, T. Peterka, J. Deng, R. Ross, and C. Jacobsen, "Parallel ptychographic reconstruction," *Opt. Express* **22**, 32082–32097 (2014).
3. J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature* **400**, 342–344 (1999).
4. S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence, "X-ray image reconstruction from a diffraction pattern alone," *Phys. Rev. B* **68**, 140101 (2003).
5. H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm," *Phys. Rev. Lett.* **93**, 023903 (2004).
6. P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, "High-resolution scanning X-ray diffraction microscopy," *Science* **321**, 379–382 (2008).
7. P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," *Nature* **494**, 68–71 (2013).
8. S. Marathe, S. S. Kim, S. N. Kim, C. Kim, H. C. Kang, P. V. Nickles, and D. Y. Noh, "Coherent diffraction surface imaging in reflection geometry," *Opt. Express* **18**, 7253–7262 (2010).
9. M. A. Pfeifer, G. J. Williams, I. A. Vartanyants, R. Harder, and I. K. Robinson, "Three-dimensional mapping of a deformation field inside a nanocrystal," *Nature* **442**, 63–66 (2006).
10. J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Optics* **21**, 2758–2769 (1982).
11. S. Marchesini, "Ab initio compressive phase retrieval," Tech. Rep. 0809.2006, arXiv (2008).
12. M. L. Moravec, J. K. Romberg, and R. G. Baraniuk, "Compressive phase retrieval," *Proc. SPIE* **6701**, 670120 (2007).
13. R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, "Potential for biomolecular imaging with femtosecond x-ray pulses," *Nature* **406**, 752–757 (2000).
14. B. Abbey, K. Nugent, G. Williams, J. Clark, A. Peele, M. Pfeifer, M. de Jonge, and I. McNulty, "Keyhole coherent diffractive imaging," *Nature Phys.* **4**, 394–398 (2008).

15. E. Candés and M. Wakin, “An introduction to compressive sampling,” *Signal Processing Magazine, IEEE* **25**, 21–30 (2008).
16. X. Huang, J. Nelson, J. Steinbrener, J. Kirz, J. J. Turner, and C. Jacobsen, “Incorrect support and missing center tolerances of phasing algorithms,” *Opt. Express* **18**, 26441–26449 (2010).
17. A. Tripathi, S. Leyffer, T. Munson, and S. M. Wild, “Visualizing and improving the robustness of phase retrieval algorithms,” *Procedia Computer Science* **51**, 815–824 (2015). *International Conference On Computational Science, (ICCS) 2015*.
18. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends in Machine Learning* **3**, 1–122 (2011).
19. E. J. Candés, M. B. Wakin, and S. P. Boyd, “Enhancing sparsity by reweighted ℓ_1 minimization,” *Journal of Fourier Analysis and Applications* **14**, 877–905 (2008).
20. T. Blumensath and M. E. Davies, “Iterative thresholding for sparse approximations,” *Journal of Fourier Analysis and Applications* **14**, 629–654 (2008).
21. T. Blumensath and M. E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied and Computational Harmonic Analysis* **27**, 265–274 (2009).
22. F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, “Optimization with sparsity-inducing penalties,” *Foundations and Trends in Machine Learning* **4**, 1–106 (2011).
23. R. H. Chan, T. F. Chan, L. Shen, and Z. Shen, “Wavelet algorithms for high-resolution image reconstruction,” *SIAM Journal on Scientific Computing* **24**, 1408–1432 (2003).
24. K. Herrity, A. Gilbert, and J. Tropp, “Sparse approximation via iterative thresholding,” in *2006 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. III (IEEE, 2006), pp. 624–627.
25. Z. Yang, C. Zhang, and L. Xie, “Robust compressive phase retrieval via L1 minimization with application to image reconstruction,” *Tech. Rep. 1302.0081, arXiv* (2013).
26. J. J. Moré and S. M. Wild, “Estimating derivatives of noisy simulations,” *ACM Transactions on Mathematical Software* **38**, 19:1–19:21 (2012).
27. X. Zhang, Y. Lu, and T. Chan, “A novel sparsity reconstruction method from Poisson data for 3D bioluminescence tomography,” *Journal of Scientific Computing* **50**, 519–535 (2011).
28. B. Dong and Y. Zhang, “An efficient algorithm for ℓ_0 minimization in wavelet frame based image restoration,” *Journal of Scientific Computing* **54**, 350–368 (2012).
29. Y. Zhang, B. Dong, and Z. Lu, “ ℓ_0 minimization for wavelet frame based image restoration,” *Math. Comput.* **82**, 995–1015 (2013).
30. A. Pein, S. Looock, G. Plonka, and T. Salditt, “Using sparsity information for iterative phase retrieval in x-ray propagation imaging,” *Opt. Express* **24**, 8332–8343 (2016).
31. R. Fan, Q. Wan, F. Wen, H. Chen, and Y. Liu, “Iterative projection approach for phase retrieval of semi-sparse wave field,” *EURASIP Journal on Advances in Signal Processing* **2014**, 1–13 (2014).
32. A. M. Maiden and J. M. Rodenburg, “An improvedptychographical phase retrieval algorithm for diffractive imaging,” *Ultramicroscopy* **109**, 1256–1262 (2009).
33. J. N. Clark, C. T. Putkunz, M. A. Pfeifer, A. G. Peele, G. J. Williams, B. Chen, K. A. Nugent, C. Hall, W. Fullagar, S. Kim, and I. McNulty, “Use of a complex constraint in coherent diffractive imaging,” *Opt. Express* **18**, 1981–1993 (2010).
34. L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: Nonlinear Phenomena* **60**, 259–268 (1992).
35. H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. H. Spence, U. Weierstall, T. Beetz, C. Jacobsen, and D. Shapiro, “High-resolution ab initio three-dimensional x-ray diffraction microscopy,” *J. Opt. Soc. Am. A* **23**, 1179–1200 (2006).
36. J. Steinbrener, J. Nelson, X. Huang, S. Marchesini, D. Shapiro, J. J. Turner, and C. Jacobsen, “Data preparation and evaluation techniques for x-ray diffraction microscopy,” *Opt. Express* **18**, 18598–18614 (2010).
37. D. Lazzaro and L. Montefusco, “Edge-preserving wavelet thresholding for image denoising,” *Journal of Computational and Applied Mathematics* **210**, 222–231 (2007).
38. S. Yi, D. Labate, G. R. Easley, and H. Krim, “A shearlet approach to edge analysis and detection,” *IEEE Transactions on Image Processing* **18**, 929–941 (2009).
39. J.-L. Starck, E. J. Candés, and D. L. Donoho, “The curvelet transform for image denoising,” *IEEE Transactions on Image Processing* **11**, 670–684 (2002).

1. Introduction

Coherent diffractive imaging (CDI) is a technique in widespread use at x-ray synchrotrons to image nanoscale materials. Traditional x-ray microscopy techniques, such as transmission x-ray microscopy (TXM), rely on condensing and objective optics. In contrast, CDI uses nonlinear optimization methods as a type of “numerical optic” to solve the phase problem, which arises

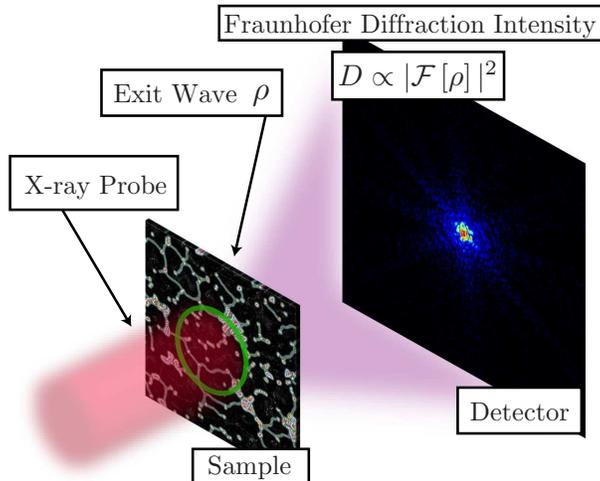


Fig. 1. CDI experimental setup: a complex-valued exit wave ρ results in a real-valued data D collected by the detector.

because of an x-ray detector's inability to measure a full complex-valued wavefield.

Under the projection approximation [1], a (complex-valued¹) exit wave ρ is formed when a plane wave x-ray beam interacts with a sample. As shown in Fig. 1, when an area detector is placed in the far field, the (real-valued) Fraunhofer diffraction intensity D is measured, where D is proportional to the squared modulus of the two-dimensional Fourier transform of ρ . In this paper, we study nonlinear optimization methods to recover the phase lost upon measurement.

An advantage of CDI over more traditional microscopy techniques is that the recovered image's spatial resolution is determined by the highest spatial frequencies measured. These spatial frequencies are in turn determined primarily by incident x-ray fluence and the contrast in both the phase and magnitude of the sample's complex valued transmission function. Another advantage is that the full complex-valued transmission function (related to the complex-valued index of refraction wherein all the interesting physics is located) is recovered. Methods such as TXM can recover only the modulus of the sample's transmission function. Furthermore, development of phase retrieval algorithms for CDI has progressed to a point where high-quality images are produced in real time and at the same speed with which one can obtain images in a TXM (see, e.g., [2]).

Recovery of the phase typically starts with a guess $\rho^{(0)}$ for the exit wave; a nonlinear optimization algorithm then iteratively refines this guess using knowledge from the experiment performed and the physics of the sample. This knowledge is typically used to define constraints for a nonlinear, nonconvex optimization formulation of the phase retrieval problem. Examples of such knowledge include the measured diffraction, as well as constraints on the sample such as known compact support [3], sparsity [4], and ptychography [5, 6, 7].

Common CDI experimental configurations for collecting data include the small angle geometry shown in Fig. 1, as well as reflection and Bragg geometries [8, 9]. With these configurations, one can use an x-ray probe with a spatial area (or volume in the reflection and Bragg cases) larger or smaller than the sample under investigation. When the x-ray probe is larger than the sample, a single diffraction pattern from a single view can be used with known sup-

¹In our figures, the phase and magnitude of complex-valued quantities are represented in an HSV colorspace, with the hue and value/brightness denoting the phase and magnitude respectively, and with the saturation component set to maximum for all image pixels.

port or sparsity [3, 4, 10, 11, 12]. This single-view approach is particularly advantageous when the x-ray probe will damage the sample, such as with an x-ray free-electron laser source [13]. Improvements in single-view algorithms can also readily be incorporated into ptychographic methods for more robust phase retrieval.

