LoDIP: Low light phase retrieval with deep image prior

Raunak Manekar[1*†], Elisa Negrini[2†], Minh Pham[3], Daniel Jacobs[3], Jaideep Srivastava[1], Stanley J. Osher[2], Jianwei Miao[3]

[1*]Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA.
[2]Mathematics, University of California, Los Angeles, Los Angeles, CA, USA.
[3]Physics and Astronomy, and California NanoSystems Institute, University of California, Los Angeles, CA, USA.

*Corresponding author(s). E-mail(s): manek009@umn.edu;
Contributing authors: enegrini@math.ucla.edu; minhrose@ucla.edu; danieljacobs@physics.ucla.edu; srivasta@umn.edu; sjo@math.ucla.edu; miao@physics.ucla.edu;
†These authors contributed equally to this work.

Abstract

Phase retrieval (PR) is a fundamental challenge in scientific imaging, enabling nanoscale techniques like coherent diffractive imaging (CDI). Imaging at low radiation doses becomes important in applications where samples are susceptible to radiation damage. However, most PR methods struggle in low dose scenario due to the presence of very high shot noise. Advancements in the optical data acquisition setup, exemplified by in-situ CDI, have shown potential for low-dose imaging. But these depend on a time series of measurements, rendering them unsuitable for single-image applications. Similarly, on the computational front, data-driven phase retrieval techniques are not readily adaptable to the single-image context. Deep learning based single-image methods, such as deep image prior, have been effective for various imaging tasks but have exhibited limited success when applied to PR. In this work, we propose LoDIP which combines the in-situ CDI setup with the power of implicit neural priors to tackle the problem of single-image low-dose phase retrieval. Quantitative evaluations demonstrate the superior performance of LoDIP on this task as well as applicability to real experimental scenarios.

Keywords: deep image prior, deep generative models, computational imaging, phase retrieval, inverse problems, low light imaging, low photon count

1 Introduction

Coherent diffractive imaging (CDI) is a lensless imaging technique [36] 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 specimens to nanoscale objects [38].

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 material, polymers or organic semiconductors, require minimal radiation exposure to prevent damage during

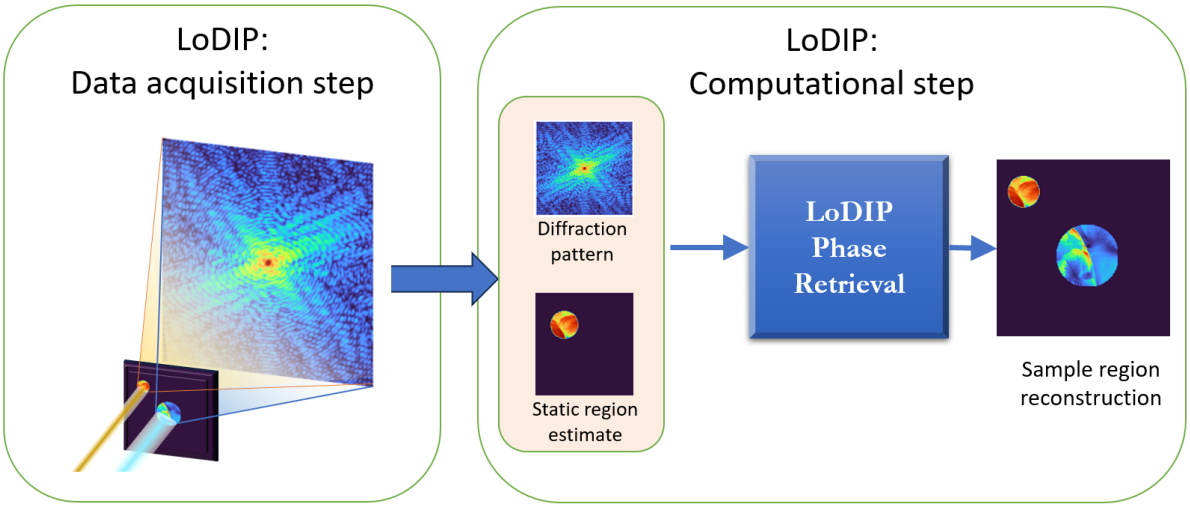

Fig. 1 Coherent Diffraction Imaging (CDI) employs a coherent X-ray beam directed at a sample, capturing the resulting diffraction pattern on a 2D detector. A computational algorithm is then applied to reconstruct the desired sample image. Inspired by in-situ CDI [28], LoDIP introduces two modifications 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 through the use of an attenuator. 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.

data acquisition [15, 16]. Thus, there is considerable interest in techniques that minimize sample radiation exposure and enable X-ray imaging at extremely low photon counts.

The main challenge is that imaging at low photon counts leads to very high shot noise in the acquired measurements (diffraction pattern) and makes the subsequent image reconstruction part (i.e. phase retrieval) very challenging. Under high shot noise, iterative-based phase retrieval algorithms, including ER, HIO, become unstable, facing challenges such as getting trapped in local minima, stagnation, and failing to converge (keep oscillating).

One category of proposed solutions for the low-dose challenge involve modifications to the optical data acquisition setup of CDI [27, 28, 41]. A notable example is in-situ CDI, as introduced by Lo et al. [28] and further explored in [29]. A dose-tolerant static region next to the sample and is illuminated with a much higher radiation dose. This results in ample light falling on the detector while keeping a low radiation dose on the sample. Multiple such measurements are taken over time. The static region present in the measurement from each time step gives a very strong time-overlap constraint for regularizing the phase retrieval optimization. However, this method requires multiple measurements. Adapting this setup for single-image scenarios results in suboptimal reconstruction (see HIO-stat column in Fig. 3).

Another possibility could be to modify the computational algorithm. Data-driven approaches have shown potential for simpler versions of phase retrieval at low-photon counts [8, 19]. But they cannot be readily adapted to the single-image context due to the requirement for a large training dataset. Deep learning based single-image methods such as deep image prior(DIP) [9, 51] are very relevant to our single-image context. While they have demonstrated success in simpler phase retrieval scenarios [4, 20, 53], they begin to struggle for more challenging situations, as evident in our experiments Fig. 3 and consistent with previous observations[48, 55].

To address these issues, in this work we propose a deep learning method designed specifically to tackle the challenging case of low-dose CDI phase retrieval. More specifically, we propose LoDIP, a single-image method for low-dose phase retrieval based on deep image prior and inspired by in-situ CDI. We use the modified

imaging setup consisting of a static region illuminated with a high dose of radiation alongside the sample of interest which is illuminated by low-light. This increases the amount of light incident on the detector, without increasing the radiation dose on the sample. Then we modify the DIP framework to exploit the additional constraints coming from this setup. Experiments show that the proposed method LoDIP can reliably work for low-dose phase retrieval both for simulated and experimental data.

