

# Low-light phase retrieval with implicit generative priors

Raunak Manekar\*, Elisa Negrini\*, Minh Pham, Daniel Jacobs, Jaideep Srivastava, Stanley J. Osher, Jianwei Miao

**Abstract**—Phase retrieval (PR) is fundamentally important in scientific imaging and is crucial for nanoscale techniques like coherent diffractive imaging (CDI). Low radiation dose imaging is essential for applications involving radiation-sensitive samples. However, most PR methods struggle in low-dose scenarios due to high shot noise. Recent advancements in optical data acquisition setups, such as in-situ CDI, have shown promise for low-dose imaging, but they rely on a time series of measurements, making them unsuitable for single-image applications. Similarly, data-driven phase retrieval techniques are not easily adaptable to data-scarce situations. Zero-shot deep learning methods based on pre-trained and implicit generative priors have been effective in various imaging tasks but have shown limited success in PR. In this work, we propose low-dose deep image prior (LoDIP), which combines in-situ CDI with the power of implicit generative priors to address single-image low-dose phase retrieval. Quantitative evaluations demonstrate LoDIP’s superior performance in this task and its applicability to real experimental scenarios.

**Index Terms**—deep generative models, computational imaging, phase retrieval, inverse problems, low light imaging, low photon count, zero-shot learning

## I. INTRODUCTION

Coherent diffractive imaging (CDI) is a lensless imaging technique [1] used for high resolution imaging at nanoscale. CDI has found broad applications across different disciplines due to its remarkable ability to provide high-resolution structural information about a wide range of specimens from biological to nanoscale objects [2]. Unlike visible light, X-rays have high penetrating power and thus can be used to image thick, unfixed specimens. However, many samples of interest for CDI, such as biological materials, polymers or organic semiconductors, require minimal radiation exposure to prevent damage during data acquisition [3, 4]. Thus, there is considerable interest in techniques that minimize sample radiation exposure and enable X-ray imaging at extremely low photon counts. However, imaging at low photon counts is very challenging as it leads to very high shot noise in the acquired measurements (diffraction pattern) and makes the subsequent image reconstruction (i.e. phase retrieval) very hard. Under high shot noise, iterative phase retrieval algorithms, including

ER, HIO, become unstable, as they often get trapped in local minima, stagnate, and fail to converge [5, 6].

**Motivation.** One category of proposed solutions for the low-dose challenge involve modifications to the optical data acquisition setup of CDI [7, 8, 9]. A notable example is in-situ CDI, as introduced by Lo et al. [9] and further explored in [10]. A dose-tolerant static region is positioned next to the sample and is illuminated with a higher radiation dose. This ensures sufficient light reaches the detector while maintaining a low radiation dose on the sample. Multiple measurements are taken over time and the static region is leveraged as a robust time-overlap constraint which regularizes the phase retrieval optimization. However, in order to leverage the time-invariant static region as real-space constraint, this method requires multiple measurements (that is multiple diffraction patterns). Adapting this setup for single-image scenarios results in suboptimal reconstructions (see HIO-stat column in Fig. 4).

On the computational front, data-driven supervised learning approaches have shown potential for simpler versions of phase retrieval at low-photon counts [11, 12, 13]. However, the necessity for a substantial dataset is a challenge to meet for numerous scientific imaging applications. Furthermore, these methods are prone to generalization problems if the training data diverges from the target domain. Consequently, deep learning-based zero-shot methods such as deep image prior (DIP) [14, 15] have been employed. These do not require pre-training, rather the primary concept is to leverage the inductive bias induced by the neural network architecture as a prior for natural images. We refer to this class of methods as implicit generative priors to highlight the absence of training data. While successful in simpler phase retrieval scenarios [16, 17, 18], these methods often struggle to converge to satisfactory solutions in more challenging realistic scenarios, such as the single-shot far-field Phase Retrieval (FFPR) studied in this work. This observation is supported by our low-dose experiments (see Fig. 4) and corroborated by prior literature [19, 20]. The difficulty can be attributed to the well-known trivial ambiguities in the phase retrieval problem [5].

**Basic Idea.** LoDIP, a single-image method for low-dose phase retrieval, modifies both the experimental setup and computational algorithm of classical CDI. We modify the imaging setup of CDI to incorporate a static region illuminated with a high dose of radiation alongside the sample of interest which is illuminated by a low dose. Then we modify the DIP framework to exploit the additional physical constraints coming from this setup. The high-dose static region increases the amount of light incident on the detector, without increasing the radiation dose on the sample. This results in a reduction of

\*these authors contributed equally to this work. *corresponding author*: R. Manekar (manek009@umn.edu) R. Manekar and J. Srivastava are with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA. (email: manek009@umn.edu, srivasta@umn.edu). S.J.Osher and E. Negrini are with the Department of Mathematics, University of California, Los Angeles, Los Angeles, CA, USA (email: sjo@math.ucla.edu, enegrini@math.ucla.edu). J. Miao, M. Pham and D. Jacobs are with the Department of Physics and Astronomy, and California NanoSystems Institute, University of California, Los Angeles, CA, USA (email: j.miao@ucla.edu, minhrose@ucla.edu, danieljacobs@physics.ucla.edu)

noise in the captured diffraction pattern. Moreover, the static region aids in the convergence of implicit generative prior by alleviating the trivial ambiguities [21].

**Innovation.** LoDIP introduces innovation by concurrently modifying both experimental setup and computational algorithm. By integrating physics-based constraints from the modified experimental setup, LoDIP effectively resolves instability and artifacts typical of conventional DIP methods in single-shot far-field phase retrieval. This ensures stable optimization and precise reconstruction from low-light diffraction patterns, leveraging the benefits of deep image prior optimization while overcoming inherent challenges in single-shot far-field phase retrieval.

Unlike in-situ CDI setup, which requires multiple diffraction patterns, LoDIP operates in a single-image framework. Additionally, in contrast to in-situ CDI, where the static region is used as a time-invariant constraint present in multiple measurements, in LoDIP, the high-dose imaging of the static region serves two novel purposes: (a) being imaged at high-dose, it provides an accurate reconstruction of the static region, which acts as a strong constraint crucial for achieving a high-quality reconstruction of the sample; and (b) enhancing the convergence of phase retrieval optimization by mitigating ambiguities arising from symmetries in the forward process, a significant challenge in phase retrieval [5, 21].

**Contributions.** Experiments demonstrate that LoDIP is effective for both simulated and experimental data. The LoDIP framework is very flexible, compatible with any implicit generative prior method, and can be tailored to diverse experimental setups. Unlike iterative phase retrieval algorithms, LoDIP does not necessitate specialized tuning to optimize the algorithmic parameters for satisfactory results, making LoDIP a user-friendly tool for a wide user community. The contributions of this work can be summarized as follows:

- A zero-shot method for low-dose phase retrieval. LoDIP uses the inductive bias of an implicit generative prior and does not require large dataset for training.
- Extensive experiments on simulated and experimental datasets demonstrate that LoDIP outperforms state-of-the-art methods across a range of image quality metrics.
- On the experimental side, LoDIP being designed for single-image phase retrieval, can produce reconstructions comparable to the original in situ CDI method without requiring a time series of measurements.

## II. RELATED WORK

### A. Coherent Diffractive Imaging

In Coherent Diffractive Imaging (CDI), an object is illuminated by a highly coherent light source. The interaction between the object and the incident wave results in the generation of a diffraction pattern, which is subsequently detected. However, while detectors capture the magnitude of the diffracted wave, the phase information is lost. As a consequence, the process of reconstructing the image of the object of interest necessitates the development of a computational algorithm designed to recover the lost phase from the acquired diffraction pattern. This is commonly referred to as the “phase

retrieval” problem. As shown in [22], if the diffraction pattern is sufficiently oversampled, the phase can be retrieved from the diffraction pattern via iterative algorithms (see for instance [5]).