In this paper, we explore the use of single-view phase retrieval methods for an extended sample, with a primary motivation being the ability to overcome difficulties encountered in ptychography when attempting to image samples that significantly change during the collection of a ptychographic dataset, or when the long-term temporal and spatial experimental stability requirements of ptychography are not met. Single-view phase retrieval methods for an extended sample have had very limited success in the literature [14] when compared with methods for an isolated sample. In Sec. 2, we identify the primary limitations of using a support defined from the approximate spatial extent of the illumination function when imaging an extended sample and in Sec. 3 we discuss sparsity in an edge-detected representation of the exit wave. We then develop in Sec. 4 a phase retrieval algorithm that overcomes these limitations by performing phase retrieval on the edge-detected representation of an exit wave and by using ideas from compressive sensing [15]. The primary idea is to recover the sparsest possible edge-detected representation of the exit wave consistent with the diffraction measurement. This process is colloquially known as “shrinkwrap” in the experimental phase retrieval community [4] and more broadly known as “compressive phase retrieval” [11, 12]. The numerical tests in Sec. 5 on a range of simulated samples illustrate the benefits and limitations of the proposed methods. We test these methods on various simulated samples by varying the sparsity of the edge-detected representation of the exit wave. Our results demonstrate that our method can reliably recover the phase of extended samples in many cases, cases where standard, single-view phase retrieval methods fail. Nontrivial noise in the measured diffraction, however, requires use of an appropriate spatial frequency filtering procedure, and this is also discussed. In Sec. 6, we discuss further prospects and extensions of the proposed methodology.

2. Support Defined from the Illuminated Area on the Sample

In single-view phase retrieval of an extended sample, one obvious choice of sample constraint is to use the approximate illuminated area on the sample as the support on the exit wave we wish to recover. As an example, in Fig. 2a we simulate an x-ray probe by solving the Fresnel integral with a circular pinhole aperture as input. This type of probe is typical in small angle transmission geometry x-ray synchrotron experiments, where sample-to-circular-pinhole distances of between a millimeter and a centimeter are used. With this probe and a simulated sample transmission function defined by generating a periodic Voronoi pattern (not shown), the exit wave ρ_{true} under the projection approximation [1] is shown in Fig. 2b (the red dashed line will be discussed shortly). The primary difficulty arising when attempting to use the approximate illuminated area on the sample as a support constraint is that the x-ray probe photon area density smoothly decays as one moves away from the center of the probe function. This smooth decay makes defining an effective support boundary ambiguous.

As an example, we use a support dynamically generated using the shrinkwrap method [4] as a constraint on the exit wave and the Fourier modulus measurement \sqrt{D} as the measurement constraint. We denote the current iterate of the exit wave as $\rho^{(k)}$, where k denotes the iteration index and the problem size is $M \times N = 512 \times 512$. We run 4,900 iterations of Fienup’s hybrid input-output method (HIO), then 100 iterations of error reduction (ER), and then a shrinkwrap update of the binary support mask, denoted by the symbol $\mathbb{I}_S^{(k)}$. This sequence of HIO+ER+shrinkwrap is repeated 20 times. When updating the binary support mask in shrinkwrap, we Gaussian blur (using $\sigma_{\text{sw}} = 0.01\sqrt{MN}$ as the standard deviation of the Gaussian low-pass filter in Fourier space) the current iterate of $|\rho^{(k)}|$ to get the quantity $\rho_{\text{blur}}^{(k)}$. We then enforce a sparsity ratio of

$\kappa = \frac{1}{MN} \|\rho_{\text{blur}}^{(k)}\|_0 = 0.15$ on $\rho_{\text{blur}}^{(k)}$ to define the binary support mask. The sparsity ratio is the fraction of nonzeros since

$$\|a\|_0 = \sum_{\mathbf{r}} \mathbb{I}_{\{a \neq 0\}}(\mathbf{r}), \quad (1)$$

where the indicator function used is defined as

$$\mathbb{I}_{\{a \neq 0\}}(\mathbf{r}) = \begin{cases} 1 & \text{if } a(\mathbf{r}) \neq 0 \\ 0 & \text{if } a(\mathbf{r}) = 0, \end{cases} \quad (2)$$

for some $a \in \mathbb{C}^{M \times N}$, and we index the spatial coordinate system of the sample by $\mathbf{r} \in \{(r_n, r_m) : n = 0, \dots, N-1, m = 0, \dots, M-1\}$. To define the binary support mask with a particular enforced sparsity ratio κ , we determine the $[MN\kappa]$ th largest value of $\rho_{\text{blur}}^{(k)}$. The binary support mask is then created when we set to zero the elements of $\rho_{\text{blur}}^{(k)}$ below this value and set to one the elements above or equal to this value. A typical result for the support boundary using this recipe is shown by the dashed red line around the ground truth exit wave ρ_{true} in Fig. 2b. A linecut of the ground truth exit wave ρ_{true} through the main diagonal is shown in Fig. 2d, with the

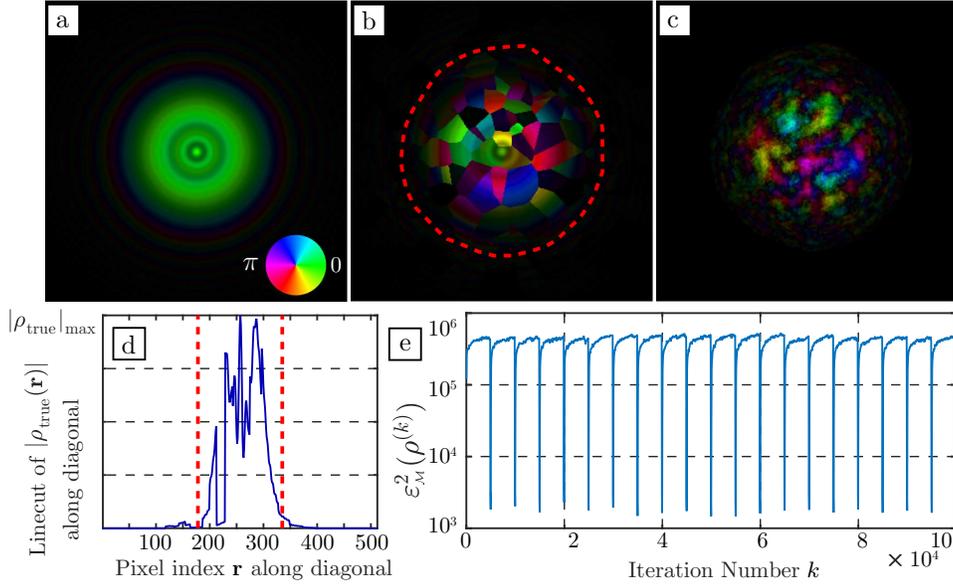


Fig. 2. Simulation results when using the approximate diameter of the probe as the support for the exit wave ρ . (a) The probe function generated by solving the Fresnel integral using a simulated circular pinhole as input. (b) The ground truth exit wave ρ_{true} generated by using the projection approximation with the probe shown in (a) and a simulated sample transmission function (not shown) defined as a Voronoi pattern. A typical boundary, where inside (outside) the red dotted line the support $\mathbb{I}_S^{(k)}$ takes on values of 1 (0), of the binary support mask \mathbb{I}_S generated by using shrinkwrap is shown as the red dotted line. (c) The unsuccessfully recovered exit wave $\rho^{(k)}$ after $k = 10^5$ iterations of combined HIO, ER, and shrinkwrap. (d) A linecut of the exit wave ρ_{true} through the main diagonal. The boundary of the support (red dotted line) in (b) is shown here also as the red dotted lines. The support boundary shown potentially can zero out features of $|\rho_{\text{true}}|$ below approximately $0.05|\rho_{\text{true}}|_{\text{max}}$. (e) The measurement metric Eq. (3) versus iteration number during the HIO+ER+shrinkwrap sequence.

support boundary again shown as the dotted red lines. The support boundary shown in Fig. 2d potentially can zero out features of $|\rho_{\text{true}}|$ below approximately $0.05|\rho_{\text{true}}|_{\text{max}}$.

The results for the recovered exit wave $\rho^{(k)}$ for $k = 10^5$ are shown in Fig. 2c. The support generated using shrinkwrap is not constraining enough to accurately recover the ground truth exit wave shown in Fig. 2b. The value of the standard error metric

$$\varepsilon_{\mathcal{M}}^2(\rho^{(k)}) = \sum_{\mathbf{q}} \left| |\tilde{\rho}^{(k)}(\mathbf{q})| - \sqrt{D(\mathbf{q})} \right|^2, \quad (3)$$

used to measure solution quality, as a function of iteration number k is shown in Fig. 2e. There, we use the notation $\tilde{\rho} = \mathcal{F}[\rho] = |\tilde{\rho}| \odot e^{i\tilde{\phi}}$, where $\mathcal{F}[\cdot]$ is the two-dimensional discrete Fourier transform, \odot is the (componentwise) Hadamard product, and we index the Fourier space coordinate system of the measurement by \mathbf{q} , so that $\tilde{\rho}(\mathbf{q}) = |\tilde{\rho}(\mathbf{q})| \odot e^{i\tilde{\phi}(\mathbf{q})}$ for $\mathbf{q} \in \{(q_n, q_m) : n = 0, \dots, N-1, m = 0, \dots, M-1\}$. If we enforce smaller sparsity ratios (i.e., $\kappa < 0.15$), the support generated is “too tight” [16], and the recovered $\rho^{(k)}$ is similarly unsatisfactory when compared with the ground truth exit wave in Fig. 2b. Enforcing larger sparsity ratios (i.e., $\kappa > 0.15$) causes prompt algorithmic stagnation when monitoring Eq. (3), indicating convergence to a poor local solution [17] and again showing unsatisfactory results when compared with the ground truth exit wave in Fig. 2b.

3. Sparse Exit Wave Representations for Phase Retrieval

Mathematical transformations can be applied to represent a group of variables in a more meaningful and useful way. An example is the wavelet transformation of an image, where many images have only a few dominant wavelet components. Ignoring the nondominant components yields lossy compression of the image into a potentially small fraction of the original number of information bytes the image comprised. The information content of the image is *sparse* in the wavelet domain. The technique of “compressive sensing” attempts to incorporate and exploit properties of this inherent sparsity. The basic idea is that by formulating optimization problems in terms of a simplified representation instead of the original representation and solving for an optimal sparse, simplified representation, one might obtain more robust results.