2 Background

2.1 Coherent Diffraction Imaging

In Coherent Diffraction 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. While detectors capture the magnitude of the diffracted wave, the phase information is lost. As a consequence, the process of reconstructing an 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. If the diffraction pattern is sufficiently oversampled [35], the phase can be in principle retrieved from the diffraction pattern via iterative algorithms [45].

In the general phase retrieval setup, given an image of an object $\boldsymbol{X} \in \mathbb{C}^{n\times n}$, the measurement process to capture a diffraction pattern $\boldsymbol{Y} \in \mathbb{R}^{m\times m}$ can be described as:

$$\boldsymbol{Y} = \left|\mathcal{F}\left(\boldsymbol{X}\right)\right|^2 \tag{2.1}$$

To meet the oversampling criteria which helps to mitigate the non-uniqueness of the phase retrieval problem, the original object $\boldsymbol{X} \in \mathbb{C}^{n\times n}$ is zero-padded to size $m \times m$, where $m = 2 \times n$. The overall goal of the phase retrieval problem is to recover the image $\boldsymbol{X}$ from the captured diffraction pattern $\boldsymbol{Y}$.

Generally, some information about the support of the object (i.e. the location of the object) within this zero-padded image is known. Let $\boldsymbol{S}_0$ be the known support information. Note that this may not be the exact support of the object and thus still leave some translation freedom. The corresponding optimization problem is given by:

$$\begin{aligned} &\min_{\hat{\boldsymbol{X}}\in\mathbb{C}^{n\times n}} \ \ell\left(\boldsymbol{Y}, \left|\mathcal{F}\left(\hat{\boldsymbol{X}}\right)\right|^2\right), \\ &s.t.\ (1-\boldsymbol{S}_0)\odot\hat{\boldsymbol{X}} = [\boldsymbol{0}]_{m\times m} \end{aligned} \tag{2.2}$$

where the first term imposes that 2.1 is satisfied, while the second term is the support constraint. The objective function in the above formulation can be considered as the magnitude constraint or data consistency term. And the support constraint utilize the information coming from oversampling, requiring the off-support values to be zero. The objective function can also be written in terms of the Fourier magnitudes (not squared), i.e. $l(\sqrt{\boldsymbol{Y}}, |\mathcal{F}(\boldsymbol{X})|)$.

Imaging at cellular or atomic scales necessitates use of X-ray radiation. However, many samples of interest for CDI, such as biological material, polymers or organic semiconductors, require minimal radiation exposure to prevent damage during data acquisition. As a consequence,there is considerable interest in techniques that minimize sample radiation exposure and enable X-ray imaging at extremely low-photon counts.

2.2 Phase retrieval for low-light imaging

The major challenge of working in the low-dose case is the strong presence of noise in the captured signal. In practice, most of the existing PR methods [2, 11, 31, 37, 40, 43] struggle to work in the low dose setting (high noise). 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 (unless additional strategies are adopted like the shrink-wrap method [33]).

There have been attempts at low-photon phase retrieval that modify the imaging setup of the problem [27, 28, 41]. While [41] relies on a phase diverse approach, [27, 28] use a template object made up of heavy metals which can withstand high energy light, into the imaging setup. Of these, in-situ CDI [28] achieves a substantial order of magnitude dose reduction over other methods. Its primary objective is to monitor dynamic processes over time, employing two distinct regions: a dynamic region encompassing the sample, which

exhibits continuous changes over time, and a static region that remains stationary. The static region is subject to significantly higher illumination compared to the sample, ensuring ample light on the detector while keeping the radiation dose low for the sample. A series of diffraction patterns is collected over time, resulting in interference between the static and dynamic regions. The static region's presence in each time step's measurement offers a strong overlap-in-time constraint for facilitating the regularization of phase retrieval optimization.

The static region generates a time-overlap constraint to reconstruct the motion of low-dose samples such as glioblastoma cells and material science sample through Oversampling Smoothness (OSS) algorithm [43]. Some existing works [8, 18] have shown that deep neural networks (DNN) coupled with the supervised learning paradigm can show improvement over the traditional algorithms in the low-light setting. However, as explained before, these require a large dataset for training the DNN. Thus, it is not straightforward to use these techniques for single-image phase retrieval setting. This work specifically addresses the single-image phase retrieval scenario, which which does not allow for the classical supervised learning techniques.

2.3 Deep Learning for phase retrieval

Previous research [8, 18] has demonstrated that deep neural networks (DNN) within a supervised learning framework can enhance performance in low-light conditions. However, these methods necessitate substantial training dataset, rendering them unsuitable for single-image phase retrieval. Our work focuses on single-image phase retrieval, a context unsuitable for traditional supervised learning techniques due to the absence of multiple images for training.

Recent data-driven deep learning methods for PR [47, 50], despite some initial successes are yet to see widespread adoption among microscopy practitioners. The main reason is that model training necessitates an appropriate dataset, huge training time, GPU resources and also relevant expertise in training deep neural networks. Moreover, the reconstruction performance on the target sample depends heavily on the quality of data provided for the model training. Additionally, each model has to be retrained to be used on samples captured under different experimental settings. Thus, PR practitioners raise a common question whether DL can be used to improve the methods for single-image phase retrieval, i.e. without requiring training on a separate dataset. This work proposes such a DL-based method which can function in a single-image setting.

2.3.1 Neural Networks as a prior

After the success of convolutional neural networks(CNNs) on image classification task [26], neural networks and particularly CNNs have improved the state-of-the-art in many computer vision tasks. Most of these were in the supervised learning paradigm and used a large dataset for training the neural networks. Gradually, neural networks also succeeded in generative modelling and approximating distributions of images [17, 23, 24]. This led to the use of such trained CNNs as a prior for regularizing inverse problems in computer vision [3]. Since these methods involve the use of CNN models that have already been pre-trained on a different task on a similar data, these can be used for inverse problems in a single-sample setting, i.e. without separate training on a collected dataset. Around the same time it was discovered that even an untrained neural network can be used as a strong prior for natural images [9, 51] and can be successfully used to regularize inverse problems on natural images.

The surprising success of untrained neural prior or deep image prior(DIP) on inverse problems in image restoration and enhancement has been followed by works aimed at understanding its underlying mechanisms [6]. and also extensions which build upon different aspects of the original deep image prior method. Subsequent research has resulted in various extensions and adaptations, including approaches involving early stopping strategies [20, 54], methods addressing spectral bias improvements [46], and innovative applications [14]. For a comprehensive overview of the extensive body of work on untrained neural priors in image enhancement, we recommend a recent comprehensive survey [30].

Untrained neural priors have also been popularly applied to various computational imaging inverse problems [7, 39, 42] including simpler versions of phase retrieval [4, 20, 49, 53]. But

these works tackle much simpler settings for PR (e.g. near-field Fresnel diffraction), opposed to the setting used in this paper (far-field Fraunhoffer diffraction). Untrained neural networks have been used for far field diffraction imaging with only limited success [49].