Representing the object of interest by the complex-valued matrix  $\mathbf{X} \in \mathbb{C}^{n \times n}$  and the captured diffraction pattern by  $\mathbf{Y} \in \mathbb{R}^{m \times m}$ , in the conventional phase retrieval configuration the diffraction pattern  $\mathbf{Y}$  is obtained through the application of the following forward process:

$$\mathbf{Y} = |\mathcal{F}(\mathbf{X})|^2 \quad (\text{II.1})$$

Here,  $\mathcal{F}$  represents the Fourier transform, converting the spatial information contained in the object matrix  $\mathbf{X}$  into the frequency domain.  $|\cdot|$  is applied element-wise. the original object  $\mathbf{X}$  is placed in an empty background such that  $m = 2 \times n$ . In simulation, this oversampling is achieved by zero-padding  $\mathbf{X}$ . The goal of the phase retrieval is to recover the image  $\mathbf{X}$  from  $\mathbf{Y}$ .

Generally, some information about the support of the object (i.e. the location of the object) within the empty background is known. Let  $\mathbf{S}_0 \in \mathbb{C}^{m \times m}$  be the known support information.  $\mathbf{S}_0$  is a matrix containing ones where the object of interest is estimated to be, zeros otherwise. In general one may not know the exact support of the object of interest, but only an estimate.

With these assumptions, the phase retrieval problem can be formulated as an optimization problem:

$$\begin{aligned} \min_{\hat{\mathbf{X}} \in \mathbb{C}^{n \times n}} \ell(\mathbf{Y}, |\mathcal{F}(\hat{\mathbf{X}})|^2), \\ \text{s.t. } (1 - \mathbf{S}_0) \odot \hat{\mathbf{X}} = [\mathbf{0}]_{m \times m} \end{aligned} \quad (\text{II.2})$$

where the first term imposes that II.1 is satisfied (magnitude constraint), while the second term is the support constraint.

### B. Phase retrieval for low-light imaging

**Challenge** The major challenge of imaging at low-photon counts is the strong presence of noise in the captured signal. In practice, most of the existing PR methods [6, 23, 24, 25, 26, 27] struggle to work in the low dose (high noise) setting. This comes in addition to the existing challenge of working without accurate support information which itself is sufficient for many of the existing algorithms to fail.

**Existing solutions** Attempts have been made at low-photon phase retrieval that modify the imaging setup of the problem [7, 8, 9]. While [7] relies on a phase diverse approach, [8, 9] use in the imaging setup a static region made of heavy metals (usually gold) which can withstand high energy light. Among these methods, in-situ CDI [9] achieves a substantial order of magnitude dose reduction over other methods. However, in-situ CDI uses a time-series of diffraction patterns which should contain a time-invariant “static” region. The static region’s presence in each time step’s measurement offers a strong overlap-in-time constraint for facilitating the convergence of phase retrieval optimization. Fourier holography [28, 29] also

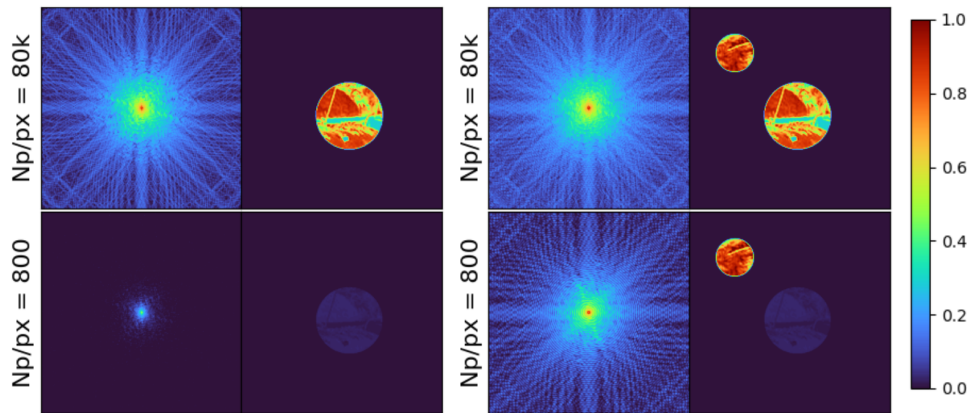


Fig. 1. Examples from the simulated data at different levels of illumination of the sample region as measured in photons per pixel (Np/px). **[Left]** The first two columns are the diffraction pattern and the image sample without a static structure. **[Right]** And the last two columns are the same sample with a static structure. **[Top]** The first row has a high illumination of 80k Np/px on both the sample and the static region. **[Bottom]** Whereas the second row has a high illumination of 80k Np/px on the static region and a low illumination of 800 Np/px on the sample. At lower illumination levels (bottom row), the sample exhibits significantly reduced pixel values. Lesser illumination results in higher presence of Poisson noise. Thus different illumination levels correspond to different levels of noise.

utilizes two regions of interest. However, unlike Fourier holography, LoDIP can handle varying illuminations and doesn't impose geometric constraints between the regions.

### C. Data-driven phase retrieval

Data-driven methods have shown promise for phase retrieval [12, 30, 31] including at low-dose [32]. Subsequently, the end-to-end supervised learning paradigm has been extended by incorporating known physical models in the neural-network training. This includes utilizing an unrolled iterative algorithm [33, 34] or by devising a self-supervised loss function [35, 36] capable of learning from measurements alone or using initial phase estimates from a physical model as inputs to the network [37, 38]. For a comprehensive review of data-driven phase retrieval, we refer the reader to [39].

However, these methods are not yet widely adopted by microscopy practitioners due to several challenges: they often require large datasets for training, extensive computational resources, and their performance is heavily reliant on the quality of the training data. Additionally, retraining is necessary to adapt the model to different experimental settings. In this study, we propose a single-image deep learning method based on deep image prior that overcomes these limitations, requiring no large dataset and being adaptable across various experimental setups.

### D. Zero-shot deep learning for phase retrieval

**Pre-trained generative priors.** Following the breakthroughs achieved by deep generative models in approximating image distributions [40, 41, 42, 43], such models have been leveraged as priors to regularize visual inverse problems [43, 44], including phase retrieval [13]. Methods based on pre-trained generative priors are commonly referred to as “zero-shot” as no additional task-specific training is performed in these methods. Despite being unsupervised and not reliant on labeled datasets, these approaches still necessitate pretraining

on extensive datasets. Additionally, akin to supervised learning methods, these pre-trained generative priors face challenges in generalization when the training data deviates from the target domain.

**Implicit generative priors.** A parallel line of work discovered that even a convolutional neural network (CNN) without training can serve as a robust prior for natural images, and can be effectively employed to regularize inverse problems on such images. Beginning with the so-called deep image prior (DIP) [14]. This has paved the way for a suite of methods including implicit neural representations (INR) [15, 45], deep decoder [17], double-DIP [46], random projection based methods [47] and enhanced versions of DIP [48]. For an exhaustive review, readers are directed to [49]. In the absence of an overarching terminology for these methods, we label them as ‘implicit generative priors’. This term highlights the inductive bias derived from the network architecture rather than the training data. It also differentiates these methods from the use of pre-trained generative priors. Both types are frequently referred to as ‘zero-shot’ deep learning methods. Finally, it highlights the generative nature of our algorithm which takes noise as input and generates a natural image.

**For Phase Retrieval.** Several of these have also been applied to computational imaging inverse problems [50, 51, 52] including simpler versions of phase retrieval [16, 17, 18, 53].<sup>1</sup> However, the settings addressed in these works are considerably simpler compared to the single-shot far-field (Fraunhofer) diffraction with high noise scenario considered in this paper. A naive application of DIP and related methods is known to struggle in such realistic phase retrieval scenarios [20, 53] .