Here, we explore the use of edge detection as a sparsifying transformation of an exit wave ρ . This idea was originally explored in [11] for use in phase retrieval. For edge detection, we take the forward-difference arrays (each of size $M \times N = 512 \times 512$)

$$E_x = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad E_y = E_x^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad (4)$$

and convolve ρ with E_x and E_y . Since this edge detection uses forward differences, we refer to the resulting quantities as *spatial derivatives* and use the notation $\partial_x \rho = E_x * \rho$, $\partial_y \rho = E_y * \rho$. The convolution theorem shows

$$\partial_j \rho = E_j * \rho = \mathcal{F}^{-1} [\mathcal{F}[E_j] \odot \mathcal{F}[\rho]] = \mathcal{F}^{-1} [\tilde{E}_j \odot \tilde{\rho}], \quad (5)$$

for $j = x, y$ and with $\tilde{E}_j = \mathcal{F}[E_j]$.

Figures 3a, 3c, and 3e show the ground truth exit wave ρ_{true} used in Fig. 2b and its respective edge-detected representations. In Fig. 3b we again show the binary support mask $\mathbb{I}_{\mathcal{S}}$ generated

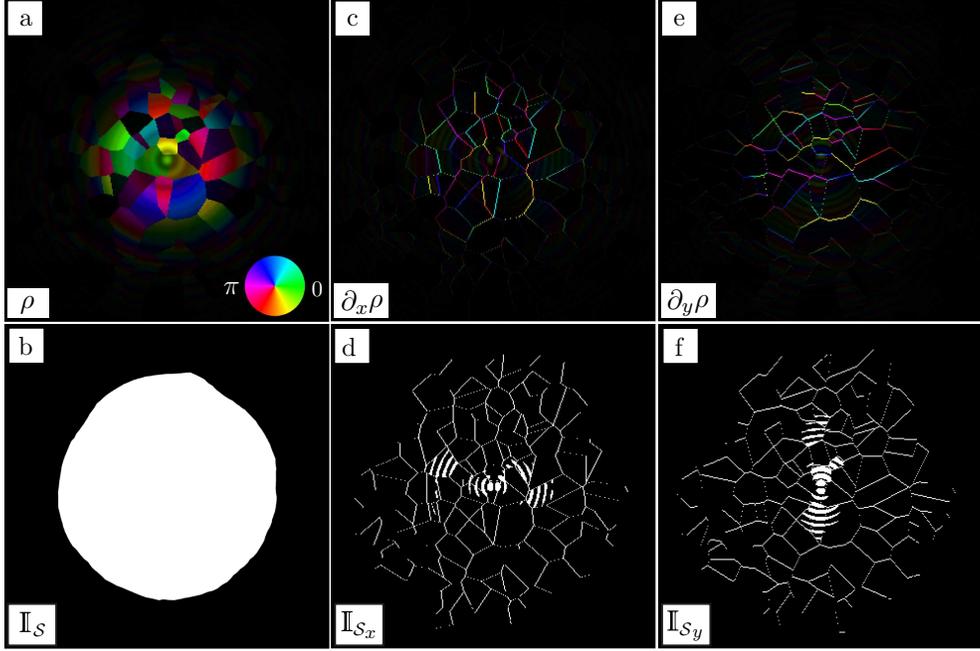


Fig. 3. (a) The ground truth exit wave ρ_{true} . (b) A support \mathbb{I}_S generated by hard thresholding $|\rho_{\text{true}}|$ by 2% of the maximum value of $|\rho_{\text{true}}|$, with $\sum_{\mathbf{r}} \mathbb{I}_S(\mathbf{r})/MN = 0.15$ being the resulting sparsity ratio and $\mathbb{I}_S(\mathbf{r})$ defined in Eq. (6). In (c,e), $\partial_x \rho_{\text{true}}$ and $\partial_y \rho_{\text{true}}$, respectively, the x and y components of the forward-difference gradient of ρ_{true} , are shown. Shown in (d,f) are the supports \mathbb{I}_{S_x} and \mathbb{I}_{S_y} generated by hard thresholding 2% of the maximum value of $|\partial_x \rho_{\text{true}}|$ and $|\partial_y \rho_{\text{true}}|$, respectively; $\frac{1}{MN} \|\mathbb{I}_{S_y}\|_0 = 0.0141$ and $\frac{1}{MN} \|\mathbb{I}_{S_x}\|_0 = 0.0133$ are the respective sparsity ratios.

by using shrinkwrap in Sec. 2 to define a support corresponding to the approximate illuminated area on the sample. In Figs. 3d and 3f, we show the supports $\mathbb{I}_{S_x}, \mathbb{I}_{S_y}$ that result from thresholding the modulus of the respective quantities in Figs. 3c and 3e to zero whenever they fall below 2% of the maximum value of the modulus:

$$\mathbb{I}_{S_j}(\mathbf{r}) = \begin{cases} 1 & \text{if } |\partial_j \rho| > 0.02 |\partial_j \rho|_{\max} \\ 0 & \text{if } |\partial_j \rho| \leq 0.02 |\partial_j \rho|_{\max}, \end{cases} \quad \text{for } j = x, y. \quad (6)$$

From Fig. 3, we can clearly see that the $\partial_x \rho$ and $\partial_y \rho$ quantities have far sparser supports when compared with ρ_{true} . We explore this property in the remainder of this paper and show that by performing phase retrieval to solve for the quantities $\partial_x \rho$ and $\partial_y \rho$ with relatively sparse supports in place of solving for the exit wave ρ with a nonsparse support, we can robustly recover the exit wave for extended samples in many cases.

4. Solving for the Sparse Representation of the Exit Wave

The alternating direction method of multipliers (ADMM) attempts to solve optimization problems by breaking them into smaller, more manageable subproblems [18]. It has the advantages that the objective function used in the optimization problem need not be differentiable and the method is simple to implement and parallelize. The convergence properties of this method are difficult to analyze for nonconvex problems, however, and more specialized algorithms can

outperform it. We begin a derivation of ADMM update rules for obtaining the sparse representations $\partial_j \rho$ of the exit wave ρ by considering the constrained, nonconvex optimization problem

$$\begin{aligned} \min_{u_x, u_y, \rho} \quad & \|u_x\|_\ell + \|u_y\|_\ell \\ \text{subject to} \quad & u_x = \partial_x \rho \\ & u_y = \partial_y \rho \\ & \rho \in \mathcal{M}, \end{aligned} \quad (7)$$

where $\ell \in \{0, 1\}$ and we define the Fourier modulus measurement constraint set

$$\mathcal{M} = \left\{ \rho : \rho = \mathcal{F}^{-1} [\sqrt{D} \odot e^{i\tilde{\phi}}] \right\} \quad (8)$$

and its associated projection operator $\Pi_{\mathcal{M}}$

$$\Pi_{\mathcal{M}}[\rho] = \mathcal{F}^{-1} [\sqrt{D} \odot e^{i\tilde{\phi}}], \quad (9)$$

with $\tilde{\rho} = \mathcal{F}[\rho] = |\tilde{\rho}| \odot e^{i\tilde{\phi}}$. The $\|\cdot\|_0$ norm is defined in Eq. (1), and for any $a \in \mathbb{C}^{M \times N}$ we define the $\|\cdot\|_1$ norm by

$$\|a\|_1 = \sum_{\mathbf{r}} |a(\mathbf{r})|. \quad (10)$$

The $\ell \in \{0, 1\}$ norms are well-known sparsity-inducing norms, with the $\|\cdot\|_1$ norm frequently used as a convex and continuous relaxation of the $\|\cdot\|_0$ norm [19]. The optimization problem in (7) seeks the ρ satisfying the Fourier modulus measurement with sparsest possible $\partial_j \rho$.

The constrained augmented Lagrangian function for Eq. (7) is

$$\begin{aligned} \mathcal{L}(u_x, u_y, \rho, \lambda_x, \lambda_y, \tau_x, \tau_y, \ell) &= \sum_{j \in \{x, y\}} \mathcal{L}_j(u_j, \rho, \lambda_j, \tau_j, \ell) \\ &= \sum_{j \in \{x, y\}} \left[\tau_j \|u_j\|_\ell + \|u_j - \partial_j \rho\|_F^2 + 2\text{Re} \left[\sum_{\mathbf{r}} \lambda_j^* \odot (u_j - \partial_j \rho) \right] \right] \\ &= \sum_{j \in \{x, y\}} \left[\tau_j \|u_j\|_\ell + \|u_j - \partial_j \rho + \lambda_j\|_F^2 - \|\lambda_j\|_F^2 \right] \end{aligned} \quad (11a)$$

$$= \sum_{j \in \{x, y\}} \left[\tau_j \|u_j\|_\ell + \|\tilde{u}_j - \tilde{E}_j \odot \tilde{\rho} + \tilde{\lambda}_j\|_F^2 - \|\lambda_j\|_F^2 \right], \quad (11b)$$

where $\tau_j \geq 0$ are parameters that control the competition between the constraints $u_j = \partial_j \rho$ and $\rho \in \mathcal{M}$ and finding the sparsest possible u_j . The $\lambda_j \in \mathbb{C}^{M \times N}$ are Lagrange multipliers (with λ_j^* denoting the complex conjugate), $\tilde{u}_j = \mathcal{F}[u_j]$, $\tilde{\lambda}_j = \mathcal{F}[\lambda_j]$, and $\tilde{\rho} = \mathcal{F}[\rho]$. We use Plancherel's theorem for the discrete Fourier transform when going from Eq. (11a) to Eq. (11b) on the Frobenius norm

$$\|a\|_F^2 = \sum_{\mathbf{r}} |a(\mathbf{r})|^2 = \sum_{\mathbf{q}} |\tilde{a}(\mathbf{q})|^2, \quad (12)$$

for any $a \in \mathbb{C}^{M \times N}$ and with $\tilde{a}(\mathbf{q}) = \mathcal{F}[a(\mathbf{r})]$.

Provided that $\ell \in \{0, 1\}$ is fixed and given initializations for $\rho^{(0)}$, $u_j^{(0)} = \partial_j \rho^{(0)}$, and $\lambda_j^{(0)}$ for

$j \in \{x, y\}$, the k th iteration of the ADMM-based method can be written as

$$\tau_j^{(k+1)} = \arg \min_{\tau_j} \psi_j(u_j^{(k)}, \rho^{(k)}, \lambda_j^{(k)}, \tau_j, \ell) \quad \text{for } j \in \{x, y\} \quad (13a)$$

$$u_j^{(k+1)} = \arg \min_{u_j} \mathcal{L}_j(u_j, \rho^{(k)}, \lambda_j^{(k)}, \tau_j^{(k+1)}, \ell) \quad \text{for } j \in \{x, y\} \quad (13b)$$

$$\rho^{(k+1)} = \arg \min_{\rho \in \mathcal{M}} \mathcal{L}(u_x^{(k+1)}, u_y^{(k+1)}, \rho, \lambda_x^{(k)}, \lambda_y^{(k)}, \tau_x^{(k+1)}, \tau_y^{(k+1)}, \ell) \quad (13c)$$

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + \beta_j(u_j^{(k+1)} - \partial_j \rho^{(k+1)}) \quad \text{for } j \in \{x, y\} \quad (13d)$$

with β_j being a damping term for the multiplier updates (typically we set $\beta_j = 1$ for both $j \in \{x, y\}$). The update rules for $u_j^{(k+1)}$ and $\rho^{(k+1)}$ are derived in Secs. 4.1 and 4.2, respectively. The function $\psi_j(\cdot)$ is defined in Sec. 4.3 and allows for effective iterative determination of the parameters τ_j . Note that we cannot minimize Eq. (11) to update the τ_j because $(\tau_x, \tau_y) = (0, 0)$ is the global minimizer of Eq. (11) when holding the other quantities constant, which is not useful for our purposes. Thus, we introduce the additional auxiliary function $\psi_j(\cdot)$ for determining the optimal values of the parameters τ_j .