In this work, we adopt the in-situ CDI setup for the single-image PR and supplement it with an untrained neural network prior in the computational algorithm. Our method is able to produce accurate reconstructions even in the low dose setting, thanks to the in-situ CDI set up and is able to produce results for single-image phase retrieval since the untrained neural network does not require any data. Experiments on simulated and experimental data demonstrate the efficacy of this method over either of these methods alone.

3 Proposed method

The proposed method LoDIP modifies the data acquisition setup of conventional CDI to incorporate a high-dose static region based on the in-situ CDI setup. While the original in-situ CDI leverages the static region as a time-invariant constraint to reconstruct a sequence of dynamic process measurements, the high-dose static region serves a very different purpose in low-dose CDI.

Firstly, it increases the available light for image formation on the detector, effectively reducing the impact of shot noise. Additionally, given its high-dose illumination, it is easy to get a good reconstruction of the static region. This known static region can be used as a strong constraint to ease the subsequent phase retrieval optimization. Moreover, it mitigates the fundamental difficulty of phase retrieval, i.e. the trivial ambiguities arising from symmeteries in the forward process [21, 32, 45].

While these adaptations significantly enhance the performance of established phase retrieval methods (e.g., HIO) at low radiation doses, our experiments (see Fig. 3) underscore the need for further improvement in the final reconstruction quality. To address this, LoDIP incorporates an untrained neural prior for phase retrieval.

3.1 LoDIP: Data acquisition 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 while the static region is exposed to the full dose of the incident illumination. Far-field diffraction patterns recorded from this setup are formed 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.

The proposed method LoDIP does not depend on the exact placement of the object and the static structure as long as they do not overlap and the support information is approximately known. Let the operator $\mathcal{S}(\cdot)$ extract the support (location of non-zero elements) of the actual object in the oversized (zero-padded) image. Then the object and the static structure should have a non-overlapping support, i.e. $\mathcal{S}(\boldsymbol{X}) \odot \mathcal{S}(\boldsymbol{U}) = [\boldsymbol{0}]_{m\times m}$.

Given the high-dose illumination on the static region and the known support, it is straightforward to obtain a high-quality static region reconstruction. This can be achieved in many ways. Here use an established iterative method, Generalized Proximal Smoothing (GPS) initalized with 1000 iterations of HIO.

3.2 LoDIP: Phase Retrieval

For LoDIP, the known static region in the above data acquisition step provides us with a useful constraint to regularize the optimization problem. By incorporating the static structure $\boldsymbol{U}$ in the original PR formulation, the measurement operator(or forward operator) is changed to :

$$\boldsymbol{Y} = |\mathcal{F}(\boldsymbol{X} + \boldsymbol{U})|^2 \tag{3.1}$$

The optimization problem now becomes

$$\begin{aligned} \min_{\hat{X}\in\mathbb{C}^{n\times n}} \ & \ell\left(\boldsymbol{Y}, \left|\mathcal{F}\left(\hat{X} + \boldsymbol{U}\right)\right|^2\right), \\ s.t. \ & (1 - \boldsymbol{S}_0) \odot \hat{\boldsymbol{X}} = [\boldsymbol{0}]_{m\times m} \end{aligned} \tag{3.2}$$

For an inverse imaging problem for recovering an image $\mathcal{X} \in \mathbb{C}^{n\times n}$ from its measurements $\mathcal{Y} \in$

$\mathbb{C}^{n\times n}$, deep image prior can be formulated as:

$$\min_{\boldsymbol{W}} \ell\left(\boldsymbol{Y}, \mathcal{A}\left(g_{\boldsymbol{W}}\left(\boldsymbol{z}\right)\right)\right), \tag{3.3}$$

where $\boldsymbol{Y} = \mathcal{A}\left(\boldsymbol{X}\right) + \eta$ and $\mathcal{A}$ is the known forward operator, η is the noise. The optimization variable $\boldsymbol{X}$ is reparameterized with a new function $g_{\boldsymbol{W}}\left(\boldsymbol{z}\right)$. $\boldsymbol{W}$ represents the learnable parameters of the new function, generally a convolutional neural network and $\boldsymbol{z}$ is a random input seed which is fixed throughout the optimization process. In this work we choose z to be the recorded diffraction pattern.

Using the flexibility of the deep image prior framework, we extend it to incorporate the constraints coming from the above setup: the known static region, the known sample support as well as the relative illumination doses on the sample vs the static region. The differential doses on the sample and the static region lead to very different range of pixel values for $\boldsymbol{X}$ and $\boldsymbol{U}$. The sample has orders of magnitude lower pixel values. Adding a scaling factor (k) to the optimization objective (Section 3.2) can bring the expected output of the neural network back to the range $[0, 1]$ which greatly helps the optimization process.

$$\begin{aligned} \min_{\boldsymbol{W}} \ell\left(\boldsymbol{Y}, \left|\mathcal{F}\left(k * g_{\boldsymbol{W}}\left(\boldsymbol{z}\right) + \boldsymbol{U}\right)\right|^2\right), \\ s.t.\left(1 - \boldsymbol{S}_0\right) \odot \hat{\boldsymbol{X}} = [\boldsymbol{0}]_{m\times m} \end{aligned} \tag{3.4}$$

Following the literature on this topic, we have used a U-Net [44] with skip connections and ReLU activation functions. But any suitable architecture can be used in the proposed setup. LoDIP can also be adapted to be used with multiple measurements of the same sample or a sequence of measurements from a dynamic processs. LoDIP can seamlessly incorporate a rough reconstruction of the sample as initialization. As a result, it can be effectively combined with other techniques designed to generate high-quality initializations for iterative phase retrieval methods [32].

Importantly, LoDIP can also be modified to work in the presence of experimental non-idealities such as the presence of a probe function by incorporating them in the forward operator. This makes LoDIP usable in diverse experimental conditions. This is in contrast to existing state-of-the-art methods such as GPS [40].

Moreover, the performance of LoDIP is not sensitive to the relative size or relative location of the sample and the static structure, nor to the specific choice of the static structure. A static structure design step such as in [1, 21] has the potential to improve the performance but is not necessary. In fact in our experiments we obtain accurate reconstructions independent of the choice of the static structure. Finally, unlike Fourier holography [1, 34], the proposed method works with different illuminations of the sample and the static structure.

4 Experimental Results and Discussion

4.1 Data

We perform experiments on three kinds of data. First, we create a simulated data[1] using a procedure similar to [5]. Stock images from the internet are used to create the sample region and the static region for each of the 50 images in the dataset. Examples of the generated images can be seen in Fig. 2. Next, we create a simulation images of biological cell using physically accurate simulations of a gold lacey for the static structure and a biological cell for the sample region[2]. Fig. 5.