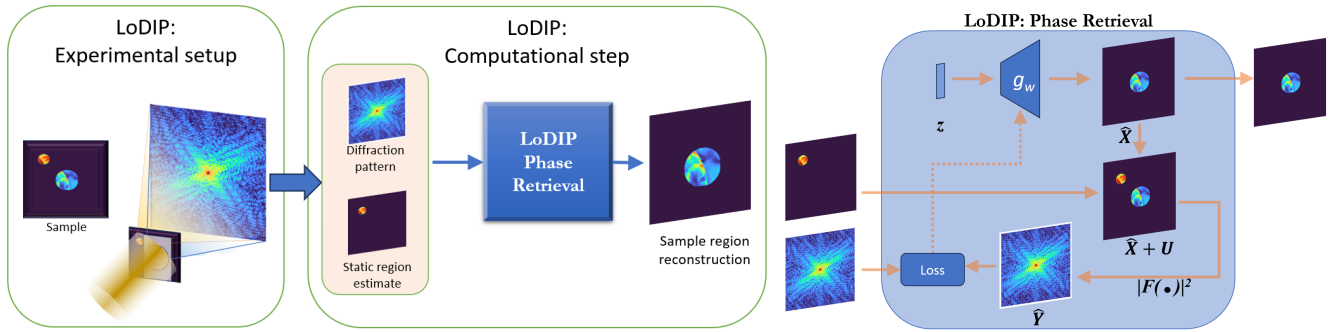


Fig. 2. Coherent Diffraction Imaging (CDI) employs a coherent X-ray beam directed at a sample, capturing the resulting diffraction pattern on a 2D detector. A computational algorithm is then applied to reconstruct the desired sample image. **(Left)** Inspired by in-situ CDI [9], LoDIP introduces two modifications to the CDI setup. First, it involves imaging the sample alongside a static region. Secondly, the static region is exposed to a high radiation dose, while the sample’s exposure is significantly reduced. A customized sample grid and holder is used with an attenuator placed on top of the sample to reduce the incident radiation dose on the sample. **(Center)** In the computational step, LoDIP takes both the diffraction pattern and an estimated reconstruction of the static region as inputs, generating a sample reconstruction as its output. **(Right)** LoDIP uses the output of an implicit generative model  $g_w$  as an estimate of the sample  $\hat{X}$  and iteratively updates the generator parameters  $w$  to refine the estimate.

### III. PROPOSED METHOD

The proposed method, which we refer to as LoDIP, modifies the data acquisition setup of conventional CDI to incorporate a high-dose static region inspired by the in-situ CDI setup. While in-situ CDI leverages the static region as a time-invariant constraint to reconstruct a sequence of measurements capturing a dynamic process, the high-dose static region serves a different purpose in LoDIP where we are interested in the single-image setting.

Firstly, it increases the available light for image formation on the detector, effectively reducing the impact of shot noise. Additionally, given the high-dose illumination, it is easy to produce an accurate reconstruction of the static region. The known static region is then used as a strong constraint to improve convergence of the phase retrieval optimization. Secondly, it mitigates the ambiguities arising from symmetries in the forward process which is a fundamental difficulty of phase retrieval, (see for instance [5, 21, 55]).

While these adaptations significantly enhance the performance of established phase retrieval methods (such as HIO) at low radiation doses, our experiments (see Fig. 3), show that LoDIP can further improve the final reconstruction quality by incorporating an implicit generative prior for phase retrieval.

#### A. LoDIP: experimental setup

The sample of interest is placed within a finite support next to a static region of heavily scattering, dose-tolerant object such as a gold (Au) pattern on an optical stage. The X-ray illumination on the dose-sensitive sample is reduced to a tolerable limit by the presence of an attenuator mounted to the sample support (see Figure 2), while the static region is exposed to the full dose of the incident illumination. The design of the customized sample support would need to include a SiN attenuator over the dose-sensitive region with a thickness tailored to the desired dose reduction factor and the

<sup>1</sup>Certain self-supervised data-driven methods have demonstrated efficacy even with a single sample [35, 54], resembling a zero-shot approach akin to an implicit generative prior. Instead of a random input seed, these methods rely on a sample-dependent input [39].

energy of the incident x-ray beam. For example, to achieve a dose reduction factor of three orders of magnitude at a soft x-ray beamline operating within the water window as considered in our simulations ( $\sim 525eV$ ), a  $Si_3N_4$  thickness of roughly  $1.5 \mu m$  is required <sup>2</sup>.

Far-field diffraction patterns recorded from this setup are formed by the interference in Fourier space between the high-dose static region and the sample. These contain Poisson noise relative to the total illumination on the detector. Given the high-dose illumination on the static region and the known support, it is possible to obtain a high-quality static region reconstruction. In this work we use an established iterative method, Generalized Proximal Smoothing (GPS) initialized with 1000 iterations of HIO.

The placement of the object and the static structure do not influence the reconstruction given by LoDIP, as long as their supports do not overlap, i.e.  $\mathcal{S}(\mathbf{X}) \odot \mathcal{S}(\mathbf{U}) = [\mathbf{0}]_{m \times m}$ , where  $\mathcal{S}(\cdot)$  extracts the support the image.

#### B. LoDIP: Phase Retrieval

For LoDIP, the known static region in the above data acquisition step provides us with a useful constraint to regularize and improve the convergence of the optimization problem. By incorporating the static structure  $\mathbf{U} \in \mathbb{C}^{k \times k}$  in the original phase retrieval formulation, the forward process becomes:

$$\mathbf{Y} = |\mathcal{F}(\mathbf{X} + \mathbf{U})|^2 \quad (\text{III.1})$$

Here  $\mathbf{X}$  and  $\mathbf{U}$  are appropriately zero-padded to have the same dimension. In a similar way the optimization problem becomes:

$$\begin{aligned} \min_{\mathbf{X} \in \mathbb{C}^{n \times n}} \ell(\mathbf{Y}, |\mathcal{F}(\mathbf{X} + \mathbf{U})|^2), \\ \text{s.t. } (\mathbf{1} - \mathbf{S}_0) \odot \mathbf{X} = [\mathbf{0}]_{m \times m} \end{aligned} \quad (\text{III.2})$$

In LoDIP this setup is further generalized by incorporating an implicit generative prior. For a given inverse problem aimed at recovering an image  $\mathbf{X} \in \mathbb{C}^{n \times n}$  from its measurements

<sup>2</sup>This estimated value was obtained using the tabulated atomic scattering factors for solid  $Si_3N_4$  [56, 57]

