Phase Retrieval Using Double Deep Image Priors

Anonymous Author(s)
Affiliation
Address
email

Abstract

Phase retrieval (PR) concerns the recovery of complex phases from complex magnitudes. We identify the connection between the difficulty level and the number and variety of symmetries in PR problems. We focus on the most difficult far-field PR (FFPR), and propose a novel method using double deep image priors. In realistic evaluation, our method outperforms all competing methods by large margins. As a single-instance method, our method requires no training data and minimal hyperparameter tuning, and hence enjoys good practicality.

1 Introduction

In scientific imaging, observable physical quantities about the object of interest are often complex-valued, e.g., when diffraction happens [1]. However, practical detectors can only record complex magnitudes, but not phases, resulting in phaseless observations. Phase retrieval (PR), broadly defined, is the nonlinear inverse problem of estimating the object of interest from the phaseless observations. PR is central to coherent diffraction imaging ((B)CDI) [2, 3], image-based wavefront sensing [4], radar and sonar sensing [5]; see the recent survey [6].

Which phase retrieval (PR)? Consider a 2D object of interest $\boldsymbol{X} \in \mathbb{C}^{m \times n}$, and a physical observation model $\mathcal{A}$ that leads to an ideal complex-valued observation $\mathcal{A}(\boldsymbol{X}) \in \mathbb{C}^{m' \times n'}$. However, the detector can only record $\boldsymbol{Y} = |\mathcal{A}(\boldsymbol{X})|^2$, where $|\cdot|^2$ denotes the elementwise squared magnitudes. In **far-field (Fraunhofer) PR (FFPR) that stems from far-field propagation and is also the focus of this paper**, $\mathcal{A}$ is the oversampled 2D Fourier transform $\mathcal{F}$ with $m' \geq 2m-1$ and $n' \geq 2n-1$ to ensure recoverability. Numerous other $\mathcal{A}$'s have been studied in the literature, notable ones including: **(1) Generalized PR (GPR)**: $\mathcal{A}(\boldsymbol{X}) = \{\langle \boldsymbol{A}_i, \boldsymbol{X} \rangle\}_{i=1}^{k}$ where $\boldsymbol{A}_i$'s are iid Gaussian or randomly-masked Fourier basis matrices [7, 8]. These elegant mathematical models do not correspond to physically feasible imaging systems so far; **(2) Near-Field (Fresnel) PR (NFPR)**: $\mathcal{A}(\boldsymbol{X}) = \mathcal{F}(\boldsymbol{X} \odot [e^{i\pi\beta(i^2+j^2)}]_{i,j})$ [9, 10], where the constant $\beta > 0$ depends on the sampling intervals, wavelength, and imaging distance [11], comes from near field propagation. Note that FFPR corresponds to $\beta \to 0$, and PR problems solved in image-based wavefront sensing for astronomical applications correspond to multi-plane near-field propagation with sequential optical aberrations [12]; **(3) Holographical PR (HPR)**: $\mathcal{A}(\boldsymbol{X}) = \mathcal{F}([\boldsymbol{X}, \boldsymbol{R}])$, where $\boldsymbol{R}$ is a known reference that is put side-by-side with the object of interest $\boldsymbol{X}$ [13]; depending on the propagation distance, near-field versions are also possible [1, Chapter 11]; **(4) Ptychography (PTY)**: $\boldsymbol{X}$ is raster scanned by a sharp illumination pattern $\boldsymbol{W}$ that is focused over a local patch of $\boldsymbol{X}$ each time. Now $\boldsymbol{Y}$ is the set of magnitude measurements $\boldsymbol{Y}_i = |\mathcal{F}(\boldsymbol{W} \odot \boldsymbol{X}(\boldsymbol{p}_i))|^2$, where $\boldsymbol{p}_i$ indexes the raster grid [14, 15].

Symmetry matters Identifiability in PR is often up to intrinsic symmetries. For example, any global phase factor $e^{i\theta}$ added to $\boldsymbol{X}$ leaves $\boldsymbol{Y}$ unchanged for FFPR, NFPR, GPR, and PTY, i.e., **global phase symmetry**. While this is the only symmetry for NFPR, GPR, and PTY, FFPR has

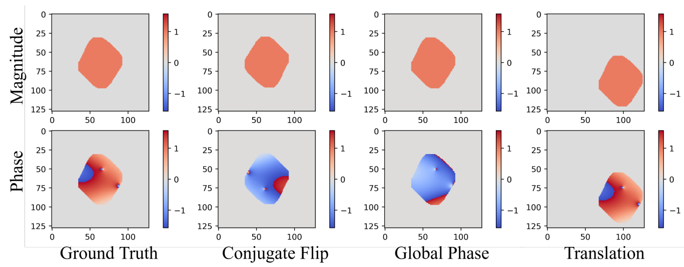
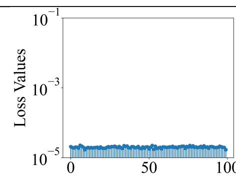
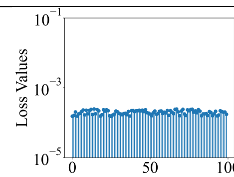
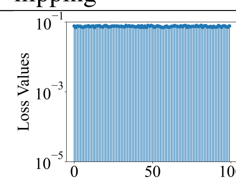


Figure 1: Illustration of the three intrinsic symmetries in FFPR on simulated complex-valued crystal data (see Section 4 for details). Any composition of 2D conjugate flipping, translation, and global phase, when applied to $\boldsymbol{X}$, leads to the same set of magnitudes $\boldsymbol{Y}$.

Table 1: Comparison of GPR, NFPR, and FFPR in terms of their symmetries and numerical solvability with the least-squares (LS) formulation combined with gradient descent.