Finally, we demonstrate the applicability of LoDIP on experimental diffraction patterns[2] from live glioblastoma cells measured with a 534 nm HeNe laser. This data was reused with permission from [28]. The static structure is a 100 µ pinhole exposed to the same incident illumination as the sample. Unlike the first two datasets above, this data has been collected using a probe. Further information about the generation of simulated data and optical laser data collection can be found in their original paper [28].

Simulation geometry

The simulation geometry used is similar to the setup of [28]. The size of the entire image (here, 256x256) corresponds to the size of the detector in the real experimental setup. Each detector pixel corresponds to an area of 10µm by 10µm.

[1] Data avaialble at https://github.com/raunakmanekar/lodip

[2] Data avaialble at https://www.physics.ucla.edu/research/imaging/ISCDI/index.html

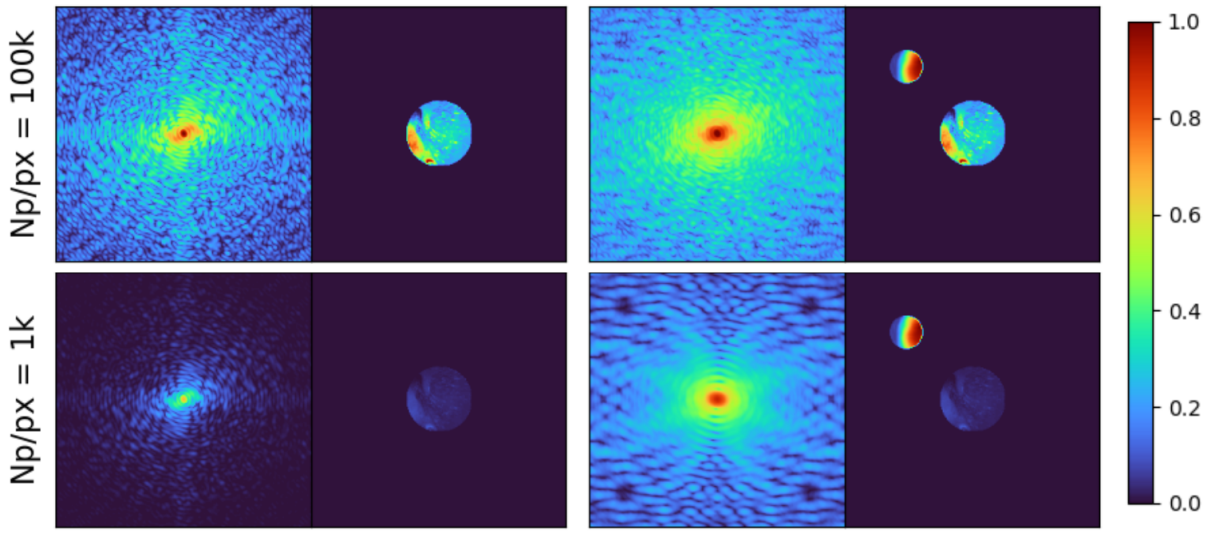

Fig. 2 Examples from the simulated data at different levels of illumination of the sample region as measured in photons per pixel (Np/px). The first two columns are the diffraction pattern and the image sample without a static structure. And the last two columns are the same sample with a static region. It can be seen that at lower illumination (bottom row), the sample has very low pixel values and it is hard to make out high resolution details.

Based on this detector pixel size and the a pinhole size of 3µm, the support of the sample region is 70x70 in pixels. Thus giving an oversampling ratio of approximately 3.6 which is higher than the required minimum 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.

Simulating low-light conditions

The illumination on the sample has been varied from 10^3 to 10^7 photons per µm^{-2}. The illumination on the static structure has been fixed at 10^{10} photons per µm^{-2}. This decides the relative scaling of the sample image and the static image. This can be seen in Fig. 2 columns two and four where the sample region becomes darker in the bottom row image, whereas the static image remains the same. Note that as the sample is imaged in lesser light, the diffraction pattern is also darker. Most of the pixels are close to zero and thus, Poisson noise is very high. The illumination in photons per pixel has been calculated by dividing this number of incident photons on the entire sample area by the total number of pixels containing the sample. The lighting conditions have been graded in number of incident photons per pixel (Np/px in Fig. 2). Based on the total number of incident photons, Poisson noise has been applied. Additive white Gaussian noise is commonly used as an approximation for Poisson noise. But this approximation is valid only at high photon counts. Thus, at low photon counts, we have to directly sample from the Poisson distribution with mean equal to the number of incident photons on a pixel.

$$N_p/px = \frac{\#\text{ incident photons}}{\#\text{ pixels}} \tag{4.1}$$

$$s = \frac{Np}{\sum_{i,j} Y_{i,j}} \tag{4.2}$$

$$\tilde{Y} = Poisson(\boldsymbol{Y} \cdot s) \cdot \frac{1}{s} \tag{4.3}$$

4.2 Results

For all experiments we compare the relative performance of the proposed method (LoDIP) with other popular methods which can be used in this setup both in the high and low photon count case. Specifically, we compare with Hybrid input-output (HIO) [12]; Deep Image Prior [51] (DIP) with no static region information; HIO-stat which is a modification of HIO to work in the LoDIP experimental setup; and Generalized proximal smoothing (GPS) [40] which is the state of the

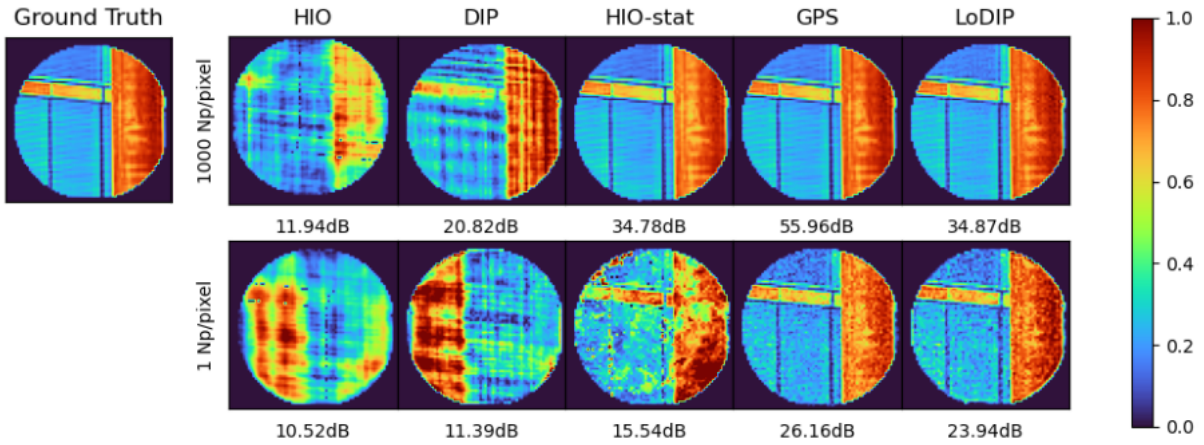

Fig. 3 Experimental Results on simulated data. **(Top row)** Reconstruction at high photon count (1000 photon/pixel). **(Bottom row)** Reconstruction at low-photon counts (1 photon/pixel). For each method we report the peak signal-to-noise ratio (PSNR), larger the better. Each image shows a zoomed-in view of only the sample region. The first two methods(HIO and DIP) use the conventional CDI setup(Section 2.1) whereas HIO-stat, GPS and LoDIP use the LoDIP setup(Section 3.1).