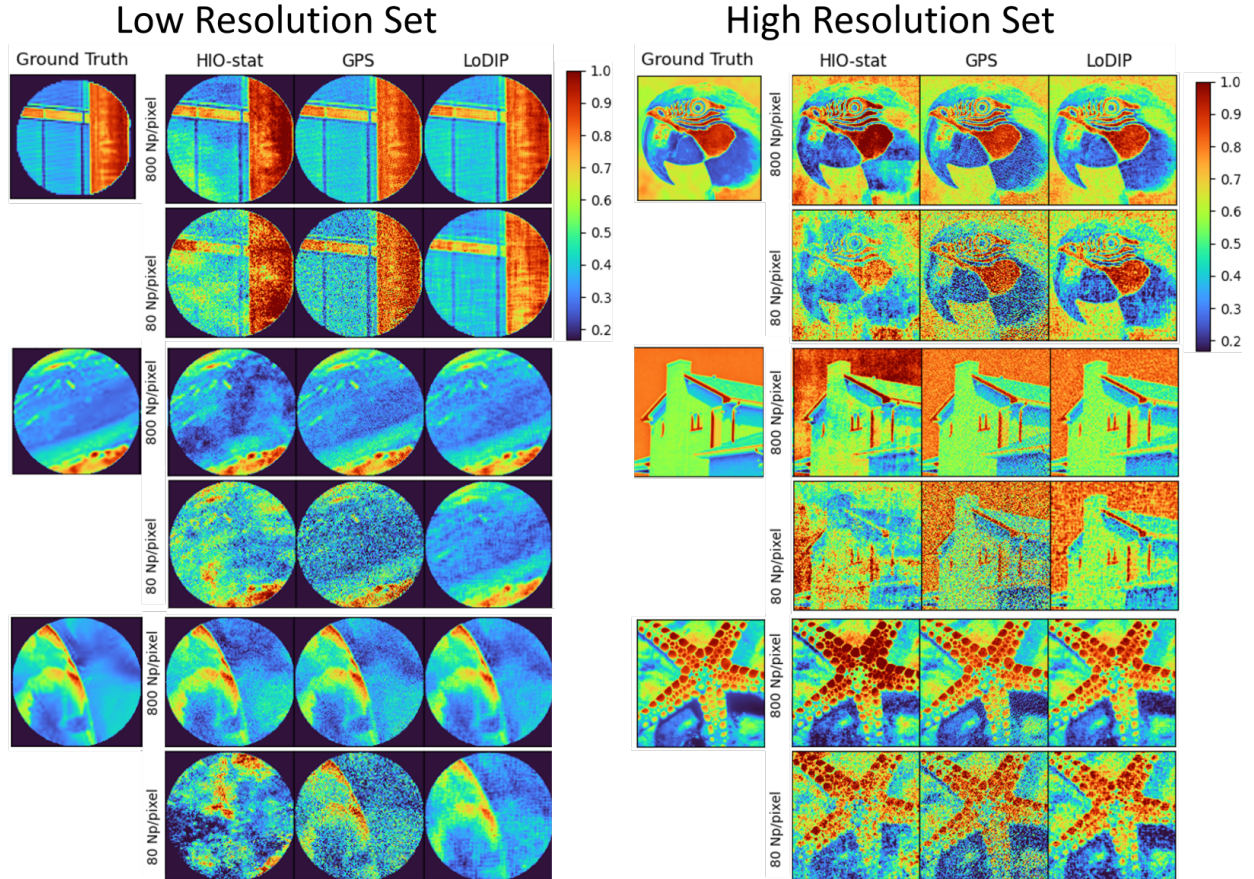


Fig. 3. Experimental Results on simulated data. **Left:** Reconstruction of Low Resolution Set images for 800 Np/pixel (Top rows) and 80 Np/pixel (Bottom rows). **Right:** Reconstruction of High Resolution Set images for 800 Np/pixel (Top rows) and 80 Np/pixel (Bottom rows). Each image shows a zoomed-in view of only the sample region.

$\mathbf{Y} \in \mathbb{C}^{m \times m}$  the implicit generative prior optimization problem is formulated as:

$$\min_{\mathbf{W}} \ell(\mathbf{Y}, \mathcal{A}(g_{\mathbf{W}}(\mathbf{z}))), \quad (\text{III.3})$$

where  $\mathbf{Y} = \mathcal{A}(\mathbf{X}) + \eta$ ,  $\mathcal{A}$  is the known forward operator and  $\eta$  is the noise. The optimization variable  $\mathbf{X}$  is re-parameterized with a new function  $g_{\mathbf{W}}(\mathbf{z})$ , generally a CNN. Here  $\mathbf{W}$  are the learnable parameters of the network and  $\mathbf{z}$  is a random input seed which is fixed throughout the optimization process.  $g_{\mathbf{W}}$  essentially works as a generative network which is optimized to output a single image.

Further, we extend this formulation to incorporate the constraints coming from the above setup: we include the known static region, the known sample support as well as the relative illumination doses on sample and static regions. Finally the optimization problem solved by LoDIP is:

$$\begin{aligned} \min_{\mathbf{W}} \ell(\mathbf{Y}, |\mathcal{F}(g_{\mathbf{W}}(\mathbf{z}) + \mathbf{U})|^2), \\ \text{s.t. } (1 - \mathbf{S}_0) \odot \mathbf{X} = [\mathbf{0}]_{m \times m} \end{aligned} \quad (\text{III.4})$$

**Extensions to LoDIP framework** The LoDIP framework is highly versatile and can accommodate a variety of modifications. Firstly, it can incorporate any suitable implicit generative prior based architecture, including variants of DIP

[58, 59]. Secondly, LoDIP can be easily tailored to different experimental conditions by integrating them into the forward operator. This adaptability is demonstrated in Section IV-B, where experimental data captured using a probe was used. In this instance, LoDIP was adjusted to include the known probe function, in contrast to existing state-of-the-art methods like GPS [6], which would necessitate significant alterations for different experimental setups.

Moreover, LoDIP can smoothly integrate a rough reconstruction of the sample as initialization, enabling it to be effectively combined with other techniques aimed at generating high-quality initializations for iterative phase retrieval methods [55]. Finally, LoDIP can be conveniently modified to function in the multiple measurements setting of the original in-situ CDI paper.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

### A. Data

We perform experiments on three kinds of data. First, we create a simulated data<sup>3</sup> using a procedure similar to [60]. Natural images are used to create the sample region and the static region for each sample in the dataset. Examples of the

<sup>3</sup>Data is available at <https://github.com/raunakmanekar/LoDIP-PR>.

TABLE I

QUANTITATIVE COMPARISON OF LoDIP AND HIO-STAT. FOR EACH METRIC, THE REPORTED VALUES REPRESENT THE MEAN OVER ALL THE SAMPLES IN THE DATASET. THE STANDARD DEVIATION REPORTED FOR PSNR SHOWS THE ROBUSTNESS OF THE METHOD TO DIFFERENT RANDOM INPUTS.

	Np/px=800			Np/px=80		
	Low resolution Set					
	PSNR $\uparrow$	SSIM $\uparrow$	$R_{real}$ $\downarrow$	PSNR $\uparrow$	SSIM $\uparrow$	$R_{real}$ $\downarrow$
LoDIP	<b>28.52</b> ( $\pm 2.28$ )	<b>0.76</b>	<b>0.07</b>	<b>25.09</b> ( $\pm 1.18$ )	<b>0.63</b>	<b>0.10</b>
HIO-stat	21.55 ( $\pm 0.77$ )	0.47	0.18	13.11 ( $\pm 1.48$ )	0.17	0.44
GPS	22.96 ( $\pm 0.20$ )	0.53	0.13	13.91 ( $\pm 0.03$ )	0.29	0.39
	High Resolution Set					
LoDIP	<b>23.56</b> ( $\pm 0.55$ )	<b>0.59</b>	<b>0.09</b>	<b>20.01</b> ( $\pm 0.65$ )	<b>0.42</b>	<b>0.14</b>
HIO-stat	19.90 ( $\pm 1.15$ )	0.57	0.13	13.99 ( $\pm 1.10$ )	0.24	0.29
GPS	19.16 ( $\pm 0.30$ )	0.45	0.15	10.83 ( $\pm 0.04$ )	0.17	0.40
	Biological cell sample					
LoDIP	<b>26.97</b> ( $\pm 0.40$ )	<b>0.68</b>	<b>0.11</b>	<b>19.08</b> ( $\pm 0.99$ )	<b>0.32</b>	<b>0.26</b>
HIO-stat	21.27 ( $\pm 1.44$ )	0.50	0.18	12.21 ( $\pm 1.62$ )	0.14	0.39
GPS	23.68 ( $\pm 0.31$ )	0.53	0.15	15.14 ( $\pm 0.04$ )	0.19	0.42

generated images can be seen in Fig. 1. Next, using the data provided in [9] we create a simulated image of biological cell using physically accurate simulations of a gold lacey for the static structure and a biological cell for the sample region, these can be found in Fig. 4, 6.