4.1. $u_j^{(k+1)}$ Update

The least-squares optimization subproblem with the sparsity-promoting $\ell \in \{0, 1\}$ -norm regularization is

$$u_j^{(k+1)} = \arg \min_{u_j} \tau_j \|u_j\|_\ell + \|u_j - b\|_F^2, \quad (14)$$

which is equivalent to minimizing \mathcal{L}_j in Eq. (11a) over u_j when we set $b = \partial_j \rho - \lambda_j$ and we drop the $\|\lambda_j\|_F$ constant. The solution to Eq. (14) can be expressed in terms of proximal mappings [20, 21, 22, 23, 24, 25]. If using $\ell = 1$, this mapping is the soft-thresholding operator

$$\mathbf{T}_1(\tau, b) = \text{sgn}(b) \odot (|b| - \tau) \odot \mathbb{I}_{\{|b| > \tau\}}, \quad (15)$$

where we use the notation of the complex signum function $\text{sgn}(b) = e^{i\gamma}$ for $b = |b| \odot e^{i\gamma}$ when $b \neq 0$ and $\text{sgn}(0) = 0$ with $\gamma \in \mathbb{R}^{M \times N}$. If using $\ell = 0$, this mapping is the hard-thresholding operator

$$\mathbf{T}_0(\tau, b) = b \odot \mathbb{I}_{\{|b| > \sqrt{\tau}\}}. \quad (16)$$

In both cases, we have an indicator function of the form

$$\mathbb{I}_{\{|b| > \mu\}}(\mathbf{r}) = \begin{cases} 1 & \text{if } |b(\mathbf{r})| > \mu \\ 0 & \text{if } |b(\mathbf{r})| \leq \mu \end{cases} \quad (17)$$

for a specified $\mu \geq 0$. We thus use the update rule

$$\begin{aligned} u_j^{(k+1)} &= \arg \min_{u_j} \tau_j^{(k)} \|u_j\|_\ell + \|u_j - \partial_j \rho^{(k)} + \lambda_j^{(k)}\|_F^2 \\ &= \mathbf{T}_\ell(\tau_j^{(k)}, \partial_j \rho^{(k)} - \lambda_j^{(k)}), \quad \text{for } j \in \{x, y\} \end{aligned} \quad (18)$$

in Eq. (13b). In addition to balancing the competition between fitting the measurement and promoting sparsity, we can interpret τ_j as thresholding parameters.

4.2. $\rho^{(k+1)}$ Update

An update for $\rho^{(k+1)}$ can be found by taking the Wirtinger (complex-valued) derivative of the $\tilde{\rho}$ -dependent term in Eq. (11b) with respect to $\tilde{\rho}$ and setting it to zero:

$$\frac{\partial}{\partial \tilde{\rho}} \sum_{j \in \{x,y\}} \|\tilde{u}_j^{(k+1)} - \tilde{E}_j \odot \tilde{\rho} + \tilde{\lambda}_j^{(k)}\|_F^2 = 0. \quad (19)$$

Solving for $\tilde{\rho}$, we obtain

$$\tilde{\rho} = \frac{\tilde{E}_x^* \odot \tilde{u}_x^{(k+1)} + \tilde{E}_x^* \odot \tilde{\lambda}_x^{(k)} + \tilde{E}_y^* \odot \tilde{u}_y^{(k+1)} + \tilde{E}_y^* \odot \tilde{\lambda}_y^{(k)}}{|\tilde{E}_y|^2 + |\tilde{E}_x|^2}, \quad (20)$$

where the division by $|\tilde{E}_y|^2 + |\tilde{E}_x|^2$ is componentwise. To satisfy the constraint in Eq. (13c) that $\rho \in \mathcal{M}$, we then project Eq. (20) onto the constraint set \mathcal{M} :

$$\rho^{(k+1)} = \Pi_{\mathcal{M}} \left[\mathcal{F}^{-1} \left[\frac{\tilde{E}_x^* \odot \tilde{u}_x^{(k+1)} + \tilde{E}_x^* \odot \tilde{\lambda}_x^{(k)} + \tilde{E}_y^* \odot \tilde{u}_y^{(k+1)} + \tilde{E}_y^* \odot \tilde{\lambda}_y^{(k)}}{|\tilde{E}_y|^2 + |\tilde{E}_x|^2} \right] \right], \quad (21)$$

which we use for the update in Eq. (13c).

4.3. $\tau_j^{(k+1)}$ Update

The choice of thresholding parameters (τ_x, τ_y) is critical to the success of our method. We determine τ_j parameters by solving Eq. (13a), where we introduce the cost function

$$\psi_j(u_j^{(k)}, \rho^{(k)}, \lambda_j^{(k)}, \tau_j, \ell) = \left\| |\tilde{E}_j| \odot \sqrt{D} - |\mathcal{F}[z^{(k)}(\tau_j)]| \right\|_F^2 + w_j^{(k)} \|z^{(k)}(\tau_j)\|_{\ell}, \quad (22)$$

with $z^{(k)}(\tau_j) = \mathbf{T}_{\ell}(\tau_j, \partial_j \rho^{(k)} - \lambda_j^{(k)})$. The two spatial terms are then combined to define

$$\psi^{(k)}(\tau_x, \tau_y) = \sum_{j \in \{x,y\}} \psi_j(u_j^{(k)}, \rho^{(k)}, \lambda_j^{(k)}, \tau_j, \ell). \quad (23)$$

The weighting parameter $w_j^{(k)} \geq 0$ controls the relative scaling of the first term in Eq. (22) with the second and is defined as

$$w_j^{(k)} = \varsigma \frac{\left\| |\tilde{E}_j| \odot \sqrt{D} - |\mathcal{F}[u_j^{(k)}]| \right\|_F^2}{\|u_j^{(k)}\|_{\ell}}, \quad (24)$$

where $\varsigma \geq 0$ is a parameter balancing the sparsity of the solution with its Fourier modulus measurement agreement. If ς is too large, then the sparsity-inducing term can dominate in Eq. (22), resulting in a solution that has a large mismatch with the $|\tilde{E}_j| \odot \sqrt{D}$ term. If ς is too small, we are allowing nonsparse solutions, which can quickly cause convergence to a poor local solution similar to that seen in Sec. 2 when using the probe diameter as a support region. The choice of ς values for our numerical experiments is discussed in Sec. 5.

We can numerically evaluate Eq. (23) as a function of (τ_x, τ_y) to gain intuition on what is necessary for finding a minimizer of Eq. (23). An example is shown in Fig. 4, where $\ell = 0$ is used; similar results are obtained if $\ell = 1$ is used. At first glance, the metric $\psi^{(k)}(\tau_x, \tau_y)$ appears to have a well-behaved landscape with a single global minimum, as seen in Fig. 4a. By zooming in on a smaller field of view in Fig. 4b (which is enclosed by the boundaries of the red box in

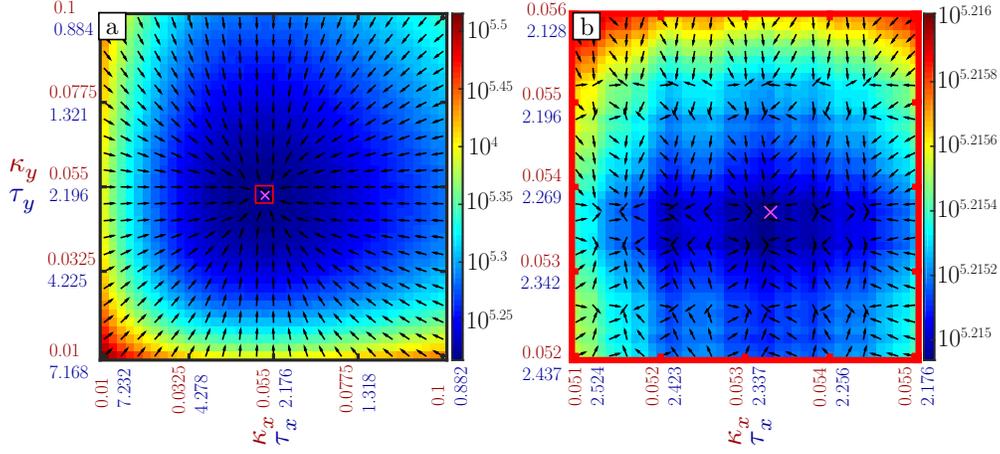


Fig. 4. At an illustrative iteration k : (a) Numerical evaluation of $\psi^{(k)}(\tau_x, \tau_y)$ over the box defined by $(\tau_x, \tau_y) = (0.882, 0.884)$ to $(7.232, 7.168)$, which corresponds to the sparsity ratio box $(\kappa_x, \kappa_y) = (0.01, 0.01)$ to $(0.1, 0.1)$. At this scale in (τ_x, τ_y) , $\psi^{(k)}$ looks well behaved. A single global minimum is indicated by the magenta \times marker; the black arrows denote the gradient (rescaled to be a unit vector) of $\psi^{(k)}(\tau_x, \tau_y)$ with respect to (τ_x, τ_y) . (b) Evaluation of $\psi^{(k)}(\tau_x, \tau_y)$ in the red box subregion shown in (a). At this scale in (τ_x, τ_y) , multiple local minima and discontinuity of the derivatives become apparent.

Fig. 4a), however, we begin to see multiple local minima. This discontinuous and nonsmooth behavior of the derivatives of $\psi^{(k)}(\tau_x, \tau_y)$ is due to the nonsmooth nature of the thresholding operators in Eq. (15) and Eq. (16).