PR model	GPR	NFPR	FFPR
Symmetry	global phase	global phase	global phase, translation, 2D conjugate flipping
Final loss of solving LS using gradient descent from 100 random initializations			

two other symmetries: translation and 2D conjugate flipping, as shown in Fig. 1 [16]. A crucial empirical observation is that **the difficulty level of a PR problem is proportional to the number of its symmetries**. To see the point, consider a natural least-squares (LS) formulation of PR: $\min_{\boldsymbol{X}\in\mathbb{C}^{m\times n}} \frac{1}{m'n'}\|\sqrt{\boldsymbol{Y}} - |\mathcal{A}(\boldsymbol{X})|\|_F^2$, with the groundtruth complex-valued 2D crystal sample in Fig. 1 as the target $\boldsymbol{X}$. On GPR with Gaussian, NFPR, and FFPR, we run gradient descent with 100 random starts respectively and record their final convergent losses. As evident from Table 1, while we can consistently find numerically satisfactory solutions for GPR and NFPR, we always find bad local solutions for FFPR—which has three symmetries. Similarly, for FFPR, the gold-standard hybrid input-output (HIO) algorithm can typically solve the problem when provided with tight support specification—translation symmetry is killed, but HIO fails when the support is loose—translation symmetry remains; see Appendix A. Moreover, prior works [17–20] also show the learning difficulties caused by these symmetries when one develops data-driven methods for solving FFPR.

Our focus on practical FFPR methods We have stressed that symmetries largely determine the difficulty level of PR. However, in previous research, there are often simplifications to FFPR, including **(1) randomized the model** $\mathcal{A}$ that only keeps the global phase symmetry, **(2) evaluation on natural images** that removes the translation symmetry and simplifies the global phase symmetry into sign symmetry [8, 21, 22]. These simplifications invariably lead to FFPR methods that do not work on practical data. **The goal of this paper is to develop practical methods for FFPR that involve all three symmetries**. In particular, we propose a novel FFPR method based on double deep image priors (see Section 3), and validate its superiority over state-of-the-art (SOTA) on realistic datasets (see Section 4).

2 FFPR: Formulation and Prior Arts

FFPR model The object of interest is $\boldsymbol{X} \in \mathbb{C}^{m\times n}$, and $\boldsymbol{Y} = |\mathcal{F}(\boldsymbol{X})|^2 \in \mathbb{R}_+^{m'\times n'}$. Here, $\mathcal{F}$ is the oversampled 2D Fourier transform. We always assume that $m' \geq 2m-1$ and $n' \geq 2n-1$, which is necessary to ensure recoverability.

Prior arts on FFPR Since we focus on practical FFPR, here we only discuss methods that have been tested on FFPR with at least partial success. **(I) Classical iterative methods**: Due to the failure of the LS, most (if not all) classical methods tackle the over-parameterized feasibility reformulation: find $\boldsymbol{Z} \in \mathbb{C}^{m'\times n'}$ s.t. $|\mathcal{F}(\boldsymbol{Z})|^2 = \boldsymbol{Y}, \mathcal{L}(\boldsymbol{Z}) = \boldsymbol{0}$, where $\mathcal{L}$ restricts $\boldsymbol{Z}$ to the zero-padding locations defined by the oversampling. **More refined support information can be naturally incorporated**

into the support constraint $\mathcal{L}(\boldsymbol{Z}) = \mathbf{0}$. These classical methods are all based on generalized alternating projection, represented by error-reduction (ER), hybrid input-output (HIO) [23], reflection average alternating reflectors (RAAR) [24], difference map (DM) [25], and oversampling smoothness (OSS) [26]. They are empirically observed to find good solutions for FFPR, provided that the support specification for $\boldsymbol{Z}$ is tight and hyperparameters are properly tuned. Alternative formulations solved by second-order methods [27, 28] are less sensitive to hyperparameters. However, all these methods require tight support specification to avoid the translation symmetry—failing so leads to spurious solutions that look like the superposition of translated copies; see Appendix B. This is addressed by the popular **shrinkwrap** trick [29] in practice, which refines the support by smoothing-and-thresholding over iterations. **(II) Data-driven methods**: **The first line of work** represents the inverse mapping from $\boldsymbol{Y}$ to $\boldsymbol{X}$ by a deep neural network (DNN) $g_{\boldsymbol{\theta}}$, which is trained either over an extensive training set $\{(\boldsymbol{Y}_i, \boldsymbol{X}_i)\}_i$, or unpaired $\{\boldsymbol{Y}_i\}_i$ and $\{\boldsymbol{X}_i\}_i$ only by observing the cycle consistency constraint: $|\mathcal{F}(g_{\boldsymbol{\theta}}(\boldsymbol{Y}))|^2 \approx \boldsymbol{Y}$ [30–35, 17–20, 36, 37]. But, as discussed in [17–20], symmetries in the problem cause substantial learning difficulties, as any $\boldsymbol{Y}$ maps to a family of equivalent $\boldsymbol{X}$'s. **The second line of work** [38–40] is tied to specific iterative methods for solving FFPR and replaces certain components of these methods with trainable DNNs. A common limitation of this line is the reliance on good initialization that is obtained from classical iterative methods. Therefore, this family can be viewed as a final refinement of the results from classical methods and does not address the essential difficulty of solving FFPR. **Both lines** suffer in generalization when the training data are not sufficiently representative.

Our method overcomes the limitations of both classical and data-driven methods. (1) **No training set**: it works with a single problem instance each time, with zero extra training data; (2) **No shrinkwrap**: we can specify the size of $\boldsymbol{X}$ directly as $\lfloor m'/2 \rfloor \times \lfloor n'/2 \rfloor$, i.e., the information-theoretic recovery limit, without worrying about the translation symmetry; (3) **Minimal tuning**: mostly we only need to tune 2 learning rates as hyperparameters, vs. the 5 or 6 hyperparameters used in HIO+ER+Shrinkwrap (HES) for practical CDI [41].