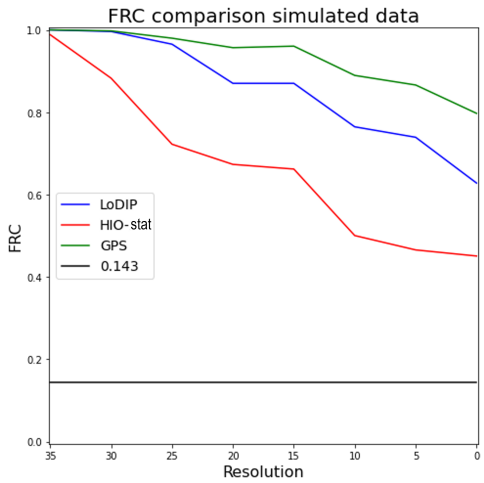

Fig. 4 Comparing FRC values for simulated data with a photon count of 1 photon per pixel. Larger FRC values indicate better reconstruction.

art method for phase retrieval. HIO, DIP do not use the LoDIP experimental setup and therefore work with diffraction patterns collected without a static region (see column 1 in Fig. 2. In contrast, HIO-stat and LoDIP work with the LoDIP setup as well as use the available static region estimate during reconstruction.

Reconstruction of simulated data

To test the robustness and generality of our method, the selected methods are tested on a dataset of 50 samples generated using the above procedure. Each sample uses a different stock image for its static structure and sample region. Table 1 shows the quantitative performance averaged over the entire test set. We measure the reconstruction accuracy using peak signal-to-noise ratio (PSNR). PSNR measures the pixel-wise error in the image domain and a larger PSNR means better reconstruction. We can see that in the low-dose case LoDIP attains better PSNR than HIO-stat and comparable PSNR to GPS.

Table 1 Quantitative comparison of LoDIP and HIO-stat performance. The reported PSNR values represent the mean of PSNR scores on 50 different simulated samples. Larger PSNR means better reconstruction.

	Np/px=1000	Np/px=1
HIO-stat	37.86dB(±3.4)	13.88dB(±2.4)
LoDIP	40.53dB(±2.8)	23.56dB(±2.4)
GPS	52.32dB(±6.3)	25.59dB(±1.43)

Fig. 3 displays a direct comparison of the results of LoDIP to reconstructions from static structure-free HIO, static structure-free DIP, and HIO with strict enforcement of the static structure. As a proof of concept, Fig. 3 uses stock images for both the static and the sample regions to create simulated diffraction patterns using the forward model described in Section 3 and Section 4.1. To improve the performance of HIO for an adequate comparison, we performed 100 independent HIO reconstructions, then averaged

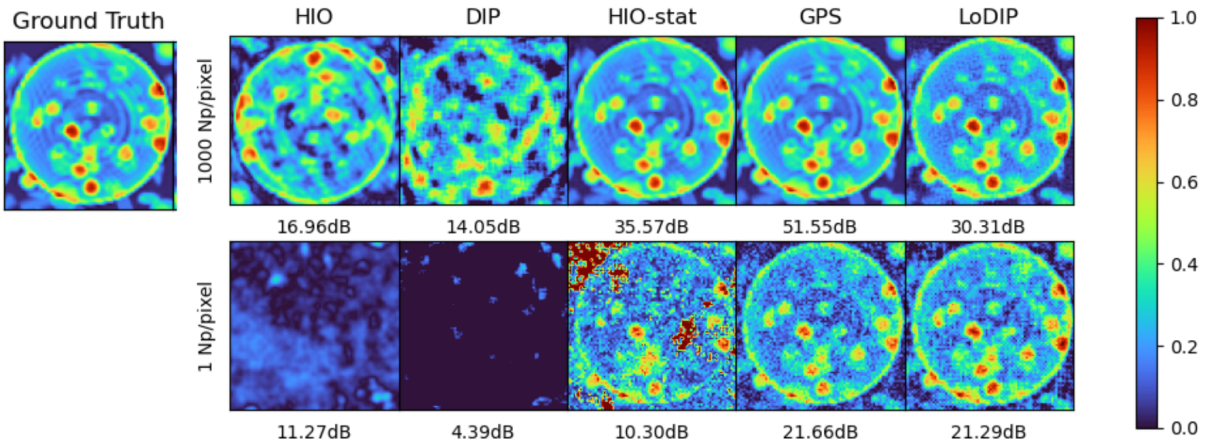

Fig. 5 Experimental Results on biological cell sample. **(Top row)** Reconstruction at high photon count (1000 photon/pixel). **(Bottom row)** Reconstruction at low-photon counts (1 photon/pixel).

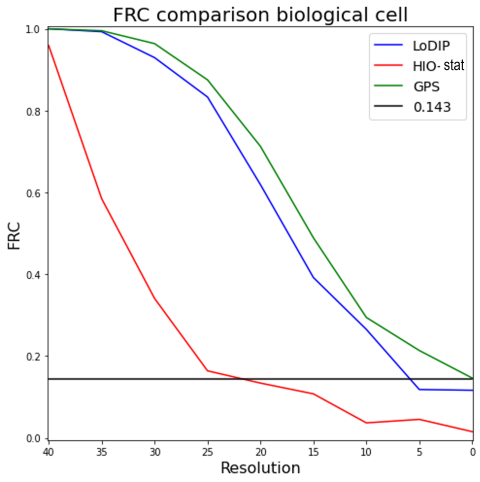

Fig. 6 Comparing FRC values for simulated data with a photon count of 1 photon per pixel. Larger FRC values indicate better reconstruction.

the best 5 of them to produce the images shown in Fig. 3.

Fig. 4 displays the Fourier Ring Correlation (FRC) for the low-photon counts case (1 photon/pixel) for HIO with static structure (HIO-stat) and LoDIP. FRC is a technique used in imaging to assess the resolution and overall quality of a reconstruction [52]. It calculates the correlation coefficient between the Fourier transforms of two images, measured within concentric rings of varying spatial frequencies. The resolution is assessed based on the drop in correlation at high frequencies, with a threshold of 0.143. Higher FRC values correspond to better resolution and reconstruction quality. Also according to this metric LoDIP obtains better resolution than HIO-stat and comparable resolution to GPS.