As a result of this nonsmooth topology with multiple local minima, if we attempt to use a gradient descent method, we can expect to quickly become stuck in one of these local solutions. One way around this problem is to use finite differences to approximate the derivatives with respect to (τ_x, τ_y) and ensure that the finite-difference steps are large enough to smooth over these spurious local solutions [26]. The larger the finite-difference steps, however, the less accurate is the approximate global minimizer that can be obtained. Consequently, instead of using a gradient descent method, we use a “polling step”-like method whereby we simply test some thresholding parameter trial values around the vicinity of the current $\tau_j^{(k)}$ in the objective function $\psi^{(k)}(\tau_x, \tau_y)$:

$$h_x^{(k)} = \arg \min_{h_x \in \{0, h_{\pm 1x}, h_{\pm 2x}, \dots\}} \psi^{(k)}(\tau_x^{(k)} + h_x, \tau_y^{(k)}) \quad (25a)$$

$$h_y^{(k)} = \arg \min_{h_y \in \{0, h_{\pm 1y}, h_{\pm 2y}, \dots\}} \psi^{(k)}(\tau_x^{(k+1)}, \tau_y^{(k)} + h_y), \quad (25b)$$

where any ties are broken lexicographically (i.e., the first tied index in the order in which the trial values of h_j are tested). The $h_{\pm nj}$, for $j \in \{x, y\}$ and $n \in \mathbb{Z}_+$, can be chosen by picking a range of sparsity ratios (e.g., $\kappa_j^{(k+1)} \in \{\kappa_j^{(k)}, \kappa_j^{(k)} \pm 0.005, \kappa_j^{(k)} \pm 0.01, \kappa_j^{(k)} \pm 0.015, \kappa_j^{(k)} \pm 0.02\}$),

with $0 < \kappa_j^{(k+1)} \leq 1$). We then can compute² the thresholding parameters corresponding to these allowable changes in sparsity ratios to get the $h_{\pm nj}$ values, and we solve Eq. (25) by direct evaluation of Eq. (23) using these $h_{\pm nj}$.

We found by experience that we need not update the thresholding parameters τ_j with the same frequency at which u_j , ρ , and λ_j are updated. If we attempt to update τ_j as frequently as the main iterates, we usually find that a polling step of $h_j = 0$ is the minimizer of $\psi^{(k)}(\tau_x, \tau_y)$ in both x and y . Therefore we introduce a parameter L that controls the frequency at which τ_j is updated to arrive at the update rule:

$$\tau_j^{(k+1)} = \begin{cases} \tau_j^{(k)} + h_j^{(k)} & \text{if } k \bmod L = 0 \\ \tau_j^{(k)} & \text{if } k \bmod L \neq 0 \end{cases} \quad \text{for } j \in \{x, y\}, \quad (26)$$

with the $h_j^{(k)}$ value computed in Eq. (25).

5. Benchmarking

Given initializations for $\rho^{(0)}$, $u_j^{(0)} = \partial_j \rho^{(0)}$, $\lambda^{(0)}$, and $\tau_j^{(0)}$, our complete algorithm to find the sparsest possible $u_j = \partial_j \rho$ subject to the constraint $\rho \in \mathcal{M}$ is

$$\tau_j^{(k+1)} = \begin{cases} \tau_j^{(k)} + h_j^{(k)} & \text{if } k \bmod L = 0 \\ \tau_j^{(k)} & \text{if } k \bmod L \neq 0 \end{cases} \quad \text{for } j \in \{x, y\} \quad (27a)$$

$$u_j^{(k+1)} = \mathbf{T}_\ell(\tau_j^{(k+1)}, \partial_j \rho^{(k)} - \lambda_j^{(k)}) \quad \text{for } j \in \{x, y\} \quad (27b)$$

$$\rho^{(k+1)} = \Pi_{\mathcal{M}} \left[\mathcal{F}^{-1} \left[\frac{\tilde{E}_x^* \odot \tilde{u}_x^{(k+1)} + \tilde{E}_x^* \odot \tilde{\lambda}_x^{(k)} + \tilde{E}_y^* \odot \tilde{u}_y^{(k+1)} + \tilde{E}_y^* \odot \tilde{\lambda}_y^{(k)}}{|\tilde{E}_y|^2 + |\tilde{E}_x|^2} \right] \right] \quad (27c)$$

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + \beta_j (u_j^{(k+1)} - \partial_j \rho^{(k+1)}) \quad \text{for } j \in \{x, y\}. \quad (27d)$$

We test this algorithm on 14 problems with varying sparsity ratios for the supports \mathbb{I}_{S_j} (as defined in Eq. (6)) of the edge-detected representations $\partial_j \rho$, with $j \in \{x, y\}$. The test problems have sparsity ratios for the (assumed unknown) \mathbb{I}_{S_j} ranging from approximately 0.005 to 0.037, as shown in Fig. 5a; some representative \mathbb{I}_{S_j} are shown in Fig. 5b, 5c, and 5d. The simulated ground truth exit waves corresponding to these sparsity ratios for the various \mathbb{I}_{S_j} are defined by using the probe function seen in Fig. 2a and from a set of sample transmission functions generated by using periodic Voronoi patterns with varying polygon centroid density in a $M \times N = 512 \times 512$ array size. The contrast of the modulus of the sample transmission functions is allowed to range between 0.1 and 0.9, while the phase can range between $\pm\pi$. Examples of these exit waves corresponding to the supports in Fig. 5e, 5f, and 5g, respectively, are shown in Fig. 5b, 5c, and 5d. By defining synthetic test problems in this way, we can systematically control the sparsity ratios for the \mathbb{I}_{S_j} in order to test performance.

The choice of $\ell \in \{0, 1\}$ is application, problem, and algorithm dependent. Therefore, in Sec. 5.1 we discuss our choice of preferred ℓ . Then in Sec. 5.2 we provide numerical results from the benchmark problems with $\ell = 0$. Section 5.3 discusses the effects of noise in the diffraction on our method.

5.1. Use of $\ell = 0$ vs. $\ell = 1$

We examine the differences in performance of our ADMM method (Eq. (27)) when faced with choosing between $\ell = 0$ or $\ell = 1$ as sparsity-inducing metrics. The $\ell = 1$ regularization choice is

²By sorting the MN values in $|\partial_j \rho^{(k)} - \lambda_j^{(k)}|$ and setting τ_j to be the $\lfloor MN\kappa_j \rfloor$ th largest value.

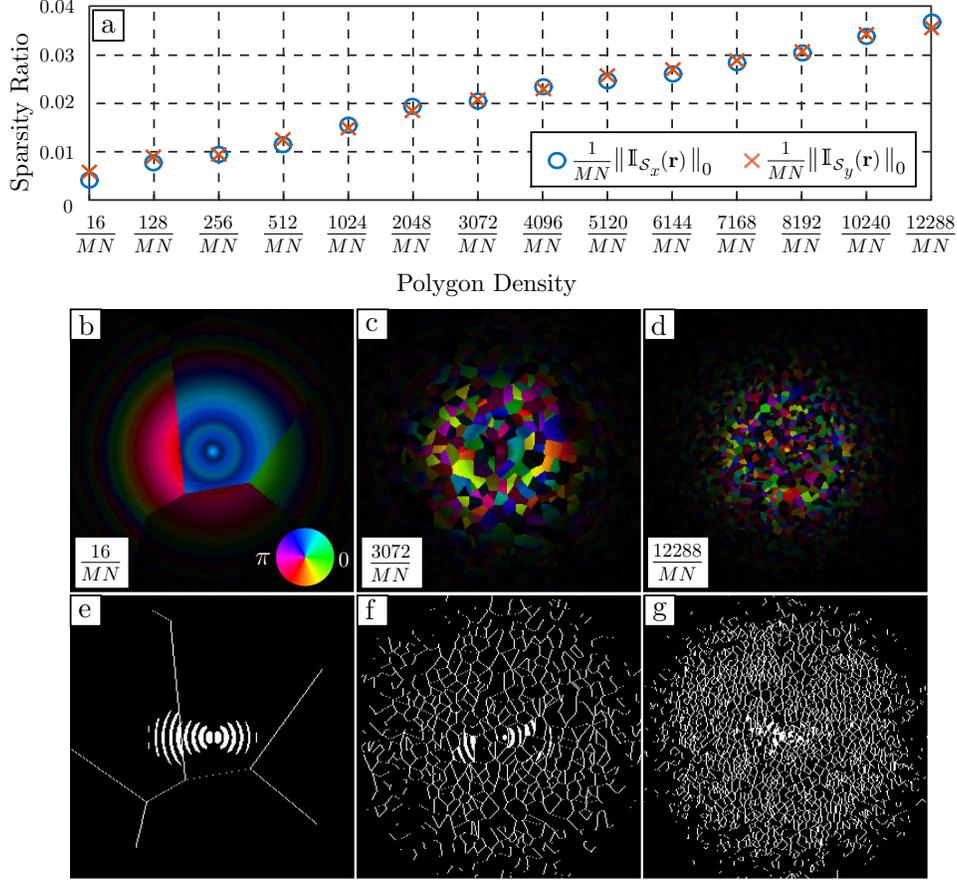


Fig. 5. (a) Sparsity ratios for the sparse representation $\partial_j \rho$ for $j \in \{x, y\}$ of the test problems we benchmark to evaluate performance of Eq. (27). To illustrate what is changing in (a), the exit wave ρ_{true} for the polygon densities $\frac{16}{MN}$, $\frac{3072}{MN}$, and $\frac{12288}{MN}$ are shown in (b), (c), and (d), respectively. The supports on the sparse representation $\partial_x \rho_{\text{true}}$ (i.e., \mathbb{I}_{S_x}) for each of these respective cases are shown in (e), (f), and (g). The probe function used to generate the exit waves is the same used in Fig. 2a, and for all the transmission functions corresponding to these exit waves the modulus contrast varies between 0.1 and 0.9, while the phase contrast varies between $\pm\pi$.

widely used as a stand-in for $\ell = 0$ regularization because using $\ell = 0$ introduces difficulties in analyzing convergence of algorithms derived from its use. Some studies [27, 28, 29], however, have noted difficulties when using $\ell = 1$ with sparse image restoration and recovery. To look at the performance of the update rules given in Eqs. (27b–27d), we attempt to recover the simulated exit wave shown in Fig. 3a by sampling over a mesh of fixed sparsity ratios (κ_x, κ_y) for some number of total iterations K and to determine the final error by using the metric in Eq. (3). We do this procedure for both $\ell = 0$ and $\ell = 1$, meaning we use either the soft-thresholding operator given in Eq. (15) or the hard-thresholding operator given in Eq. (16) to update the u_j in Eq. (27b).