3 Our method: FFPR using double DIPs

Deep image prior (DIP) for visual inverse problems DIP and variants [42] parameterize visual objects as outputs of DNNs—typically structured convolutional networks to favor spatially smooth structures, i.e., $\boldsymbol{x} = G_{\boldsymbol{\theta}}(\boldsymbol{z})$, where $\boldsymbol{z}$ is normally a random but fixed seed, and $G_{\boldsymbol{\theta}}$ is a trainable DNN paramaterized by $\boldsymbol{\theta}$. For a visual inverse problem of the form $\boldsymbol{y} \approx f(\boldsymbol{x})$ where $\boldsymbol{y}$ is the observation and f is the observation model, the classical regularized data-fitting formulation $\min_{\boldsymbol{x}} \ \ell(\boldsymbol{y}, f(\boldsymbol{x})) + \lambda\Omega(\boldsymbol{x})$ can now be empowered by DIP and turned into $\min_{\boldsymbol{\theta}} \ \ell(\boldsymbol{y}, f \circ G_{\boldsymbol{\theta}}(\boldsymbol{z})) + \lambda\Omega \circ G_{\boldsymbol{\theta}}(\boldsymbol{z})$. This simple idea has recently claimed numerous successes in computer vision and imaging; see, e.g., [43–46]. **A salient feature of DIP is the strong structured prior it imposes through DNNs, with zero extra data**! Although the theoretical understanding of DIP is still far from complete, current theories attribute its success to two aspects: (1) **structured priors** imposed by convolutional and upsampling operations, and (2) **global optimization** due to significant overparameterization and first-order methods [47, 48].

Applying DIP to FFPR As shown in Table 1, solving the LS formulation using gradient descent always gets trapped in bad local minimizers. It is then tempting to try DIP, as (1) the objects we try to recover in scientific imaging are visual objects and probably can be blessed by the structured priors enforced by DIP, and (2) more importantly, the issue we encounter in solving the LS is exactly about global optimization, which could be eliminated by overparameterization in DIP. In fact, systematic evaluation of solving $\min_{\boldsymbol{\theta}} \ \|\sqrt{\boldsymbol{Y}} - |\mathcal{F} \circ G_{\boldsymbol{\theta}}(\boldsymbol{z})|\|_F^2$ where $G_{\boldsymbol{\theta}}(\boldsymbol{z}) \in \mathbb{C}^{m \times n}$ in Figs. 2 and 3 shows that it is already competitive compared to the gold-standard HES, although it struggles to reconstruct complicated complex phases.

Double DIPs boost the performance For FFPR applications such as CDI, $\boldsymbol{X}$ as a complex-valued object can often be naturally split into two parts with disparate complexity levels. For example, in Bragg CDI on crystals, the magnitude part on the support is known to have uniform values, but the phase part can have complex spatial patterns due to strains [49–51]; in CDI on live cells, the nonnegative real part contains useful information, and the imaginary part acts like small-magnitude noise [52]. In these cases, due to the apparent asymmetry in complexity, it makes sense to parameterize

$\boldsymbol{X}$ as two separate DIPs [53, 46] instead of one:

$$\boldsymbol{X} = G^1_{\boldsymbol{\theta}_1}(\boldsymbol{z}_1)e^{iG^2_{\boldsymbol{\theta}_2}(\boldsymbol{z}_2)}, \text{ or } \boldsymbol{X} = G^1_{\boldsymbol{\theta}_1}(\boldsymbol{z}_1) + iG^2_{\boldsymbol{\theta}_2}(\boldsymbol{z}_2). \tag{1}$$

This can be justified as balancing the learning paces: with a single DIP, “simple" part is learned much faster than the “complex” part; with double DIPs, we can balance the learning paces by making the learning rate for the “simple" part relatively small compared to that for the “complex" part. We observe a substantial performance boost in Figs. 2 and 3 due to the double-DIP parametrization.

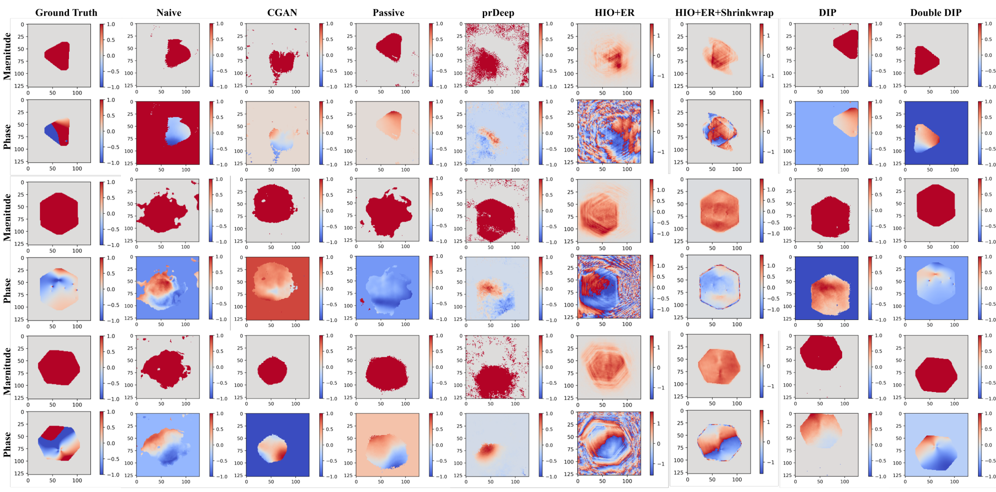


Figure 2: Visual comparison of reconstruction results by different methods on 2D crystal data