Reconstruction of biological sample

As biological samples represent a major potential application of the proposed method, 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 [28]. The results shown in Fig. 5 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. Fig. 5 shows the reconstructions from the same four methods as in Fig. 3 at a dose of 1 and 1000 photons/pixel. The static region is estimated using the procedure described in Section 3.1 and assumed to be unknown a priori.

Finally, Fig. 6 displays the Fourier Ring Correlation (FRC) for the low-photon counts case (1 photon/pixel) for HIO with static structure (HIO-stat) and LoDIP. Again LoDIP gives better results than HIO-stat and comparable results to GPS.

Reconstruction from experimental data

Finally, we perform a proof-of-concept experiment to demonstrate the proposed method on experimentally captured diffraction patterns (shown in Fig. 7 left column). Unlike the previous experiments, the object is complex-valued and the

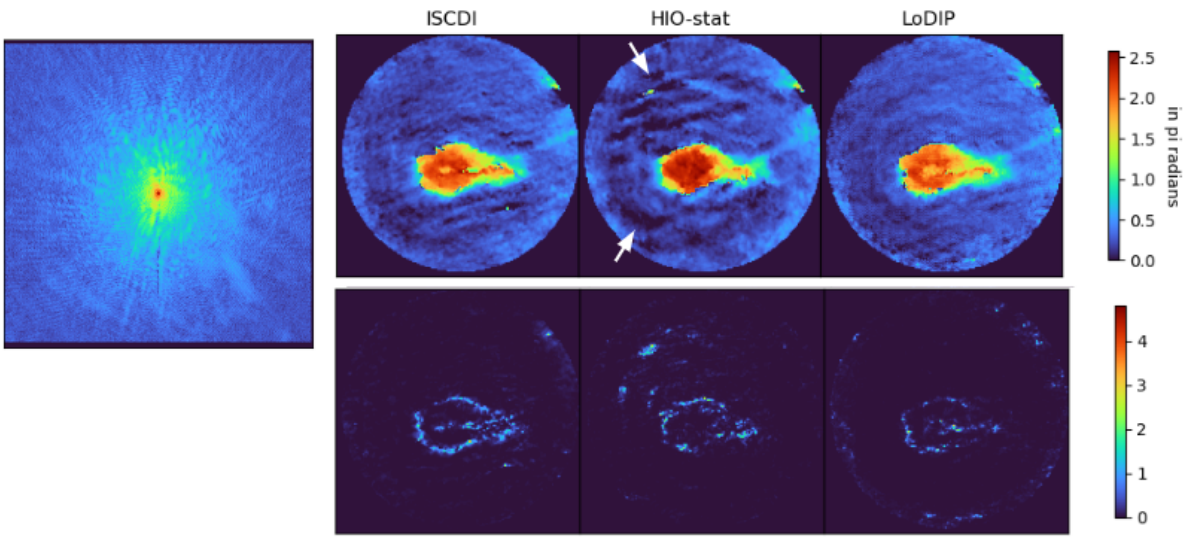

Fig. 7 Reconstruction using experimental data **(Left)** Experimentally captured diffraction pattern. **(Right)** Means (Top) and variance (Bottom) calculated of the top five reconstructions from 20 independent runs. The proposed method LoDIP uses a single diffraction pattern to produce a reconstruction of comparable quality as in-situ CDI which uses 50 samples. HIO reconstruction has similar R_F but contains severe artefacts (indicated by the white arrows)

.

optical setup includes a probe. The in-situ CDI method [28] uses 50 diffraction patterns with a fixed time-invariant static structure required in all the images. HIO-stat and LoDIP use a single diffraction pattern. A comparison with GPS is not possible here since its established version does not support Fresnel propagation. In these experiments, LoDIP and HIO-stat have been modified to incorporate the known probe function (Section 4.1). Since there is no ground truth available, we use the relative error (R-factor) calculated in the Fourier domain:

$$R_F(\hat{\boldsymbol{X}}) = \frac{\sum_{i,j} \left| |\mathcal{F}\left(\hat{\boldsymbol{X}}\right)|_{i,j} - \boldsymbol{Y}_{i,j} \right|}{\sum_{i,j} Y_{i,j}} \tag{4.4}$$

The R-factor R_F is a measure of the similarity between the captured diffraction pattern $\boldsymbol{Y}$ to the Fourier magnitudes $|\mathcal{F}\left(\hat{\boldsymbol{X}}\right)|$ of the reconstruction $\hat{\boldsymbol{X}}$. A smaller R-factor represent a better reconstruction. Table 2 shows the average R_F from 20 independent reconstructions of a single diffraction pattern. LoDIP gives performance comparable to in-situ CDI without requiring multiple diffraction patterns. Fig. 7 shows the means and variance image obtained from the top 5 reconstructions out of 20 independent runs. While the reconstruction quality is visually comparable to the other methods, we can see from the bottom row of Fig. 7 that LoDIP reconstructions have lower variance and thus are more consistent. Note also that the reconstruction from HIO-stat has comparable R_F but shows strong visible artefacts.

Table 2 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
In-situ CDI	30.22%(±0.97%)
HIO-stat	32.70%(±0.03%)
LoDIP	33.20%(±0.03%)

4.3 Discussion

The LoDIP approach presented in this work combines the optimization framework of untrained neural priors with an accurate forward model to recover the image of a sample from a measured

far-field diffraction pattern under conditions that prevent proper convergence in well-characterized iterative algorithms. We have demonstrated that this method yields an improvement in reconstruction under low-dose conditions without the need for labeled training data, tuning of hyperparameters, or supervision from a specialist in phase retrieval algorithms.

For the high illumination case of 1000 photons/pixel, it can be seen that both HIO-stat and LoDIP perform comparably both in terms of PSNR and visually (Fig. 3, top row), GPS in this case performs the best. In the low-photon regime (1 photons/pixel) the performance of all methods degrades considerably. LoDIP, both visually and quantitatively, produces comparable results to GPS and superior result to HIO-stat. This is clear both from the PSNR (Fig. 3, bottom row) and FRC values (Fig. 4). The generality of the method is also verified quantitatively in Table 1, where the average PSNR across 50 simulated samples with different samples and static structures is largest for LoDIP. This demonstrates that the method is applicable to a wide variety of samples and static structure patterns.