In Fig. 6, we show the resulting final error using the updates in Eqs. (27b–27d) for both $\ell = 0$ and $\ell = 1$ over a fixed mesh for (κ_x, κ_y) . For both, a 21×21 (equally spaced in both parameters) box is used, ranging from $(\kappa_x, \kappa_y) = (0.005, 0.005)$ to $(0.255, 0.255)$ for $\ell = 0$ (hard

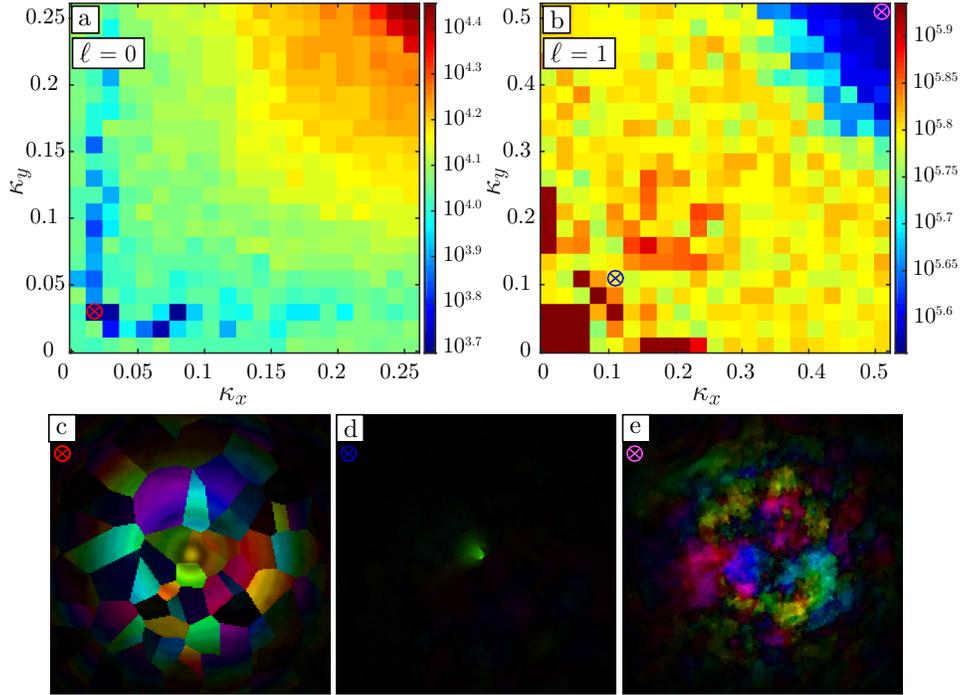


Fig. 6. (a-b) Fourier modulus measurement metric Eq. (3) vs a mesh of fixed sparsity ratios (κ_x, κ_y) . (a) Use of hard thresholding and the ADMM updates over the box of fixed sparsity ratios from $(0.005, 0.005)$ to $(0.255, 0.255)$. In this case, a minimum (located by the magenta circle) is clearly defined at a small highly constraining sparsity ratio pair allowing prompt convergence to the ground truth exit wave ρ_{true} . (b) Use of soft thresholding and the ADMM updates over the box of fixed sparsity ratios from $(0.005, 0.005)$ to $(0.505, 0.505)$. (c) Recovered exit wave using hard thresholding corresponding to the minimum (red circle/cross) in (a) at sparsity ratio pair $(\kappa_x, \kappa_y) = (0.0175, 0.03)$. (d) Recovered exit wave using soft thresholding corresponding to the minimum (blue circle/cross) in (a) at sparsity ratio pair $(\kappa_x, \kappa_y) = (0.105, 0.105)$ (e) Recovered exit wave using soft thresholding corresponding to the minimum (magenta circle/cross) in (a) at sparsity ratio pair $(\kappa_x, \kappa_y) = (0.505, 0.505)$.

thresholding) and from $(\kappa_x, \kappa_y) = (0.005, 0.005)$ to $(0.505, 0.505)$ for $\ell = 1$ (soft thresholding). For each of the 441 (κ_x, κ_y) pairs, $K = 4 \times 10^4$ such iterations are performed. After these K iterations, we start from the prescribed (κ_x, κ_y) values and linearly ramp up (κ_x, κ_y) over the next 10^4 iterations to a final sparsity ratio³ of $(\kappa_x, \kappa_y)_{\text{final}} = (0.255, 0.255)$ (hard thresholding) or $(\kappa_x, \kappa_y)_{\text{final}} = (0.505, 0.505)$ (soft thresholding). The purpose of this ramping up of sparsity ratios is to allow the binary-valued support regions, generated by the thresholding operations on the u_j , to grow larger and larger (less sparse), which enables the ADMM method to converge to the nearest stationary point defined by Eq. (11). In our experience, we observed that the solution quality, when compared with the measurement \sqrt{D} , as well as the ground truth exit wave ρ_{true} , is worse if the ramping is not performed; that is, a larger value is obtained when using the error metric given by Eq. (3) as well as when using the metric $\|\rho_{\text{true}} - \rho\|_F^2$, after subpixel image

³For the extreme values of this box in both cases, no ramping is performed.

registration and global phase offset correction on ρ are performed.

Figures 6a and 6b show that the use of hard thresholding is much more effective (has much smaller residual error) compared with the use of soft thresholding in the ADMM method in Eqs. (27b–27d). In Fig. 6a the minimum indicated by a red circle and cross corresponds to a sparsity ratio pair of $(\kappa_x, \kappa_y) = (0.0175, 0.03)$. The recovered exit wave after ramping up to $(\kappa_x, \kappa_y)_{\text{final}} = (0.255, 0.255)$ is shown in Fig. 6c and is—apart from 180° rotation symmetry, complex conjugation, and global phase offset on ρ —closely matched to the ground truth exit wave ρ_{true} shown in Fig. 3a. A typical exit wave reconstruction using soft thresholding is shown in Fig. 6d, which corresponds to the blue circle/cross at $(\kappa_x, \kappa_y) = (0.105, 0.105)$; here the use of soft thresholding appears to cause the exit wave to be dominated by relatively few image pixels, which have much larger moduli values than all other image pixels. Similar effects have been reported in wavelet-based image restoration for other highly ill-posed problems when using $\ell = 1$ sparsity-inducing regularization [27]. Other studies have shown that using $\ell = 0$ regularization leads to better edge preservation in image restoration (exactly the features we want to take advantage of) when compared with $\ell = 1$ regularization [28, 29]. The minimum of Fig. 6b is located at the magenta circle/cross at $(\kappa_x, \kappa_y) = (0.505, 0.505)$, and the corresponding exit wave reconstruction is shown in Fig. 6e; here the sparsity ratio appears to be too large (too weakly constraining), causing quick stagnation to a local solution in a fashion similar to that seen when using the approximate beam diameter as a support in Sec. 2.

As a result of this clear superiority of hard thresholding when used with Eq. (27b) compared with the use of soft thresholding, for the remainder of this paper we will consider the use only of hard thresholding. We note that other algorithms exist that successfully use soft thresholding to promote sparsity [30, 25, 31]. We further note as a reminder that the use of soft thresholding comes from $\ell = 1$ regularization as a sparsity-inducing term and that its use is as a convex and continuous relaxation of $\ell = 0$ regularization in order to make the problem simpler to analyze; both regularization terms are effective at promoting sparsity to varying degrees depending on the particulars of the phase retrieval algorithm used and type of sparsifying transformation. We believe that the use of weighted $\ell = 1$ regularization methods [19] should be explored; these methods use prior iterations of the exit wave ρ and a modified $\ell = 1$ regularization term that can closely approximate the $\ell = 0$ regularization term. These weighted $\ell = 1$ regularization methods also retain the ease of analytic computations, continuity, and convexity. To address the use of $\ell = 0$ versus $\ell = 1$ regularization to promote sparsity in the phase retrieval problem, researchers need to carry out further systematic benchmarking and comparison of these methods in order to compare relative effectiveness at successfully recovering the exit wave and their sparse edge-detected representations, but these activities are beyond the scope of this paper.

5.2. Benchmark Results for $\ell = 0$

We now provide full benchmark results on our test problem with $\ell = 0$. We use $\zeta \simeq 0.1$ when using Eq. (24). Its value was determined by numerical experimentation and works well with the test problems we consider in this section. For other synthetic test problems or if attempting phase retrieval on experimental data, other values for ζ may perform better, and further exploratory tuning of this parameter should be performed.

We run 50 independent trials, each with a different $M \times N$ complex-valued uniform random number array for $\rho^{(0)}$ for each of the 14 synthetic test problems with varying polygon density. We initialize $\lambda^{(0)}$ to be the $M \times N$ array of zeros, and we use as initial sparsity ratios $(\kappa_x^{(0)}, \kappa_y^{(0)}) = (0.05, 0.05)$, with initial thresholding parameters $(\tau_x^{(0)}, \tau_y^{(0)})$ computed from the initial $(\kappa_x^{(0)}, \kappa_y^{(0)})$ and $\partial_j \rho^{(0)}$. We update the thresholding parameters τ_j every $L = 50$ iterations and use 5 polling steps in both x and y with allowable sparsity ratio changes of $\kappa_j^{(k+1)} \in \{\kappa_j^{(k)}, \kappa_j^{(k)} \pm 0.005, \kappa_j^{(k)} \pm 0.01\}$. We allow $K_0 = 2 \times 10^4$ iterations of

Eq. (27) to be performed, after which we cease using Eq. (27a) to update the thresholding parameters τ_j . Beginning from the current $(\kappa_x^{(K_0)}, \kappa_y^{(K_0)})$, we ramp up the sparsity ratios to final values of $(\kappa_x^{(K_f)}, \kappa_y^{(K_f)}) = (0.25, 0.25)$ in the same manner as that discussed in Sec. 5.1, with $K_f = 2.5 \times 10^4$. results are shown in Fig. 7a, where for each of the 50 trials for each test problem polygon density, we sort the final values of the metric $\epsilon_{\text{true}}^2 = \|\rho_{\text{true}} - \rho^{(K_f)}\|_F^2$ in descending order. We have also corrected the recovered $\rho^{(K_f)}$ for subpixel shift and global phase offsets when compared with ρ_{true} when computing ϵ_{true}^2 .