4 Experiments Results

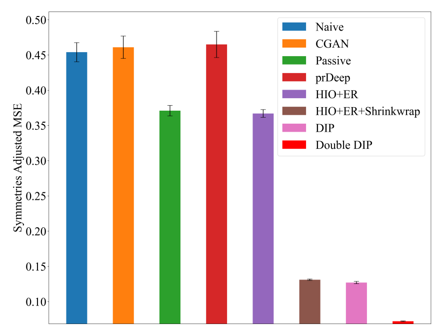


Figure 3: Quantitative comparison of reconstruction results by different methods on 2D simulated crystal data by symmetry-adjusted MSE

We first compare our Double-DIP method with multiple SOTA methods for FFPR, including Naive [36], CGAN [54], Passive [32], prDeep [38], HIO+ER, HIO+ER+Shrinkwrap (HES), and (single-)DIP on simulated 2D data for Bragg CDI on crystals. The final form of our learning objective for this task is:

$$\min_{\boldsymbol{\theta}_1,\boldsymbol{\theta}_2} \|\sqrt{\boldsymbol{Y}} - |\mathcal{F} \circ G^1_{\boldsymbol{\theta}_1}(\boldsymbol{z}_1)e^{iG^2_{\boldsymbol{\theta}_2}(\boldsymbol{z}_2)}|\|^2_F. \tag{2}$$

To ensure that the evaluation data reflect real-world complexity, we simulate 2D complex-valued crystal data in Bragg CDI applications [3]. The dataset is generated by first creating 2D convex and nonconvex shapes based on random scattering points in a 110×110 grid on a 128×128 background. The complex magnitudes are uniformly 1, and the complex phases are determined by projecting the simulated 2D displacement fields (due to crystal defects) onto the corresponding momentum transfer vectors. To maximize the diversity, the dataset contains diverse shapes and different numbers and densities of crystal defects that directly determine the complexity of the phases. **Although our double-DIP method is a single-instance method that requires no training data, the dataset is large enough to support data-driven methods, such as Passive and prDeep**. For methods that require a training set, we provide 8000 samples. All methods are tested on 50 samples. From both visual (Fig. 2) and quantitative (Fig. 3) results, it is evident that: **(1)** all data-driven methods, including Naive, CGAN, Passive, prDeep, perform poorly. We believe that this is due to either the learning difficulty caused by the three symmetries [17–20] or the bad initialization given by HIO (i.e., for prDeep); **(2)** HES, DIP, and our double-DIP are the top three methods. HES deals with translation symmetry by explicitly iterative refining the support, whereas DIP and ours do not need tight support estimation at all, substantially reducing parameter tuning; **(3)** Our method wins HES and DIP by a large margin. Although the latter two perform reasonably well in magnitude estimation, their phase estimations are typically off for complicated instances. Appendix C contains evaluation on 3D simulated Bragg CDI data.

References

[1] Joseph W. Goodman, *Introduction to Fourier Optics*, Freeman & Company, W. H., 4th edition, 2017.

[2] Jianwei Miao, Pambos Charalambous, Janos Kirz, and David Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, pp. 342–344, 1999.

[3] I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, and J. A. Pitney, "Reconstruction of the shapes of gold nanocrystals using coherent x-ray diffraction," *Physical Review Letters*, vol. 87, no. 19, pp. 195505, oct 2001.

[4] D. Russell Luke, James V. Burke, and Richard G. Lyon, "Optical wavefront reconstruction: Theory and numerical methods," *SIAM Review*, vol. 44, no. 2, pp. 169–224, jan 2002.

[5] Philippe Jaming, "Phase retrieval techniques for radar ambiguity problems," *The Journal of Fourier Analysis and Applications*, vol. 5, no. 4, pp. 309–329, jul 1999.

[6] Yoav Shechtman, Yonina C. Eldar, Oren Cohen, Henry Nicholas Chapman, Jianwei Miao, and Mordechai Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, may 2015.

[7] Emmanuel J Candes, Yonina C Eldar, Thomas Strohmer, and Vladislav Voroninski, "Phase retrieval via matrix completion," *SIAM review*, vol. 57, no. 2, pp. 225–251, 2015.

[8] Albert Fannjiang and Thomas Strohmer, "The numerics of phase retrieval," *Acta Numerica*, vol. 29, pp. 125–228, may 2020.

[9] Fei Wang, Yaoming Bian, Haichao Wang, Meng Lyu, Giancarlo Pedrini, Wolfgang Osten, George Barbastathis, and Guohai Situ, "Phase imaging with an untrained neural network," *Light: Science & Applications*, vol. 9, no. 1, may 2020.

[10] Yuhe Zhang, Mike Andreas Noack, Patrik Vagovic, Kamel Fezzaa, Francisco Garcia-Moreno, Tobias Ritschel, and Pablo Villanueva-Perez, "PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets," *Optics Express*, vol. 29, no. 13, pp. 19593, jun 2021.

[11] Tomoyoshi Shimobaba and Tomoyoshi Ito, *Computer Holography Acceleration Algorithms and Hardware Implementations*, Taylor & Francis Group, 2019.

[12] J. R. Fienup, "Phase-retrieval algorithms for a complicated optical system," *Applied Optics*, vol. 32, no. 10, pp. 1737, apr 1993.

[13] David A. Barmherzig and Ju Sun, "Towards practical holographic coherent diffraction imaging via maximum likelihood estimation," *Optics Express*, vol. 30, no. 5, pp. 6886, feb 2022.

[14] Pierre Thibault, Martin Dierolf, Andreas Menzel, Oliver Bunk, Christian David, and Franz Pfeiffer, "High-resolution scanning x-ray diffraction microscopy," *Science*, vol. 321, no. 5887, pp. 379–382, jul 2008.

[15] Stefano Marchesini, Andre Schirotzek, Chao Yang, Hau tieng Wu, and Filipe Maia, "Augmented projections for ptychographic imaging," *Inverse Problems*, vol. 29, no. 11, pp. 115009, oct 2013.