The same trend can be observed for the reconstruction of biological sample in Fig. 5 and Fig. 6. Despite this being an inherently harder problem, since the sample has higher resolution detail than the simulated data, LoDIP is able to accurately reconstruct the cell even at low-photon counts. Moreover, the static structure used in this case was an estimated realistic gold lacey pattern and the fact that the LoDIP method was able to recover the sample at both illumination levels, demonstrates that the method is not sensitive to the design of the static structure. The results on a simulated biological sample indicate the potential of the LoDIP method for applications to real samples in the low dose regime. The proposed method LoDIP shows no more variance than in-situ CDI. The static region estimate is obtained using the procedure described in Section 3.1 and not assumed to be known a priori. Finally, in this more realistic data setting of biological sample at low-photon counts the accuracy of the reconstruction given by GPS and LoDIP are almost identical.

We note explicitly that, while LoDIP and GPS give comparable reconstruction accuracy in the low-dose regime and only LoDIP can be used for reconstruction on the experimental data. The existing versions of GPS are not readily applicable in the presence of experimental non-idealities such as existenc of a probe function. On the other hand, the experiments on real data (Fig. 7) demonstrate that LoDIP is a much more flexible framework and can be used in diverse experimental conditions.

Limiting radiation damage

While tolerable radiation limits that avoid detrimental interference with cellular processes vary across cell types, previous studies have found that doses above 10^3 Gray can cause disruption of chloroplast structure in *Chlorella* algae and 90 percent cell death in *D. radiodurans* [10, 22, 25]. Doses above 10^4 Gray have been associated with loss of membrane function and cell viability in Chinese hamster ovarian cells, loss of myofibril contractile ability in rabbit muscle cells, and total loss of internal structure in *Chlorella* [10, 13, 25].

For the biological sample Fig. 5, the low-dose and high dose settings have been defined as 1 and 1000 photons/pixel respectively. The illumniation levels used for these settings correspond to a radiation dose of 1.97 x 10^3 Gray and 1.97 x 10^4 Gray, respectively. Thus, the experimental results in Fig. 5 show that LoDIP can be used at sufficiently low radiation doses for imaging delicate samples.

Comparison with time-overlapping methods

In-situ CDI concerns non-destructive imaging of samples under ambient conditions, allowing for acquisition of a time series and subsequent investigation of time-dependent processes. The time-overlapping information serves as a useful constraint to help the reconstruction algorithm. As LoDIP performs single-instance phase retrieval, the requirement for a time-invariant static structure is relaxed. This property is very useful for experimental setups where a truly static structure is difficult to ensure or for which a static structure is incompatible with the collection procedure (e.g. in collecting tomographic data, which requires shifting or tilting both the sample and the static structure).

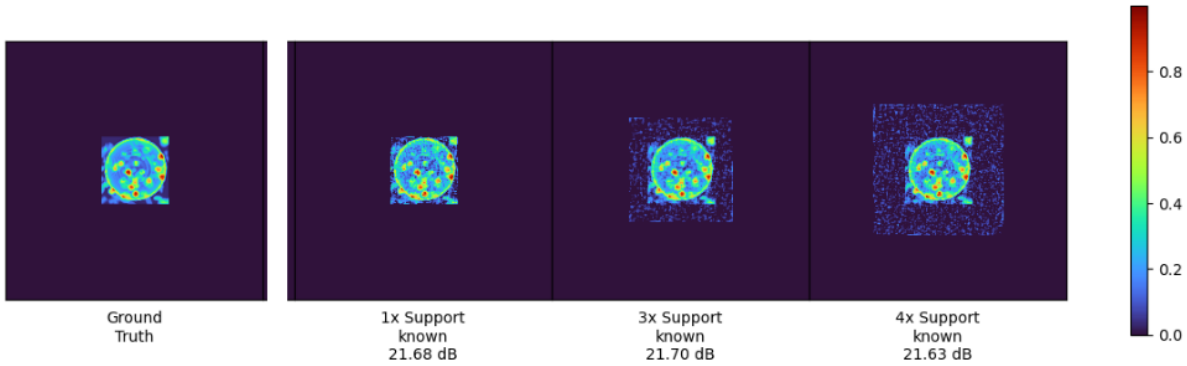

Fig. 8 Reconstruction using LoDIP without knowledge of accurate support. The columns from left to right are (1) The ground truth, (2) reconstruction with knowledge of the exact support, (3) reconstruction with an approximate support of 3x the actual size, (4) 4x the actual size. Both visually and quantitatively, the LoDIP reconstruction is robust to inaccurate specification of the support. PSNR is calculated only over the exact sample region.

Reconstruction without accurate support

In practice, image reconstruction without precise support information presents a common and challenging problem. Here we investigate the robustness of the proposed LoDIP method when dealing with only approximately known support.

In Fig. 8, the second column shows reconstruction when accurate size and shape of the support is known. The last two columns show the reconstruction when the size of the known support estimate is 3 times and 4 times the size of the actual support. From these results we can see that LoDIP reconstructions are robust to inaccurate support specification. Most widely used phase retrieval algorithms struggle when accurate support information is not available [33, 49]. While this can be partially mitigated by employing heuristic subroutine such as shrinkwrap [33] to refine the support estimate every few iterations, it adds the number of hyperparameters neeeded to be tuned. In contrast, LoDIP does not need such an additional support refinement component.

5 Conclusion and Future Work

Emerging deep learning methods present several opportunities to improve upon and expand the scope of computational imaging techniques that require processing steps framed as mathematical optimization problems. The novel LoDIP method, with its novel data acquisition setup and neural prior based phase retrieval, offers a robust approach for phase retrieval at low photon counts. As demonstrated by the PSNR and FRC curves in our simulations, the neural prior used in LoDIP outperforms other iterative methods such as HIO in image reconstruction under low-dose conditions. The LoDIP experimental setup, being derived from In-Situ CDI allows for robust phase retrieval from diffraction patterns without the resolution limitations of holography [28]. These qualities indicate that the LoDIP technique is suitable for extensions to a number of phase retrieval applications. We anticipate applications of the LoDIP method to X-ray imaging of dose-sensitive samples of importance in a variety of fields, such as organic semiconductors relevant to modern perovskite solar cells and battery materials, as well as biological samples concerning cellular interactions and life cycles.

Acknowledgments. The work is supported by STROBE: A National Science Foundation Science & Technology Center under grant number DMR 1548924. E.N. is also partially supported by the Simons Postdoctoral program at IPAM and NSF DMS 1925919 and by 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 within this paper.

References

[1] D. A. Barmherzig, J. Sun, P.-N. Li, T. J. Lane, and E. J. Candes. Holographic phase retrieval and reference design. *Inverse Problems*, 35(9):094001, 2019.

[2] H. H. Bauschke, P. L. Combettes, and D. R. Luke. Phase retrieval, error reduction algorithm, and fienup variants: a view from convex optimization. *Journal of the Optical Society of America A*, 19(7):1334, jul 2002. doi: 10.1364/josaa.19.001334.

[3] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. Compressed sensing using generative models. In *International Conference on Machine Learning*, pages 537–546. PMLR, 2017.