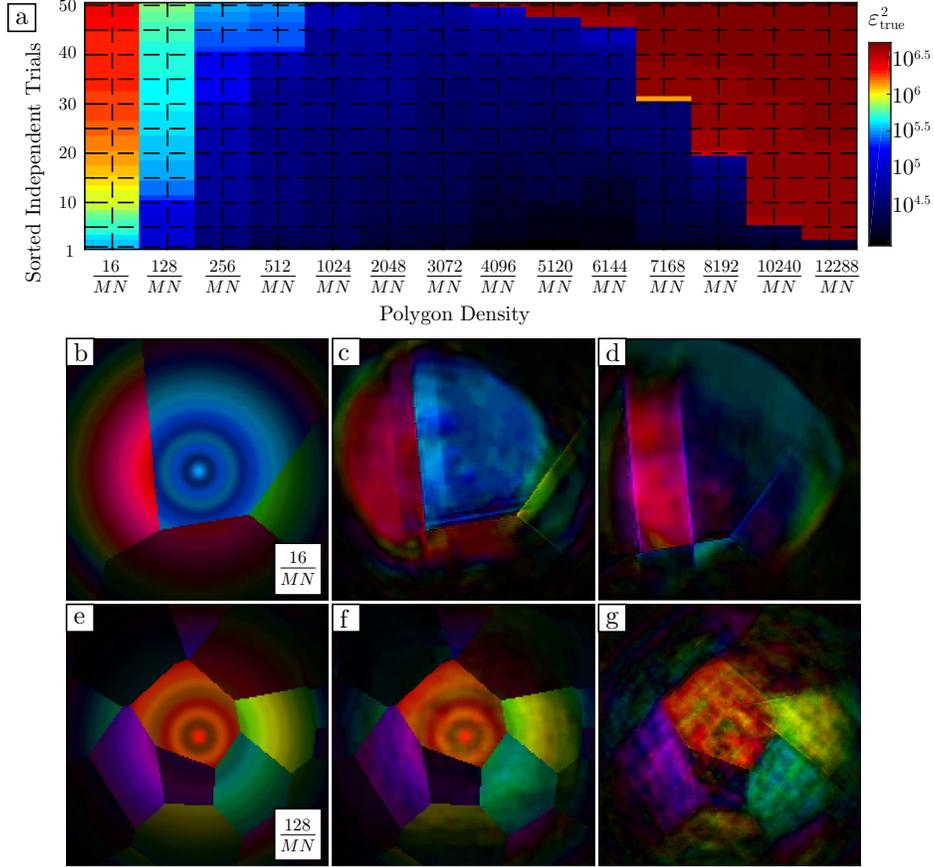


Fig. 7. (a) Sorted $\epsilon_{\text{true}}^2 = \|\rho_{\text{true}} - \rho^{(k)}\|_F^2$ values vs different sparsity ratios in $\partial_j \rho$ for $j \in \{x, y\}$ after $k = 2.5 \times 10^4$ iterations and using 50 independent trials with different starting random exit waves. For $\frac{16}{MN}$, (b) the ρ_{true} , (c) $\rho^{(k)}$ with lowest ϵ_{true}^2 value, and (d) $\rho^{(k)}$ with largest ϵ_{true}^2 value. For $\frac{128}{MN}$, (e) the ρ_{true} and (f) $\rho^{(k)}$ with lowest ϵ_{true}^2 value, and (g) $\rho^{(k)}$ with largest ϵ_{true}^2 value.

The results of these benchmarks give us some insight into the performance of Eq. (27) over a range of synthetic test problems of varying sparsity upon the forward-difference edge detection in Eq. (4). Fig. 7a shows three regions of notable and differing behavior in ϵ_{true}^2 . The first is an “ambiguous” region, where results are inconclusive and are between success and failure to varying degrees. The reason for the difficulty in this region lies in our attempt to use edge detection to reconstruct very smoothly varying features (discussed further below). The second and third regions correspond to where, for a clear majority of starting points, we unambiguously

recover or fail to adequately recover an exit wave quantitatively and qualitatively close to the corresponding ρ_{true} . (The criteria used to determine what “adequate” means is discussed below.)

The “ambiguous” region, which we denote by R_A , corresponds to polygon densities of $\frac{p_1}{MN}$ for $p_1 \in \{16, 128, 256, 512\}$ and where $10^5 \lesssim \epsilon_{\text{true}}^2 \lesssim 10^{5.5}$. For $p_1 = 16$, the ground truth exit wave is shown in Fig. 7b; a reconstruction with $\epsilon_{\text{true}}^2 \simeq 10^{5.5}$ is shown in Fig. 7c. The reconstruction clearly has recovered the locations of the edges correctly, albeit with some numerical artifact degradation, but has not successfully recovered the smoothly varying regions on the exit wave, which look like scaled and phase-shifted regions of the probe function (again shown in Fig. 2a). For $p_1 = 128$, the ground truth exit wave is shown in Fig. 7e, and a reconstruction with $\epsilon_{\text{true}}^2 \simeq 10^5$ is shown in Fig. 7f. We see that edges are again faithfully reconstructed and that smoothly varying features of the exit wave, which look like scaled-down/phase-shifted regions of the probe function, are again degraded. However, the degradation in these smoothly varying regions in Fig. 7f is not nearly as severe when compared with the $p_1 = 16$ case shown in Fig. 7c, and only finer details of the probe are degraded.

We conclude that using forward differences is not a proper sparsifying transformation for the smooth features encountered in the test problems in the “ambiguous” region R_A , and the inability of Eq. (27) to recover these smoothly varying features using edge detection is expected. The reason is that the forward-difference edge-detected representation of the smoothly varying regions gets set to zero in the thresholding step of Eq. (27b) as these regions will likely fall below the current thresholding parameter values τ_j , thereby destroying any information in these regions. Likely what is needed to remove this inability to recover smoothly varying regions of the exit wave far from the requisite edges, which the forward-difference edge-detection scheme was *designed* to effectively recover, is to perform iterative separation of the sample transmission function and the probe, as is done in ptychography [6, 32], and to introduce additional information (constraints) on the optimization problem. For example, we can constrain further the sample transmission function using additional thresholding to promote sparsity in some edge-detected representations of the sample transmission function (this can be in addition to sparsity promotion by using forward differences of the exit wave used here), as well as using index of refraction limits [33]. We can also further constrain the reconstruction for the probe function using a white-field (sample-out) measurement, which is the modulus squared of the probe function numerically propagated to the area detector. Other regularization schemes and algorithmic modifications can be made as well, such as total variational denoising methods on the $\partial_j \rho$ (using $\partial_j^2 \rho$, that is, curvature sparsity information), which promote smoothness in an image while simultaneously maintaining edges [34].

The second region of interest, R_S , corresponds to Voronoi polygon densities of $\frac{p_2}{MN}$ for $256 \leq p_2 \leq 12288$ and where $\epsilon_{\text{true}}^2 \lesssim 10^5$. The “dark blueish to black” region R_S corresponding to $\epsilon_{\text{true}}^2 \lesssim 10^5$ in Fig. 7a is a region where we consider the recovery of the exit wave successful (example reconstructions not shown) and robust (in that virtually all starting points yield an accurate solution). One qualitative criterion we can use to define success is the following: after subpixel image registration and global phase offset correction, for particular p_2 values with $\epsilon_{\text{true}}^2 \lesssim 10^5$, the recovered exit waves and the known ground truth exit waves are virtually indistinguishable upon visual inspection. A more quantitative criterion we use to determine success here is computing the phase retrieval transfer function (PRTF; plots not shown). For a particular p_2 , we compute a single averaged reconstruction from the multiple trial runs with $\epsilon_{\text{true}}^2 \lesssim 10^5$ (to average out remnant numerical artifacts from a particular trial run). All PRTFs computed in this way yield values between $\simeq 0.7$ and 1 for all spatial frequencies. In the literature, spatial frequency regions of the PRTF falling below $\simeq 0.5$ are typically considered as unreliably recovered spatial frequency content [35, 36]. Since for the cases mentioned we have PRTFs well above this cutoff at all spatial frequencies, we can be confident these reconstructions are

successful.

The third region of interest, R_F , in Fig. 7a corresponds to the exit waves that we consider unsuccessfully recovered, using both criteria of simply comparing the reconstructions to the ground truth exit waves using visual inspection and using PRTFs. These are Voronoi polygon densities of $\frac{p_1}{MN}$ for $p_1 \in \{16, 128, 256, 512\}$ with $10^{5.5} \lesssim \epsilon_{\text{true}}^2$ and $\frac{p_3}{MN}$ for $4096 \leq p_3$ with $10^{6.4} \lesssim \epsilon_{\text{true}}^2$. A result for $p_1 = 16$ and with $\epsilon_{\text{true}}^2 \simeq 10^{6.4}$, shown in Fig. 7d, indicates that the construction has very roughly recovered the correct location of the edges but also appears to suffer greatly from other numerical artifacts common to phase retrieval due to translational Fourier transform symmetries [17], as well as failing to recover any smoothly varying features of the probe function. A result for $p_1 = 128$ and with $\epsilon_{\text{true}}^2 \simeq 10^{5.8}$ is shown in Fig. 7g; hence, edge locations are adequately if imperfectly recovered, but smoothly varying features are unacceptably degraded when compared with the ground truth exit wave in Fig. 7e as well as the “borderline” successfully recovered exit wave in Fig. 7f.

The almost exclusively “dark reddish” region in Fig. 7a for the p_3 values listed and with $10^{6.4} \lesssim \epsilon_{\text{true}}^2$ forms a sharp boundary with the R_S region discussed above, which is different from the more gradual boundary that the R_S region shares with the R_A region. The R_F region is different from the unsuccessfully recovered p_1 values in R_A in that no recognizable edges are recovered and the reconstructions look poor (similar to that seen in Fig. 2c). The reconstructions corresponding to the R_F region indicate that the support implicitly generated while evaluating Eq. (27b) behaves as if it is no longer sparse enough, which was the cause of failure when using the approximate probe function diameter as support (discussed in Sec. 2). Note, however, that for the p_3 values listed in the R_F region, there remains some probability of a successful reconstruction, ranging from 4% for $p_3 = 12288$ all the way to 96% for $p_3 = 4096$. The nature of this boundary between “sparse enough” and “not sparse enough” determining probabilistic reconstruction success using Eq. (27) is unclear, however, and requires further study.

5.3. Noise, Edge Detection, and Sparsity

We next discuss the effects of noise on sparsity and the edge detection methods used in this paper. We define as before the noise-free diffraction intensity as D and the diffraction intensity with Poisson noise as D_n . The noisy measurement set \mathcal{M}_n and projection operator $\Pi_{\mathcal{M}_n}$ are defined as in Eq. (8) and Eq. (9), but using D_n in place of D . In Fig. 8a, we compute the azimuthal average of the diffraction intensities using

$$\langle F(q_r, q_\theta) \rangle_{q_\theta} = \frac{1}{N(q_r)} \sum_{\mathbf{q}} F(q_r, q_\theta) \odot A(q_r, q_\theta) \quad (28)$$

for some $F(\mathbf{q}) = F(q_n, q_m) = F(q_r, q_\theta) \in \mathbb{R}_+^{M \times N}$, where (q_r, q_θ) are spatial frequency polar coordinates, $q_r = \sqrt{q_m^2 + q_n^2}$ and $q_\theta = \tan^{-1}(q_m/q_n)$, $A(q_r, q_\theta)$ is an $M \times N$ binary mask with ones within the annulus of width one pixel at radius q_r and zeros elsewhere, and $N(q_r) = \sum_{\mathbf{q}} A(q_r, q_\theta)$ is the number of pixels within the annulus defined by the constant radial spatial frequency q_r , which in the continuous limit is the circumference of the circle $2\pi q_r$. The azimuthal average of the noise-free diffraction intensity D is shown in Fig. 8a as the black curve, while the azimuthal average of the diffraction intensity with Poisson noise D_n is shown by the green curve; the signal at high spatial frequencies becomes lost in the noise at approximately $q_r \simeq 10^2 \Delta q_r$, where Δq_r is the spatial frequency radial pixel size in units of inverse length.