[16] Tamir Bendory, Robert Beinert, and Yonina C. Eldar, "Fourier phase retrieval: Uniqueness and algorithms," in *Compressed Sensing and its Applications*, pp. 55–91. Springer International Publishing, 2017.

[17] Kshitij Tayal, Chieh-Hsin Lai, Vipin Kumar, and Ju Sun, "Inverse problems, deep learning, and symmetry breaking," *arXiv:2003.09077*, Mar. 2020.

[18] Raunak Manekar, Kshitij Tayal, Zhong Zhuang, Chieh-Hsin Lai, Vipin Kumar, and Ju Sun, "Breaking symmetries in data-driven phase retrieval," in *OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP)*. 2021, Optica Publishing Group.

[19] Kshitij Tayal, Chieh-Hsin Lai, Raunak Manekar, Zhong Zhuang, Vipin Kumar, and Ju Sun, "Unlocking inverse problems using deep learning: Breaking symmetries in phase retrieval," in *NeurIPS 2020 Workshop on Deep Learning and Inverse Problems*, 2020.

[20] Raunak Manekar, Zhong Zhuang, Kshitij Tayal, Vipin Kumar, and Ju Sun, "Deep learning initialized phase retrieval," in *NeurIPS 2020 Workshop on Deep Learning and Inverse Problems*, 2020.

[21] D. Russell Luke, Shoham Sabach, and Marc Teboulle, "Optimization on spheres: Models and proximal algorithms with computational performance comparisons," *SIAM Journal on Mathematics of Data Science*, vol. 1, no. 3, pp. 408–445, jan 2019.

[22] S. Marchesini, "Invited article: A unified evaluation of iterative projection algorithms for phase retrieval," *Review of Scientific Instruments*, vol. 78, no. 1, pp. 011301, jan 2007.

[23] James R Fienup, "Phase retrieval algorithms: a comparison," *Applied optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[24] D Russell Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse problems*, vol. 21, no. 1, pp. 37, 2004.

[25] Veit Elser, "Phase retrieval by iterated projections," *JOSA A*, vol. 20, no. 1, pp. 40–55, 2003.

[26] Jose A Rodriguez, Rui Xu, C-C Chen, Yunfei Zou, and Jianwei Miao, "Oversampling smoothness: an effective algorithm for phase retrieval of noisy diffraction intensities," *Journal of applied crystallography*, vol. 46, no. 2, pp. 312–318, 2013.

[27] Stefano Marchesini, "Phase retrieval and saddle-point optimization," *JOSA A*, vol. 24, no. 10, pp. 3289–3296, 2007.

[28] Zhong Zhuang, Gang Wang, Yash Travadi, and Ju Sun, "Phase retrieval via second-order nonsmooth optimization," in *ICML Workshop on Beyond First-Order Methods for Machine Learning*, 2020.

[29] 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," *Physical Review B*, vol. 68, no. 14, pp. 140101, oct 2003.

[30] Ayan Sinha, Justin Lee, Shuai Li, and George Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117–1125, 2017.

[31] Alexander Scheinker and Reeju Pokharel, "Adaptive 3d convolutional neural network-based reconstruction method for 3d coherent diffraction imaging," *Journal of Applied Physics*, vol. 128, no. 18, pp. 184901, 2020.

[32] Henry Chan, Youssef SG Nashed, Saugat Kandel, Stephan O Hruszkewycz, Subramanian KRS Sankaranarayanan, Ross J Harder, and Mathew J Cherukara, "Rapid 3d nanoscale coherent imaging via physics-aware deep learning," *Applied Physics Reviews*, vol. 8, no. 2, pp. 021407, 2021.

[33] Ross Harder, "Deep neural networks in real-time coherent diffraction imaging," *IUCrJ*, vol. 8, no. Pt 1, pp. 1, 2021.

[34] Longlong Wu, Shinjae Yoo, Ana F Suzana, Tadesse A Assefa, Jiecheng Diao, Ross J Harder, Wonsuk Cha, and Ian K Robinson, "Three-dimensional coherent x-ray diffraction imaging via deep convolutional neural networks," *npj Computational Materials*, vol. 7, no. 1, pp. 1–8, 2021.

[35] Yudong Yao, Henry Chan, Subramanian Sankaranarayanan, Prasanna Balaprakash, Ross J Harder, and Mathew J Cherukara, "Autophasenn: unsupervised physics-aware deep learning of 3d nanoscale bragg coherent diffraction imaging," *npj Computational Materials*, vol. 8, no. 1, pp. 1–8, 2022.

[36] Mathew J Cherukara, Youssef SG Nashed, and Ross J Harder, "Real-time coherent diffraction inversion using deep generative networks," *Scientific reports*, vol. 8, no. 1, pp. 1–8, 2018.

[37] Yuhe Zhang, Mike Andreas Noack, Patrik Vagovic, Kamel Fezzaa, Francisco Garcia-Moreno, Tobias Ritschel, and Pablo Villanueva-Perez, "Phasegan: a deep-learning phase-retrieval approach for unpaired datasets," *Optics express*, vol. 29, no. 13, pp. 19593–19604, 2021.

[38] Christopher Metzler, Phillip Schniter, Ashok Veeraraghavan, et al., "prdeep: Robust phase retrieval with a flexible deep network," in *International Conference on Machine Learning*. PMLR, 2018, pp. 3501–3510.

[39] Çağatay Işıl, Figen S Oktem, and Aykut Koç, "Deep iterative reconstruction for phase retrieval," *Applied optics*, vol. 58, no. 20, pp. 5422–5431, 2019.

[40] Yaotian Wang, Xiaohang Sun, and Jason Fleischer, "When deep denoising meets iterative phase retrieval," in *International Conference on Machine Learning*. PMLR, 2020, pp. 10007–10017.

[41] Marcus C. Newton, Steven J. Leake, Ross Harder, and Ian K. Robinson, "Three-dimensional imaging of strain in a single ZnO nanorod," *Nature Materials*, vol. 9, no. 2, pp. 120–124, dec 2009.