Finally, we demonstrate the applicability of LoDIP on experimental diffraction patterns from live glioblastoma cells (see Fig. 7). In this setup, the collimated beam with a wavelength of 534 nm incident onto a pair of 100  $\mu\text{m}$  pinholes with an edge-to-edge separation distance of 100  $\mu\text{m}$ . This produced a pair of beams incident onto the sample 400  $\mu\text{m}$  downstream of the dual pinhole aperture with the same intensity, as no dose reduction was performed in this experiment. The diffraction intensity from the sample exit wave was recorded using a CCD detector placed at the focal point of a lens adjacent to the sample. The static structure is a 100  $\mu\text{m}$  pinhole exposed to the same incident illumination as the sample. Unlike the first two datasets above, this data has been collected using a probe. This dataset is sourced from Lo et al. [9]. Further information about the optical laser data collection can be found in the original paper.

*a) Simulated data generation:* The experimental setup is similar to the one in [9]. The size of the entire image (here, 512x512) corresponds to the size of the detector in the real experimental setup. The support of the sample region is 170x170 pixels. This gives an oversampling ratio of approximately 3.0 which is higher than the theoretical requirement of 2. The sample and the static structure have non-overlapping support and the static structure is set to have half the radius of the sample region (however, the method is independent of the static region size).

*b) Simulating low-light conditions:* Based on previous studies on low-dose phase retrieval [9] and [10], the illumination on the sample has been varied from  $2.5 \times 10^5$  to  $2.5 \times 10^9$  photons per  $\mu\text{m}^{-2}$ . The illumination on the static structure has been fixed at  $2.5 \times 10^9$  photons per  $\mu\text{m}^{-2}$ . We show results on the low doses of  $2.5 \times 10^6$  and  $2.5 \times 10^7$  photons per  $\mu\text{m}^{-2}$  which corresponds to 80 and 800 photons per pixel (Np/px) based on the simulation geometry.

## B. Results

**Methods** For all experiments we compare the relative performance of the proposed method (LoDIP) with other popular methods which can be used in this setup. Specifically, we compare with Hybrid input-output (HIO) [61]; Deep Image Prior [14] (DIP) with no static region information; HIO-stat which is a modification of HIO that uses a static region; and Generalized proximal smoothing (GPS) [6] which is the state of the art method for phase retrieval. HIO and DIP work with diffraction patterns collected without a static region (see column 1 in Fig. 1. Rest of the methods, namely HIO-stat, GPS and LoDIP use the static region, however LoDIP also exploits neural network priors.

**Metrics.** Following previous works on reconstruction with noisy measurements [62], we measure the reconstruction accuracy for all experiments in this paper using using Peak Signal-to-Noise Ratio (PSNR). Higher values of PSNR indicate better fidelity; Structural Similarity Index (SSIM) [63] which measures the similarity between two images, evaluating luminance, contrast, and structure for enhanced accuracy in assessing perceptual image quality (higher the better); R-factor ( $R_{real}$ ) [6] which measures the degree of agreement between observed and predicted data, commonly used in the assessment of image reconstruction quality (smaller the better).

*a) Reconstruction of natural images:* Two commonly used natural image datasets for testing phase retrieval algorithms were selected: the Set12 dataset [62, 63, 64] and internet-sourced stock images used in [60]. For each image, diffraction patterns were first generated using the forward model outlined in Section III and Section IV-A. Results were presented for two low-dose scenarios: 800 and 80 photons per pixel. Different doses correspond in practice to different amounts of Poisson noise. Good reconstruction across dose amounts shows the robustness of the method with respect to different noise levels.

We observed that the Set12 images showed higher resolution and finer details compared to the stock images, hence we distinguish them as the high and low resolution sets, respectively. The complementary characteristics of these datasets allow for a thorough evaluation of each method's denoising capability and achievable resolution. Across each dataset, we compared reconstructions from all methods using 12 samples,

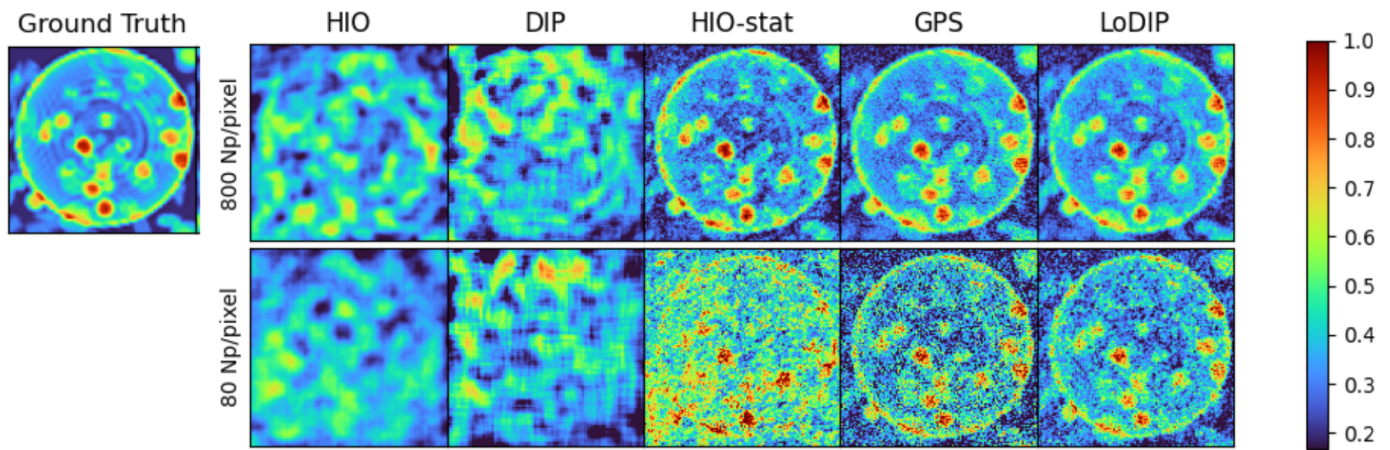


Fig. 4. Experimental Results on biological cell sample. (**Top row**) Reconstruction at 800 Np/pixel. (**Bottom row**) Reconstruction at 80 Np/pixel).

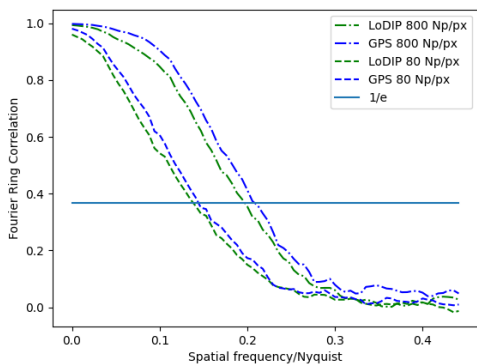


Fig. 5. Comparison of FRC values for GPS and LoDIP reconstructions on biological cell at 800 Np/px and 80 Np/px.

each utilizing a different image for its static structure and sample region.

Table I shows results averaged over the entire test set of 12 images for the low and high resolutions sets. Since DIP and HIO without reference regions did not provide good results in this case we omit those methods here. LoDIP attains better results across all metrics for both data sets and photon counts. The better performance of LoDIP is especially clear in the low photon counts case where LoDIP accuracy is approximately double the accuracy of the other methods. Finally, as expected, the error is higher for the high resolution set as more fine detail has to be reconstructed.

**Discussion** Fig. 3 displays a comparison of the results of LoDIP, HIO-stat and GPS for images in the low resolution set (left) and high resolution set (right). To study the robustness of our method with respect to the experimental setup we used a circular mask for low resolution set and a square one for the high resolution set. GPS and LoDIP yield the best results, however, GPS generates noisier images compared to LoDIP, highlighting LoDIP’s strong denoising ability. The figure shows also the choice of the mask for the sample region does not influence the reconstruction quality for LoDIP (and for the other methods).

We also note that iterative methods such as HIO-stat and GPS require parameter tuning by experts in the field and often multiple independent runs of the algorithm are necessary to obtain good reconstruction. On the other hand, LoDIP hyper-parameter tuning is minimal and convergence is usually obtained in one run.