[4] 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*, 7(6): 559–562, 2020.

[5] 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*, 130(1):016101, 2023.

[6] Z. Cheng, M. Gadelha, S. Maji, and D. Sheldon. A bayesian perspective on the deep image prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5443–5451, 2019.

[7] M. Z. Darestani and R. Heckel. Accelerated mri with un-trained neural networks. *IEEE Transactions on Computational Imaging*, 7: 724–733, 2021.

[8] M. Deng, S. Li, A. Goy, I. Kang, and G. Barbastathis. Learning to synthesize: robust phase retrieval at low photon counts. *Light: Science & Applications*, 9(1):36, 2020.

[9] U. Dmitry, A. Vedaldi, and L. Victor. Deep image prior. *International Journal of Computer Vision*, 128(7):1867–1888, 2020.

[10] M. Du and C. Jacobsen. Relative merits and limiting factors for x-ray and electron microscopy of thick, hydrated organic materials. *Ultramicroscopy*, 184: 293–309, 2018. ISSN 0304-3991. doi: https://doi.org/10.1016/j.ultramic.2017.10.003. URL https://www.sciencedirect.com/science/article/pii/S030439911730339X.

[11] J. R. Fienup. Reconstruction of an object from the modulus of its fourier transform. *Optics letters*, 3(1):27–29, 1978.

[12] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758, aug 1982. doi: 10.1364/ao.21.002758.

[13] T. W. Ford. Imaging cells using soft x-rays. In *From Cells to Proteins: Imaging Nature across Dimensions: Proceedings of the NATO Advanced Study Institute on From Cells to Proteins: Imaging Nature across Dimensions Pisa, Italy 12–23 September 2004*, pages 167–185. Springer, 2005.

[14] 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*, pages 11026–11035, 2019.

[15] E. F. Garman and M. Weik. X-ray radiation damage to biological macromolecules: further insights. *Journal of Synchrotron Radiation*, 24(1):1–6, Jan 2017. doi: 10.1107/S160057751602018X. URL https://doi.org/10.1107/S160057751602018X.

[16] 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*, 19(6):875–886, Nov 2012. doi: 10.1107/S090904951203943X. URL https://doi.org/10.1107/S090904951203943X.

[17] 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*, 63(11):139–144, 2020.

[18] A. Goy, K. Arthur, S. Li, and G. Barbastathis. Low photon count phase retrieval using deep learning. *Physical Review Letters*, 121(24), dec 2018. doi: 10.1103/physrevlett.121.243902.

[19] A. Goy, K. Arthur, S. Li, and G. Barbastathis. Low photon count phase retrieval using deep learning. *Physical review letters*, 121(24):243902, 2018.

[20] R. Heckel and P. Hand. Deep decoder: Concise image representations from untrained non-convolutional networks. *arXiv preprint arXiv:1810.03982*, 2018.

[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*, pages 425–441. Springer, 2020.

[22] H. Ito, H. WATANABE, M. TAKEHISA, and H. IIZUKA. Isolation and identification of radiation-resistant cocci belonging to the genus ¡i¿deinococcus¡/i¿ from sewage sludges and animal feeds. *Agricultural and Biological Chemistry*, 47(6):1239–1247, 1983. doi: 10.1271/bbb1961.47.1239.

[23] D. P. Kingma and P. Dhariwal. Glow: Generative flow with invertible 1x1 convolutions. *Advances in neural information processing systems*, 31, 2018.

[24] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[25] J. Kirz, C. Jacobsen, and M. Howells. Soft x-ray microscopes and their biological applications. *Quarterly reviews of biophysics*, 28 (1):33–130, 1995.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[27] T.-Y. Lan, P.-N. Li, and T.-K. Lee. Method to enhance the resolution of x-ray coherent diffraction imaging for non-crystalline biosamples. *New Journal of Physics*, 16(3): 033016, 2014.

[28] 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*, 9(1):1826, 2018.

[29] 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.

[30] 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.

[31] D. R. Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse problems*, 21(1):37, 2004.

[32] 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.

[33] 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*, 68(14):140101, 2003.

[34] 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*, 256(5059):1009–1012, 1992.

[35] J. Miao, D. Sayre, and H. Chapman. Phase retrieval from the magnitude of the fourier transforms of nonperiodic objects. *JOSA A*, 15(6):1662–1669, 1998.

[36] J. Miao, P. Charalambous, J. Kirz, and D. Sayre. Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature*, 400(6742):342–344, 1999.

[37] J. Miao, J. Kirz, and D. Sayre. The oversampling phasing method. *Acta Crystallographica Section D: Biological Crystallography*, 56(10): 1312–1315, 2000.

[38] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane. Beyond crystallography: Diffractive imaging using coherent x-ray light sources. *Science*, 348(6234):530–535, 2015.

[39] 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.

[40] M. Pham, P. Yin, A. Rana, S. Osher, and J. Miao. Generalized proximal smoothing (gps) for phase retrieval. *Optics Express*, 27 (3):2792–2808, 2019.

[41] 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*, 106(1):013903, 2011.

[42] 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.

[43] 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*, 46(2):312–318, 2013.

[44] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[45] 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*, 32(3):87–109, may 2015. doi: 10.1109/msp.2014.2352673.

[46] 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*, 130(4):885–908, 2022.

[47] A. Sinha, J. Lee, S. Li, and G. Barbastathis. Lensless computational imaging through deep learning. *Optica*, 4(9):1117, sep 2017. doi: 10.1364/optica.4.001117.

[48] K. Tayal, C.-H. Lai, V. Kumar, and J. Sun. Inverse problems, deep learning, and symmetry breaking. *arXiv preprint arXiv:2003.09077*, 2020.

[49] 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*, pages JW2A–37. Optica Publishing Group, 2021.

[50] T. Uelwer, A. Oberstraß, and S. Harmeling. Phase retrieval using conditional generative adversarial networks. *arXiv:1912.04981*, 2019.

[51] D. Ulyanov, A. Vedaldi, and V. Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9446–9454, 2018.

[52] M. van Heel and M. Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3): 250–262, 2005. ISSN 1047-8477. doi: https://doi.org/10.1016/j.jsb.2005.05.009. URL https://www.sciencedirect.com/science/article/pii/S1047847705001292.

[53] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ. Phase imaging with an untrained neural network. *Light: Science & Applications*, 9(1):77, 2020.

[54] H. Wang, T. Li, Z. Zhuang, T. Chen, H. Liang, and J. Sun. Early stopping for deep image prior. *arXiv preprint arXiv:2112.06074*, 2021.

[55] Z. Zhuang, D. Yang, F. Hofmann, D. Barmherzig, and J. Sun. Practical phase retrieval using double deep image priors. *arXiv:2211.00799*, 2022.