[42] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky, “Deep image prior,” *International Journal of Computer Vision*, vol. 128, no. 7, pp. 1867–1888, 2020.

[43] Adnan Qayyum, Inaam Ilahi, Fahad Shamshad, Farid Boussaid, Mohammed Bennamoun, and Junaid Qadir, “Untrained neural network priors for inverse imaging problems: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[44] Taihui Li, Zhong Zhuang, Hengyue Liang, Le Peng, Hengkang Wang, and Ju Sun, “Self-validation: Early stopping for single-instance deep generative priors,” *arXiv preprint arXiv:2110.12271*, 2021.

[45] Hengkang Wang, Taihui Li, Zhong Zhuang, Tiancong Chen, Hengyue Liang, and Ju Sun, “Early stopping for deep image prior,” *arXiv preprint arXiv:2112.06074*, 2021.

[46] Zhong Zhuang, Taihui Li, Hengkang Wang, and Ju Sun, “Blind image deblurring with unknown kernel size and substantial noise,” *arXiv preprint arXiv:2208.09483*, 2022.

[47] Reinhard Heckel and Mahdi Soltanolkotabi, “Denoising and regularization via exploiting the structural bias of convolutional generators,” *arXiv preprint arXiv:1910.14634*, 2019.

[48] Reinhard Heckel and Mahdi Soltanolkotabi, “Compressive sensing with un-trained neural networks: Gradient descent finds a smooth approximation,” in *International Conference on Machine Learning*. PMLR, 2020, pp. 4149–4158.

[49] JN Clark, X Huang, R Harder, and IK Robinson, “High-resolution three-dimensional partially coherent diffraction imaging,” *Nature communications*, vol. 3, no. 1, pp. 1–6, 2012.

[50] Felix Hofmann, Edmund Tarleton, Ross J Harder, Nicholas W Phillips, Pui-Wai Ma, Jesse N Clark, Ian K Robinson, Brian Abbey, Wenjun Liu, and Christian E Beck, “3d lattice distortions and defect structures in ion-implanted nano-crystals,” *Scientific reports*, vol. 7, no. 1, pp. 1–10, 2017.

[51] David Yang, Mark T Lapington, Guanze He, Kay Song, Minyi Zhang, Clara Barker, Ross J Harder, Wonsuk Cha, Wenjun Liu, Nicholas W Phillips, et al., “Refinements for bragg coherent x-ray diffraction imaging: Electron backscatter diffraction alignment and strain field computation,” *arXiv preprint arXiv:2203.17015*, 2022.

[52] Gijs Van Der Schot, Martin Svenda, Filipe RNC Maia, Max Hantke, Daniel P DePonte, M Marvin Seibert, Andrew Aquila, Joachim Schulz, Richard Kirian, Mengning Liang, et al., “Imaging single cells in a beam of live cyanobacteria with an x-ray laser,” *Nature communications*, vol. 6, no. 1, pp. 1–9, 2015.

[53] Yosef Gandelsman, Assaf Shocher, and Michal Irani, ““double-DIP”: Unsupervised image decomposition via coupled deep-image-priors,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. jun 2019, IEEE.

[54] Tobias Uelwer, Alexander Oberstraß, and Stefan Harmeling, “Phase retrieval using conditional generative adversarial networks,” in *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021, pp. 731–738.

[55] Tomoyoshi Shimobaba and Tomoyoshi Ito, *Computer Holography Acceleration Algorithms and Hardware Implementations*, Taylor & Francis Group, 2019.

A Failure of HIO when the support specification is loose

The result is presented in Fig. 4.

B Failure patterns of classical iterative methods on FFPR

The result is presented in Fig. 5.

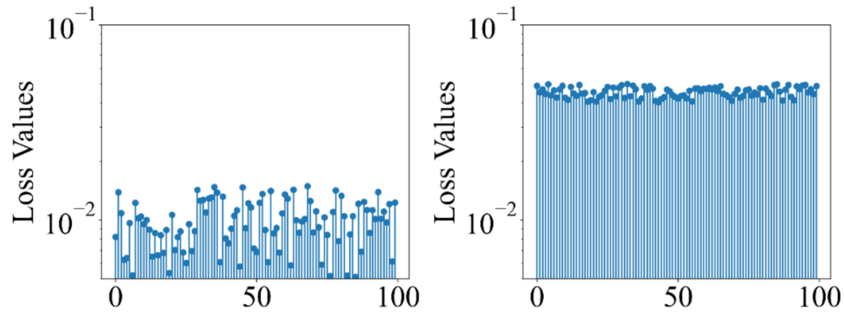


Figure 4: HIO to solve FFPR with vs without precise support. We plot the final least-squares losses over 100 random starts. $\boldsymbol{X}$ is the groundtruth in Fig. 1.

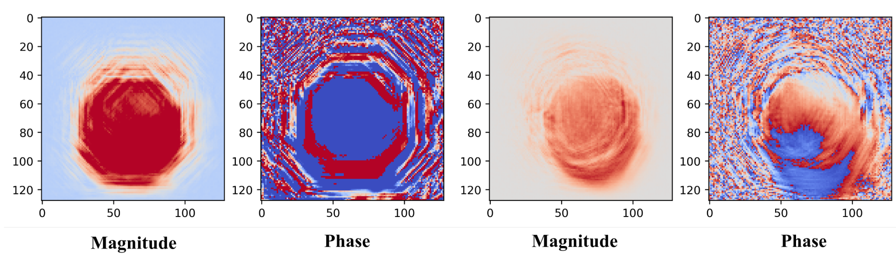


Figure 5: Two failure examples when solving FFPR using classical iterative methods without precise support specification and without shrinkwrap.