*b) Reconstruction of biological cell sample:* Computational microscopy represents a major potential application of the proposed method. Following previous works on PR for CDI [9, 10], we evaluated the performance of LoDIP on realistic and physically accurate simulated data for a prototypical live cell. The static structure and biological cell were simulated as done in [9]. The results in this section use a simulated 20-nm thick gold lacey pattern as the static structure and a simulated cell consisting of a vesicle containing water and protein aggregates. The static region is unknown and estimated using the procedure described in Section III-A.

**Discussion** Similar to the case of natural images, Fig. 4 shows that GPS and LoDIP produce the best reconstructions in both cases with LoDIP providing a less noisy reconstructed image than GPS especially at 80 Np/px. HIO and DIP without reference regions fail in the reconstruction. The last section of Table I shows that LoDIP outperforms the other methods across all metrics.

**Comparing Resolution.** Fig. 5 displays the Fourier Ring Correlation (FRC) for both photon counts cases for LoDIP and GPS. Higher FRC values correspond to better resolution. The resolution of the reconstruction obtained by LoDIP is comparable to the one obtained by GPS. This indicates that the denoising capability of LoDIP results in only a minimal loss of fine-scale details.

**Reconstruction without precise support** PR without precise support information presents a common and challenging problem. From Fig. 6, we can see that LoDIP reconstructions are robust to inaccurate support specification. Widely used phase retrieval algorithms struggle when accurate support information is not available [20, 53] and require additional measures for support estimation[65].

*c) Reconstruction from experimental data:* Finally, we demonstrate the proposed method on experimentally cap-

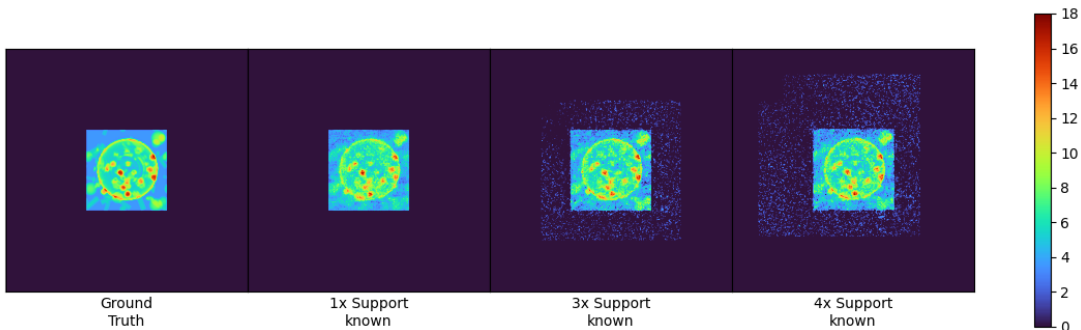


Fig. 6. Reconstruction using LoDIP at 80 Np/px without knowledge of accurate support. The columns from left to right are (1) The ground truth, (2) reconstruction with knowledge of the exact support (PSNR=20.08), (3) reconstruction with an approximate support of 3x the actual size (PSNR=20.06), (4) 4x the actual size (PSNR=19.85). Both visually and quantitatively, the LoDIP reconstruction is robust to inaccurate specification of the support. PSNR is calculated only over the exact sample region.

tured diffraction patterns (one of them is shown in Fig. 7 left column). Unlike the previous experiments, the object is complex-valued and the optical setup includes a probe. We note explicitly that in this dataset the static and sample images have the same high dose, so this is not exactly the setup for which our method is designed. In fact, our method is especially advantageous in the low dose setting, as the previous experiments show. Nonetheless, this setup is very close to the one of interest so we included it in this work.

**Comparison with in-situ CDI.** The original in-situ CDI method [9] uses 50 diffraction patterns with a fixed time-invariant static structure required in all the images. HIO-stat and LoDIP use only a single diffraction pattern and perform single-image phase retrieval. This is an advantage over in-situ CDI since as in many experimental setups a truly static structure is difficult to ensure because of the data collection procedure (e.g. in collecting tomographic data, which requires shifting or tilting both the sample and the static structure). In this section, the reconstruction given by in-situ CDI is used as a proxy for the ground truth. In fact, a comparison between in-situ CDI and LoDIP would be unfair since LoDIP uses 50 times less data (only one diffraction pattern instead of 50).

**Comparison with GPS.** A comparison with GPS is not possible for this data since the current version of GPS does not support Fresnel propagation and adding this would require major modifications which are outside of the scope of this work. In these experiments, LoDIP and HIO-stat have been modified to incorporate the known probe function.

**Metrics.** Since there is no ground truth available, to evaluate our results we use R-factor in the Fourier domain. This measures the disagreement between the captured diffraction pattern  $\mathbf{Y}$  to the Fourier magnitudes  $|\mathcal{F}(\hat{\mathbf{X}})|$  of the reconstruction  $\hat{\mathbf{X}}$  and is defined as:

$$R_F(\hat{\mathbf{X}}) = \frac{\sum_{i,j} \left| |\mathcal{F}(\hat{\mathbf{X}})|_{i,j} - Y_{i,j} \right|}{\sum_{i,j} Y_{i,j}} \quad (\text{IV.1})$$

The reconstruction obtained through in-situ CDI yields an R-factor of 30.22%(±0.97%). As explained before in-situ CDI utilizes 50 diffraction patterns for its reconstruction.

Comparing this R-factor directly with LoDIP, which only uses a single diffraction pattern, would be unfair due to the substantial difference in data quantity needed to obtain the reconstruction. However, we can use in-situ CDI R-factor as a benchmark to assess the performance of our method.

**Discussion.** The average  $R_F$  from 20 independent reconstructions based on a single diffraction pattern using both LoDIP and HIO-stat is presented in Table II. Remarkably, LoDIP achieves a comparable R-factor to in-situ CDI, highlighting its efficacy even without the need for multiple diffraction patterns, making it less data-intensive. Interestingly, HIO-stat also demonstrates a comparable R-factor to LoDIP in this scenario. It's important to note, however, that this success is primarily attributed to the high photon count setting of this experiment, where many methods perform well. It is crucial also to note that, as demonstrated in previous experiments, HIO-stat's performance diminishes in low photon count settings, providing less accurate results than LoDIP in this more realistic case. Finally, Fig. 7 shows the averaged top 5 reconstructions out of 20 independent runs for all methods. We see that HIO-stat introduces visible artifacts in its reconstruction while LoDIP does not (again the in-situ CDI reconstruction is treated as a proxy for the ground truth in this case). Therefore, these results highlight the robustness and adaptability of LoDIP, showcasing its ability to easily accommodate changes in the data acquisition setup and maintain reliable reconstructions across varying experimental conditions.

TABLE II  
QUANTITATIVE COMPARISON ON EXPERIMENTAL DATA. THE MEAN AND STANDARD DEVIATION OF R-FACTOR( $R_F$ ) VALUES ARE CALCULATED OVER 20 INDEPENDENT RECONSTRUCTIONS EACH OF 5 DIFFERENT SAMPLES.