To best understand the effect noise has on edge detection using the forward difference convolution matrices in Eq. (4), we view edge detection as a spatial frequency filter in Fourier space. In Fig. 8c we show $|\mathcal{F}[E_x]| = |\tilde{E}_x|$; viewed in this way, forward-difference edge detection in the x direction can be seen as spatial high-pass filtering preferentially emphasizing high spatial frequencies in the x direction while attenuating spatial frequency content in the y direction. The

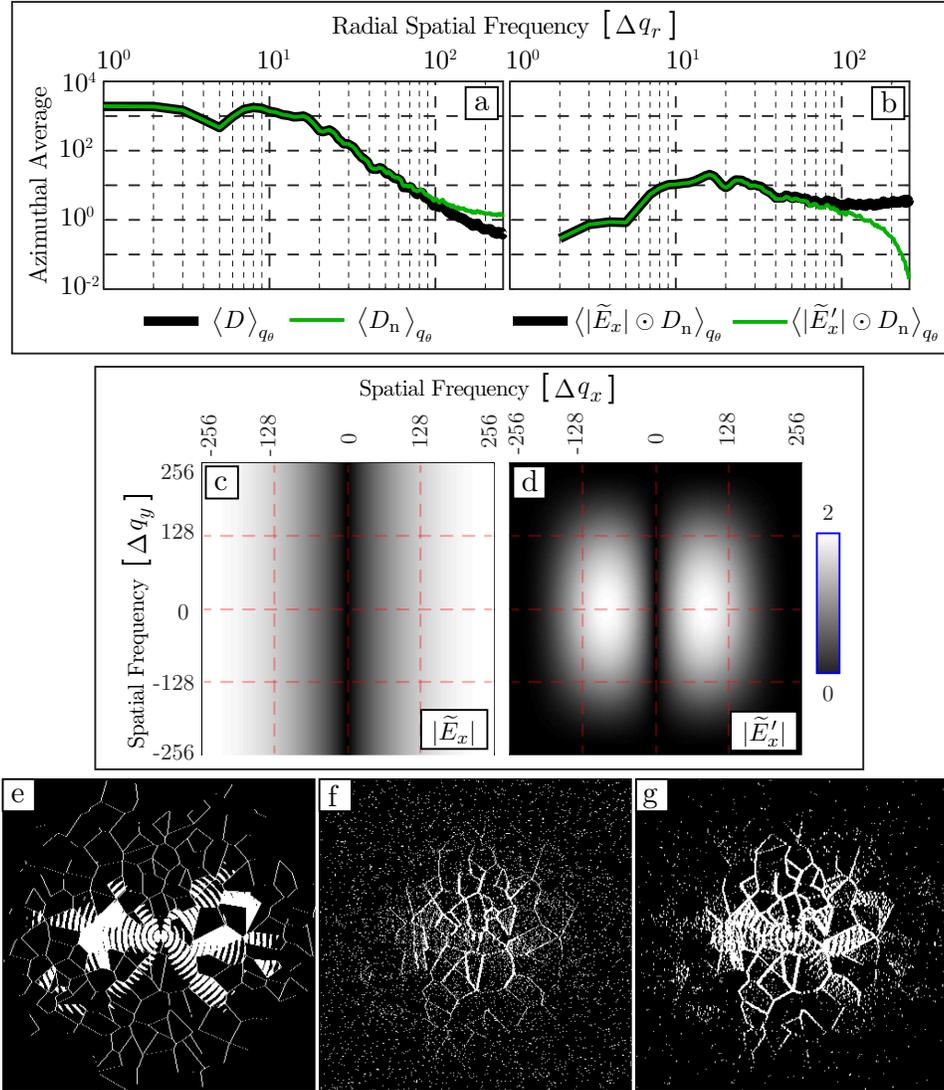


Fig. 8. Effects of Poisson degraded diffraction intensity on the sparsity of the supports for the sparse representations $\partial_j \rho$. (a) Azimuthal average of diffraction patterns without noise D (black curve) and with noise D_n (green curve). (b) Azimuthal average of the noisy diffraction pattern filtered in Fourier space using the forward difference \tilde{E}_x (black curve) and Sobel \tilde{E}'_x (green curve) edge detection convolution matrices. (c) $|\tilde{E}_x|$, (d) $|\tilde{E}'_x|$. (e) An example noise-free support for $|\partial_x \rho_{\text{true}}|$, (f) the support of $|\partial_x \Pi_{\mathcal{M}_n}[\rho_{\text{true}}]|$ using forward difference edge detection, and (g) the $|\partial_x \Pi_{\mathcal{M}_n}[\rho_{\text{true}}]|$ using Sobel edge detection. For (e-g), hard thresholding is used to generate the supports with an enforced sparsity ratio of $\kappa_x = 0.05$.

azimuthal average of $|\mathcal{F}[\partial_x \Pi_{\mathcal{M}_n}[\rho]]| = |\tilde{E}_x| \odot \sqrt{D_n}$ is shown as the black curve in Fig. 8b. An example of a support generated by hard thresholding using an enforced sparsity ratio of 5% for the noise-free case is shown in Fig. 8e, while the support generated for the noisy case, also using an enforced sparsity ratio of 5%, is shown in Fig. 8f. The noise-degraded regions at high

spatial frequencies (where the true diffraction intensity signal is essentially lost in the noise) are being amplified, giving a highly noncompact support region that in turn looks noisy, with image pixels in the support randomly distributed and inconsistent with the known ground truth support shown in Fig. 8e.

To alleviate this amplification of noise-degraded regions at high spatial frequency, we can use other types of edge detection besides forward differences. One well-known method of edge detection that has some noise suppressing properties is a Sobel filter

$$E'_x = \frac{1}{4} \begin{bmatrix} -1 & 0 & 1 & 0 & \cdots & 0 \\ -2 & 0 & 2 & 0 & \cdots & 0 \\ -1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{M \times N}, \quad (29)$$

where $E'_y = E'^T_x$ and \cdot^T is the transpose. The behavior of E'_x in Fourier space is shown in Fig. 8d, where $|\mathcal{F}[E'_x]| = |\tilde{E}'_x|$. The Sobel filter viewed in Fourier space acts as a band-pass filter, attenuating both low and high spatial frequencies (which are degraded from the noise) in the y direction. We also show the azimuthal average of $|\mathcal{F}[\partial'_x \Pi_{\mathcal{M}_n}[\rho]]| = |\tilde{E}'_x| \odot \sqrt{D_n}$ as the green curve in Fig. 8b; the attenuation of the noise-degraded regions at high spatial frequencies $q_r \geq 10^2 \Delta q_r$ is apparent. An example of a support generated by hard thresholding using an enforced sparsity ratio of 5% for the noisy case is shown in Fig. 8g. The support generated by using a Sobel filter for edge detection appears much more compact and retains features of the known ground truth in Fig. 8e better than when using forward differences.

We conclude that when noise is present in a diffraction measurement and significantly degrades particular spatial frequency content of the measurement, care must be taken to choose an appropriate edge detection method to promote sparsity. The algorithm presented in Eq. (27) can readily be modified to use thresholding on a sparse representation of the exit wave ρ using any type of edge detection, for example, explicitly defined Fourier space filters like those discussed here, as well as more sophisticated methods that combine image denoising techniques and wavelet [37], shearlet [38], or curvelets [39] transforms on the exit wave ρ . Further benchmarking and study using these other edge detection methods, using sparsity promotion through thresholding both on the modulus of the sparse representation of ρ and on the phase [30], should be explored, but this work is beyond the scope of this paper.

6. Conclusions

Developing innovative imaging methods utilizing phase retrieval is especially crucial and timely given the current planning and commissioning of high brightness coherent x-ray sources across the world. Especially for x-ray free electron lasers, the experimental arrangement explored in this work will be highly convenient and natural to use. New methods such as those presented here will help to unlock the full potential of these sources by allowing for robust phase retrieval under experimental conditions currently thought to be infeasible and/or too difficult to work with.

We show a way forward for robust single-view CDI of extended samples probed with plane wave illumination using phase retrieval. This case, which is particularly relevant to flash imaging by x-ray free-electron laser sources and for study of materials and life sciences specimens when scanning approaches (ptychography) are impractical, has previously been considered intractable. Our method exploits edge detection to arrive at a sparse representation of the sample exit wave (one with far fewer dominant pixels) and uses thresholding of the modulus of this

sparse representation to implicitly generate a sparse support. This sparse support is much more highly constraining when compared to using a support on the original non-sparse exit wave, which is the essence of the method: we use phase retrieval solve for the simpler, sparser representation, not the original non-sparse exit wave.

We derive an algorithm based on ADMM to solve for the sparse edge-detected representations of the exit wave, and benchmark it on test problems with varying support sparsity (dominant nonzero pixels) in the edge-detection representation. We discuss and identify characteristics in the test problems where the method succeeds and fails, as well as a borderline region where success is ambiguous. This ambiguous region corresponds to spatial exit wave features where edges have been successfully recovered but other smoothly varying regions away from edges are not. We then examine the effects of Poisson noise when performing edge detection using the forward-differences approach and discuss how Poisson noise that is dominant at high spatial frequencies adversely affects use of thresholding to define an effective support on the edge-detected representations. We then discuss a simple example of noise suppressing edge detection using a Sobel filter, as well as some other prospects to perform edge detection in the presence of noise to generate sparse supports on the edge-detected representations.

The key insight of this work is that simple mathematical transformations can drastically simplify how the relevant information content in an image is represented. We anticipate that this general approach to constraining and regularizing the optimization problem can be extended by using other types of sparsity-inducing transformations, such as wavelets, curvelets, and shearlets. In addition to enabling imaging extended samples by single-view CDI, we expect that incorporation of this concept into multiple-view (e.g., ptychographic) algorithms will be beneficial by supplementing ptychographic overlap constraints with a sparse support constraint on the edge-detected representation of the current view.

Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research under Contract No. DE-AC02-06CH11357. The authors are grateful to Sven Leyffer for insightful discussions.

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.