C Evaluation on 3D simulated Bragg CDI crystal data

We will not continue considering data-driven methods, due to their clear performance deficiency on 2D data and the considerable cost to obtain sufficiently representative training sets for 3D. We only compare HES, which is the gold-standard used in Bragg CDI practice, with our double-DIP method. Since both methods can work with single instances and need no training data, here we quickly compare their performance qualitatively on a single 3D simulated crystal instance (the simulation process is similar to the 2D case), as shown in Fig. 6. It is obvious that even with Shrinkwrap iteratively refining support, HES still struggles to get the support right. By contrast, our double-DIP method obtains sharp support recovery and good phase estimation.

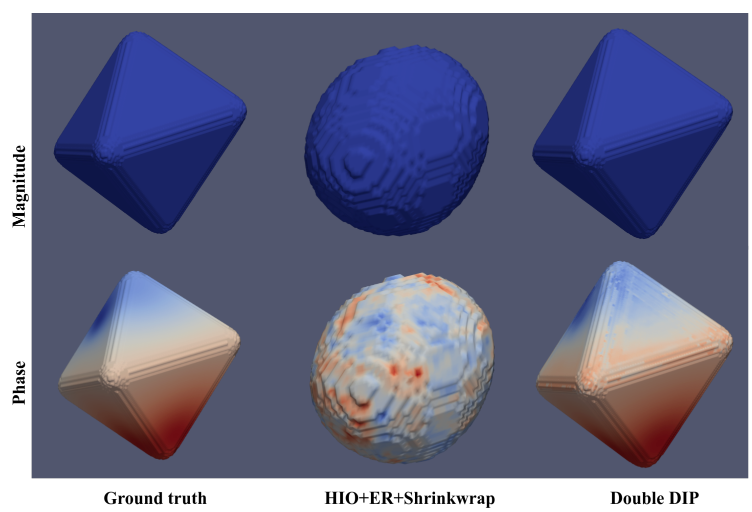


Figure 6: Visual comparison of reconstruction results by HIO+ER+Shrinkwrap and our method on a 3D simulated crystal instance

D Fresnel and Fraunhofer approximations to the diffraction formula

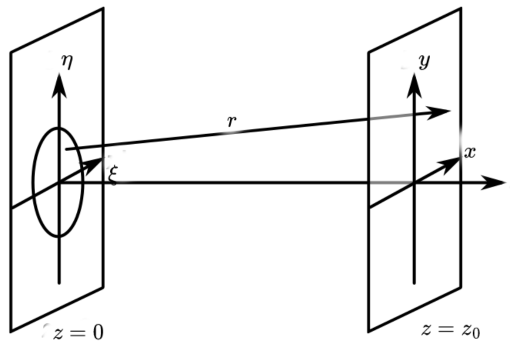


Figure 7: A schematic diffraction imaging system with parallel aperture and imaging planes (plot adapted from `https://commons.wikimedia.org/wiki/File:Diffraction_geometry.svg` under the Creative Commons Attribution-Share Alike 3.0 Unported license.)

In this section, we clarify the difference between the near-field and far-field models for phase retrieval, following [1, 55]. Consider the propagation of a monochromatic wave (with wavelength λ) from an aperture plane $z = 0$ to a parallel imaging plane $z = z_0$ with $z_0 \gg \lambda$; see Fig. 7. Let $U(x, y, z)$ denote the wave field. The celebrated Rayleigh-Sommerfeld diffraction formula dictates that

$$\boxed{U(x,y,z_0) = \frac{1}{2\pi} \iint U(\xi,\eta,0) \left(\frac{1}{r(\xi,\eta)} - ik \right) \frac{z_0}{r^2(\xi,\eta)} e^{ikr(\xi,\eta)} \; d\xi d\eta}, \tag{3}$$

where $r(\xi, \eta) \doteq \sqrt{z_0^2 + (x-\xi)^2 + (y-\eta)^2}$, $k \doteq \frac{2\pi}{\lambda}$ is the wavenumber, and the effective domain of the double integral is the aperture Ω. Moreover, write the domain of the image plane as $\Sigma \subset \mathbb{R}^2$. Eq. (3) has an equivalent form:

$$\boxed{U(x,y,z_0) = \iint \widehat{U}(f_X, f_Y, 0) e^{ikz_0\sqrt{1-(\lambda f_X)^2-(\lambda f_Y)^2}} e^{i2\pi(f_X x + f_Y y)} \; df_X df_Y}, \tag{4}$$

where $\widehat{U}(f_X, f_Y, 0) = \iint U(x, y, 0) e^{-i2\pi(f_X x + f_Y y)} \; df_X df_Y$ is the 2D Fourier transform of the planar field $U(x, y, 0)$. The equivalence is due to the convolution theorem: write $r^\circ(x, y) \doteq \sqrt{z_0^2 + x^2 + y^2}$ and $h(x, y, z_0) \doteq \frac{1}{2\pi}(\frac{1}{r^\circ(x,y)} - ik) \frac{z_0}{[r^\circ(x,y)]^2} e^{ikr^\circ(x,y)}$. Then Eq. (3) can be written as $U(x, y, z_0) = U(x, y, 0) * h(x, y, z_0)$. The equivalence is clear once we recognize that $\widehat{r^\circ}(f_X, f_Y, z_0) \doteq \iint h(x, y, z_0) e^{-i2\pi(f_X x + f_Y y)} \; df_X df_Y = e^{ikz_0\sqrt{1-(\lambda f_X)^2-(\lambda f_Y)^2}}$.