	$R_F$
HIO-stat	34.70%(±0.03%)
LoDIP	33.20%(±0.03%)

## V. CONCLUSION AND FUTURE WORK

Emerging deep learning methods present several opportunities to improve upon and expand the scope of computa-



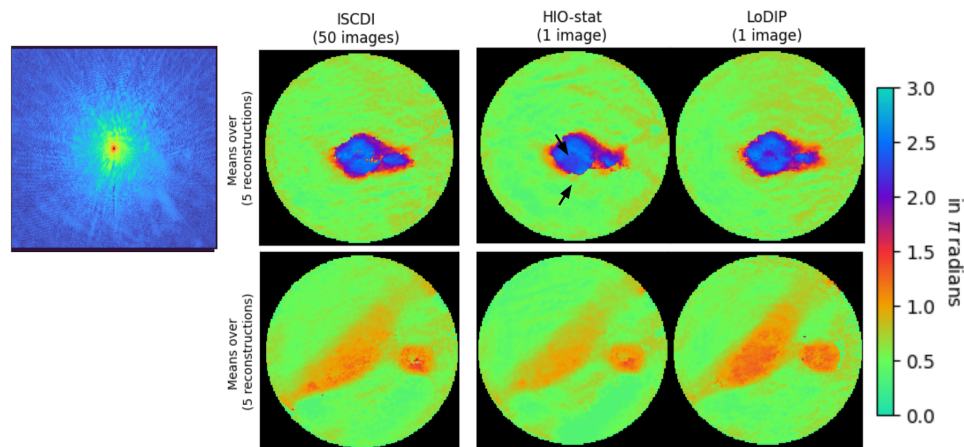


Fig. 7. Reconstruction of experimental data (**Left**) An experimentally captured diffraction pattern. (**Right**) Reconstruction of two samples from different methods. The mean of the top five reconstructions from 20 independent runs is shown in the figure. The static and sample images in this example have the same high dose. LoDIP is especially advantageous in the low dose setting, as the previous experiments show, so as expected, in this high dose setting LoDIP obtains comparable results as in-situ CDI and HIO in terms of  $R_F$ . However, LoDIP uses a single diffraction pattern to produce a reconstruction of comparable quality as in-situ CDI which uses 50 samples. Moreover, while HIO-stat reconstruction has similar  $R_F$  as LoDIP, it contains visible artefacts (indicated by the arrows).

tional imaging techniques framed as mathematical optimization problems. The LoDIP approach, introduced in this work, integrates the optimization framework of implicit generative priors with physical constraints from a modified CDI setup. This combination results in high-SNR reconstructions while preserving resolution at low photon counts in strong presence of noise where traditional methods are known to struggle. Experiments demonstrate that LoDIP’s performance is consistent across different natural images, realistic biological cells and experimental data. We expect the LoDIP method to find applications in X-ray imaging of dose-sensitive samples across diverse fields including organic semiconductors and biological specimens.

#### ACKNOWLEDGMENTS

This work was supported by STROBE, a National Science Foundation Science & Technology Center, under grant number DMR-1548924, and the U.S. Air Force Office of Scientific Research Multidisciplinary University Research Initiative (MURI) program under Award No. FA9550-23-1-0281. E.N. was partially supported by the Simons Postdoctoral program at IPAM, NSF DMS-1925919, and AFOSR MURI FA9550-21-1-0084. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported in this paper.

#### REFERENCES

- [1] 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*, vol. 400, no. 6742, pp. 342–344, 1999.
- [2] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane, “Beyond crystallography: Diffractive imaging using coherent x-ray light sources,” *Science*, vol. 348, no. 6234, pp. 530–535, 2015.
- [3] G. N. George, I. J. Pickering, M. J. Pushie, K. Nienaber, M. J. Hackett, I. Ascone, B. Hedman, K. O. Hodgson, J. B. Aitken, A. Levina, C. Glover, and P. A. Lay, “X-ray-induced photo-chemistry and X-ray absorption spectroscopy of biological samples,” *Journal of Synchrotron Radiation*, vol. 19, no. 6, pp. 875–886, Nov 2012. [Online]. Available: <https://doi.org/10.1107/S090904951203943X>
- [4] E. F. Garman and M. Weik, “X-ray radiation damage to biological macromolecules: further insights,” *Journal of Synchrotron Radiation*, vol. 24, no. 1, pp. 1–6, Jan 2017. [Online]. Available: <https://doi.org/10.1107/S160057751602018X>
- [5] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, “Phase retrieval with application to optical imaging: A contemporary overview,” *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, may 2015.
- [6] M. Pham, P. Yin, A. Rana, S. Osher, and J. Miao, “Generalized proximal smoothing (gps) for phase retrieval,” *Optics Express*, vol. 27, no. 3, pp. 2792–2808, 2019.
- [7] C. T. Putkunz, J. N. Clark, D. J. Vine, G. J. Williams, M. A. Pfeifer, E. Balaur, I. McNulty, K. A. Nugent, and A. G. Peele, “Phase-diverse coherent diffractive imaging: High sensitivity with low dose,” *Physical review letters*, vol. 106, no. 1, p. 013903, 2011.
- [8] T.-Y. Lan, P.-N. Li, and T.-K. Lee, “Method to enhance the resolution of x-ray coherent diffraction imaging for non-crystalline bio-samples,” *New Journal of Physics*, vol. 16, no. 3, p. 033016, 2014.
- [9] Y. H. Lo, L. Zhao, M. Gallagher-Jones, A. Rana, J. J. Lodico, W. Xiao, B. Regan, and J. Miao, “In situ coherent diffractive imaging,” *Nature communications*, vol. 9, no. 1, p. 1826, 2018.
- [10] X. Lu, M. Pham, E. Negrini, D. Davis, S. J. Osher, and J. Miao, “Computational microscopy beyond perfect lenses,” *arXiv preprint arXiv:2306.11283*, 2023.
- [11] A. Goy, K. Arthur, S. Li, and G. Barbastathis, “Low photon count phase retrieval using deep learning,” *Physical review letters*, vol. 121, no. 24, p. 243902, 2018.
- [12] M. Deng, S. Li, A. Goy, I. Kang, and G. Barbastathis, “Learning to synthesize: robust phase retrieval at low photon counts,” *Light: Science & Applications*, vol. 9, no. 1, p. 36, 2020.
- [13] C. A. Metzler and G. Wetzstein, “Deep s 3 pr: Simultaneous source separation and phase retrieval using deep generative models,” in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021, pp. 1370–1374.
- [14] D. Ulyanov, A. Vedaldi, and V. Lempitsky, “Deep image prior,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9446–9454.
- [15] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, “Implicit neural representations with periodic activation functions,” *Advances in neural information processing systems*, vol. 33, pp. 7462–7473, 2020.
- [16] G. Jagatap and C. Hegde, “Algorithmic guarantees for inverse imaging with untrained network priors,” *Advances in neural information processing systems*, vol. 32, 2019.
- [17] R. Heckel and P. Hand, “Deep decoder: Concise image representations from untrained non-convolutional networks,” *arXiv preprint*

- arXiv:1810.03982*, 2018.
- [18] E. Bostan, R. Heckel, M. Chen, M. Kellman, and L. Waller, “Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network,” *Optica*, vol. 7, no. 6, pp. 559–562, 2020.
- [19] K. Tayal, C.-H. Lai, V. Kumar, and J. Sun, “Inverse problems, deep learning, and symmetry breaking,” *arXiv preprint arXiv:2003.09077*, 2020.
- [20] Z. Zhuang, D. Yang, F. Hofmann, D. Barmherzig, and J. Sun, “Practical phase retrieval using double deep image priors,” *arXiv:2211.00799*, 2022.
- [21] R. Hyder, Z. Cai, and M. S. Asif, “Solving phase retrieval with a learned reference,” in *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXX 16*. Springer, 2020, pp. 425–441.
- [22] J. Miao, D. Sayre, and H. Chapman, “Phase retrieval from the magnitude of the fourier transforms of nonperiodic objects,” *JOSA A*, vol. 15, no. 6, pp. 1662–1669, 1998.
- [23] J. R. Fienup, “Reconstruction of an object from the modulus of its fourier transform,” *Optics letters*, vol. 3, no. 1, pp. 27–29, 1978.
- [24] J. Miao, J. Kirz, and D. Sayre, “The oversampling phasing method,” *Acta Crystallographica Section D: Biological Crystallography*, vol. 56, no. 10, pp. 1312–1315, 2000.
- [25] H. H. Bauschke, P. L. Combettes, and D. R. Luke, “Phase retrieval, error reduction algorithm, and fienu variants: a view from convex optimization,” *Journal of the Optical Society of America A*, vol. 19, no. 7, p. 1334, jul 2002.
- [26] D. R. Luke, “Relaxed averaged alternating reflections for diffraction imaging,” *Inverse problems*, vol. 21, no. 1, p. 37, 2004.
- [27] J. A. Rodriguez, R. Xu, C.-C. Chen, Y. Zou, and J. 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.
- [28] I. McNulty, J. Kirz, C. Jacobsen, E. H. Anderson, M. R. Howells, and D. P. Kern, “High-resolution imaging by fourier transform x-ray holography,” *Science*, vol. 256, no. 5059, pp. 1009–1012, 1992.
- [29] D. A. Barmherzig, J. Sun, P.-N. Li, T. J. Lane, and E. J. Candes, “Holographic phase retrieval and reference design,” *Inverse Problems*, vol. 35, no. 9, p. 094001, 2019.
- [30] A. Sinha, J. Lee, S. Li, and G. Barbastathis, “Lensless computational imaging through deep learning,” *Optica*, vol. 4, no. 9, p. 1117, sep 2017.
- [31] T. Uelwer, A. Oberstraß, and S. Harmeling, “Phase retrieval using conditional generative adversarial networks,” *arXiv:1912.04981*, 2019.
- [32] A. Goy, K. Arthur, S. Li, and G. Barbastathis, “Low photon count phase retrieval using deep learning,” *Physical Review Letters*, vol. 121, no. 24, dec 2018.
- [33] C.-J. Wang, C.-K. Wen, S.-H. L. Tsai, and S. Jin, “Phase retrieval with learning unfolded expectation consistent signal recovery algorithm,” *IEEE Signal Processing Letters*, vol. 27, pp. 780–784, 2020.
- [34] F. Zhang, X. Liu, C. Guo, S. Lin, J. Jiang, and X. Ji, “Physics-based iterative projection complex neural network for phase retrieval in lensless microscopy imaging,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 10 523–10 531.
- [35] L. Boominathan, M. Maniparambil, H. Gupta, R. Baburajan, and K. Mitra, “Phase retrieval for fourier ptychography under varying amount of measurements,” *arXiv preprint arXiv:1805.03593*, 2018.
- [36] R. Manekar, K. Tayal, V. Kumar, and J. Sun, “End-to-end learning for phase retrieval,” *ICML workshop on ML Interpretability for Scientific Discovery*, 2020.
- [37] Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, “Phase recovery and holographic image reconstruction using deep learning in neural networks,” *Light: Science & Applications*, vol. 7, no. 2, pp. 17 141–17 141, 2018.
- [38] I. Kang, F. Zhang, and G. Barbastathis, “Phase extraction neural network (phenn) with coherent modulation imaging (cmi) for phase retrieval at low photon counts,” *Optics Express*, vol. 28, no. 15, pp. 21 578–21 600, 2020.
- [39] K. Wang, L. Song, C. Wang, Z. Ren, G. Zhao, J. Dou, J. Di, G. Barbastathis, R. Zhou, J. Zhao *et al.*, “On the use of deep learning for phase recovery,” *Light: Science & Applications*, vol. 13, no. 1, p. 4, 2024.
- [40] D. P. Kingma and M. Welling, “Auto-encoding variational bayes,” *arXiv preprint arXiv:1312.6114*, 2013.
- [41] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.
- [42] D. P. Kingma and P. Dhariwal, “Glow: Generative flow with invertible 1x1 convolutions,” *Advances in neural information processing systems*, vol. 31, 2018.
- [43] H. Chung, J. Kim, M. T. Mccann, M. L. Klasky, and J. C. Ye, “Diffusion posterior sampling for general noisy inverse problems,” *arXiv preprint arXiv:2209.14687*, 2022.
- [44] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, “Compressed sensing using generative models,” in *International Conference on Machine Learning*. PMLR, 2017, pp. 537–546.
- [45] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis,” *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.
- [46] Y. Gandelsman, A. Shocher, and M. Irani, ““ double-dip”: unsupervised image decomposition via coupled deep-image-priors,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11 026–11 035.
- [47] T. Li, H. Wang, Z. Zhuang, and J. Sun, “Deep random projector: Accelerated deep image prior,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 18 176–18 185.
- [48] Z. Shi, P. Mettes, S. Maji, and C. G. Snoek, “On measuring and controlling the spectral bias of the deep image prior,” *International Journal of Computer Vision*, vol. 130, no. 4, pp. 885–908, 2022.
- [49] Y. Lu, Y. Lin, H. Wu, Y. Luo, X. Zheng, and L. Wang, “All one needs to know about priors for deep image restoration and enhancement: A survey,” *arXiv preprint arXiv:2206.02070*, 2022.
- [50] M. Z. Darestani and R. Heckel, “Accelerated mri with un-trained neural networks,” *IEEE Transactions on Computational Imaging*, vol. 7, pp. 724–733, 2021.
- [51] A. Qayyum, I. Ilahi, F. Shamshad, F. Boussaid, M. Bennamoun, and J. Qadir, “Untrained neural network priors for inverse imaging problems: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.
- [52] G. Ongie, A. Jalal, C. A. M. R. G. Baraniuk, A. G. Dimakis, and R. Willett, “Deep learning techniques for inverse problems in imaging,” *IEEE Journal on Selected Areas in Information Theory*, 2020.
- [53] K. Tayal, R. Manekar, Z. Zhuang, D. Yang, V. Kumar, F. Hofmann, and J. Sun, “Phase retrieval using single-instance deep generative prior,” in *Applied Industrial Spectroscopy*. Optica Publishing Group, 2021, pp. JW2A–37.
- [54] C. Wang, C.-H. Lin, and S. Lucey, “Deep nrsfm++: Towards 3d reconstruction in the wild,” *arXiv:2001.10090*, 2020.
- [55] R. Manekar, Z. Zhuang, K. Tayal, V. Kumar, and J. Sun, “Deep learning initialized phase retrieval,” in *NeurIPS 2020 Workshop on Deep Learning and Inverse Problems*, 2020.
- [56] E. Gullikson. X-ray interactions with matter. [Online]. Available: <https://henke.lbl.gov/opticalconstants/>
- [57] B. Henke, E. Gullikson, and J. Davis, “X-ray interactions: Photoabsorption, scattering, transmission, and reflection at  $e = 50$ -30,000 eV,  $z = 1$ -92,” *Atomic Data and Nuclear Data Tables*, vol. 54, no. 2, pp. 181–342, 1993. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0092640X83710132>
- [58] Y. Sun, J. Liu, M. Xie, B. Wohlberg, and U. S. Kamilov, “Coil: Coordinate-based internal learning for imaging inverse problems,” *arXiv preprint arXiv:2102.05181*, 2021.
- [59] M. Chen, P. Lin, Y. Quan, T. Pang, and H. Ji, “Unsupervised phase retrieval using deep approximate mmse estimation,” *IEEE Transactions on Signal Processing*, vol. 70, pp. 2239–2252, 2022.
- [60] D. J. Chang, C. M. O’Leary, C. Su, D. A. Jacobs, S. Kahn, A. Zettl, J. Ciston, P. Ercius, and J. Miao, “Deep-learning electron diffractive imaging,” *Physical review letters*, vol. 130, no. 1, p. 016101, 2023.
- [61] J. R. Fienup, “Phase retrieval algorithms: a comparison,” *Applied Optics*, vol. 21, no. 15, p. 2758, aug 1982.
- [62] C. A. Metzler, P. Schniter, A. Veeraraghavan, and R. G. Baraniuk, “prdeep: Robust phase retrieval with a flexible deep network,” *arXiv preprint arXiv:1803.00212*, 2018.
- [63] H. Lawrence, D. A. Barmherzig, M. Eickenberg, and M. Gabriele, “Low-photon holographic phase retrieval via a deep decoder neural network,” in *Optical Sensors*. Optica Publishing Group, 2021, pp. JTU5A–19.
- [64] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [65] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. Spence, “X-ray image reconstruction from a diffraction pattern alone,” *Physical Review B*, vol. 68, no. 14, p. 140101, 2003.