For Eq. (3), $\frac{1}{r(\xi,\eta)} - ik = \frac{1}{r(\xi,\eta)} - \frac{i2\pi}{\lambda} \approx \frac{2\pi}{i\lambda}$ as $r(\xi, \eta) \gg \lambda$ in practice. Hence we have the simplified form called the **Huygens-Fresnel principle**:

$$\boxed{U(x,y,z_0) = \frac{z_0}{i\lambda} \iint U(\xi,\eta,0) \frac{1}{r^2(\xi,\eta)} e^{ikr(\xi,\eta)} \; d\xi d\eta}. \tag{5}$$

We can derive two approximations to Eq. (5), i.e., the Fresnel (i.e., near-field) and Fraunhofer (i.e., far-field) approximations. Noting that $\sqrt{1+\varepsilon} = 1 + \varepsilon/2 - \varepsilon^2/8 + O(\varepsilon^3)$ for $|\varepsilon| \ll 1$, we can approximate

$$r(\xi,\eta) \doteq \sqrt{z_0^2 + (x-\xi)^2 + (y-\eta)^2} = z_0 \sqrt{1 + \left(\frac{x-\xi}{z_0}\right)^2 + \left(\frac{y-\eta}{z_0}\right)^2} \tag{6}$$

by its low-order Taylor expansions, provided that $(x-\xi)^2 + (y-\eta)^2 \ll z_0^2$ for all $(\xi,\eta) \in \Omega$. First, we have $r^2(\xi,\eta) \approx z_0^2$ using only the 0-th order expansion. For the exponential term, since k is normally large, we use the 1-st order expansion:

$$\exp\left(ikr(\xi,\eta)\right) \approx \exp\left[ik\left(z_0 + \frac{(x-\xi)^2}{2z_0} + \frac{(y-\eta)^2}{2z_0}\right)\right] = e^{ikz_0}\exp\left[\frac{ik}{2z_0}\left((x-\xi)^2+(y-\eta)^2\right)\right], \tag{7}$$

which is acceptable when

$$\frac{z_0 k}{8}\left(\left(\frac{x-\xi}{z_0}\right)^2 + \left(\frac{y-\eta}{z_0}\right)^2\right)^2 \ll 1 \Longleftrightarrow z_0^3 \gg \frac{k}{8}\left[(x-\xi)^2+(y-\eta)^2\right]^2 \quad \forall\,(\xi,\eta)\in\Omega, (x,y)\in\Sigma. \tag{8}$$

So we arrive at the famous Fresnel approximation

$$\boxed{\begin{aligned} U(x,y,z_0) &\approx \frac{e^{ikz_0}}{i\lambda z_0}\iint U(\xi,\eta,0)\exp\left[\frac{ik}{2z_0}\left((x-\xi)^2+(y-\eta)^2\right)\right]\,d\xi d\eta \\ &= \frac{e^{ikz_0}}{i\lambda z_0}e^{\frac{ik}{2z_0}\left(x^2+y^2\right)}\iint\left[U(\xi,\eta,0)e^{\frac{ik}{2z_0}\left(\xi^2+\eta^2\right)}\right]e^{-\frac{ik}{z_0}(x\xi+y\eta)}\,d\xi d\eta \\ &\qquad\qquad \textbf{Fresnel approximation—forward}\end{aligned}}\,.$$

If moreover $\frac{ik}{2z_0}\left(\xi^2+\eta^2\right) \ll 1 \Longleftrightarrow z_0 \gg \frac{k}{2}\left(\xi^2+\eta^2\right) \quad \forall\,(\xi,\eta)\in\Omega$, then $e^{\frac{ik}{2z_0}\left(\xi^2+\eta^2\right)} \approx 1$ and hence we obtain:

$$\boxed{\begin{aligned} U(x,y,z_0) \approx \frac{e^{ikz_0}}{i\lambda z_0}e^{\frac{ik}{2z_0}\left(x^2+y^2\right)}\iint U(\xi,\eta,0)e^{-\frac{ik}{z_0}(x\xi+y\eta)}\,d\xi d\eta \\ \textbf{Fraunhofer approximation—forward}\end{aligned}}\,.$$

By assuming $(\lambda f_X)^2 + (\lambda f_Y)^2 \ll 1$ and so $\sqrt{1-(\lambda f_X)^2-(\lambda f_Y)^2} \approx 1-(\lambda f_X)^2/2-(\lambda f_Y)^2/2$, we can approximate Eq. (4) as

$$\boxed{\begin{aligned} U(x,y,z_0) \approx e^{ikz_0}\iint \widehat{U}(f_X,f_Y,0)e^{-i\pi z_0\lambda(f_X^2+f_Y^2)}e^{i2\pi(f_X x+f_Y y)}\,df_X df_Y \\ \textbf{Fresnel approximation—backward}\end{aligned}}\,.$$

This is equivalent to the forward form of Fresnel approximation, due to

$$\mathcal{F}\left(\frac{1}{i\lambda z_0}e^{\frac{ik}{2z_0}\left(x^2+y^2\right)}\right) = e^{-i\pi\lambda z_0(f_X^2+f_Y^2)}, \tag{9}$$

and the convolution theorem again.

So when we measure the field intensity on the image plane,

$$\begin{aligned} |U(x,y,z_0)|^2 &\propto \left|\iint\left[U(\xi,\eta,0)e^{\frac{ik}{2z_0}\left(\xi^2+\eta^2\right)}\right]e^{-\frac{ik}{z_0}(x\xi+y\eta)}\,d\xi d\eta\right|^2 &(10)\\ &\propto \left|\iint \widehat{U}(f_X,f_Y,0)e^{-i\pi z_0\lambda(f_X^2+f_Y^2)}e^{i2\pi(f_X x+f_Y y)}\,df_X df_Y\right|^2 &(11)\end{aligned}$$

according to the Fresnel (near-field) approximation, and

$$|U(x,y,z_0)|^2 \propto \left|\iint U(\xi,\eta,0)e^{-\frac{ik}{z_0}(x\xi+y\eta)}\,d\xi d\eta\right|^2 \tag{12}$$

by the Fraunhofer (far-field) approximation. Detailed discussion of discretization and computation can be found in [1, Chapter 5].