



Iterative phase retrieval with a sensor mask

LI SONG AND EDMUND Y. LAM*

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China

*elam@eee.hku.hk

Abstract: As an important inverse imaging problem in diffraction optics, Fourier phase retrieval aims at estimating the latent image of the target object only from the magnitude of its Fourier measurement. Although in real applications alternating methods are widely-used for Fourier phase retrieval considering the constraints in the object and Fourier domains, they need a lot of initial guesses and iterations to achieve reasonable results. In this paper, we show that a proper sensor mask directly attached to the Fourier magnitude can improve the efficiency of the iterative phase retrieval algorithms, such as alternating direction method of multipliers (ADMM). Furthermore, we refer to the learning-based method to determine the sensor mask according to the Fourier measurement, and unrolled ADMM is used for phase retrieval. Numerical results show that our method outperforms other existing methods for the Fourier phase retrieval problem.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Recovering the missing phase from intensity-only diffraction patterns, known as phase retrieval, is an important and challenging problem in many imaging applications, such as crystallography [1], microscopy [2], holography [3,4], etc. This is because conventional optical sensors, such as charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS), can only capture the intensity information of the light proportional to the amount of photons. Moreover, measuring the phase of the electromagnetic field at $\sim 10^{15}$ Hz is nearly impossible, as no electronic device can follow such a high frequency easily.

On the other hand, the phase information is of great importance to the reconstruction of the target image. A classic example reviewed in Ref. [5] shows how swapping the Fourier phase of two images will lead to very different reconstruction results, even though the intensity pattern remains the same. Furthermore, the phase information may also reveal important structural information of the target object. For example, in biological imaging, cells are often transparent, and reconstruction of the phase information is of great importance to the analysis of the shape and structure of these cells [6].

Since only the intensity information is captured, one strategy is to encode the phase information in the diffraction pattern with additional optical setups, which is the principle behind various forms of holography [7,8]. Another strategy is to harness the prior knowledge of the target object in order to retrieve the phase information computationally with a single diffraction pattern directly, which is the focus of this work.

Mathematically, phase retrieval aims at estimating the latent image $v(x, y)$ from the noisy intensity measurement $z(x, y)$ with a given measurement operator $g(\cdot)$ related to the imaging model in the complex domain

$$z(x, y) = |g(v(x, y))| + e(x, y), \quad (1)$$

where $e(x, y)$ is the additive noise. Due to the existence of the operator $|\cdot|$ and the fact that the feasible set of the missing phase is not convex, this inverse problem is highly ill-posed and non-convex. In real applications, the measurement operator is often simplified as a linear transformation with a system matrix A .

Different optical setups will lead to different system matrices. There are mainly two different kinds: the first one is the conventional Fourier diffraction measurement [9], where A is the discrete Fourier matrix. This setup comes from the Fraunhofer diffraction equation, which indicates that the far field is related to the Fourier transform of the near field of the object. The second one is the coded diffraction measurement [10], in which A is often a concatenation of different sampling matrices. In this work, we focus on the Fourier phase retrieval problem. In such a setup, theoretically, the imaging resolution is limited by diffraction, which is determined by the wavelength of the light source, as well as the shape of the pupil and the phase across the pupil.

The combination of Fourier phase retrieval, oversampling, and high-frequency laser sources (such as X-ray free-electron laser, XFEL [11]) leads to an active research area known as coherent diffractive imaging (CDI) [12]. In this system, the imaging sensor, such as a pn-CCD [13], is used to capture the far-field diffraction patterns. Generally, its size is much larger than the target object, making the reconstruction problem highly overdetermined. As is presented in [14], a complete image has high redundancy, and removing a substantial proportion of its pixels actually makes it even easier to extract the high-level features, with higher accuracy and fewer iterations in the training process. High level features consider the probability of observing different concepts in the image, which can be easily recognized by humans [15]. Meanwhile, low-level features, or pixel-wise features, which capture the local appearance and texture statistics based on the pixels [15], are critical to the quality of the image reconstruction. Besides, Guerrero et al. [16] provide theoretical guarantees for phase retrieval from the coded diffraction patterns with equi-spaced masks, including binary masks. Kocsis et al. [17] introduce a random binary mask in the phase imaging system for wavefront separation, leading to higher resolution imaging results. Hence, by incorporating a proper sensor mask into the CDI system, we can potentially boost the iterative phase retrieval algorithm with easier feature extraction. Some researches also adopt different kinds of masks in the CDI system [18–20], but their masks are often set near the object plane.

In this work, we propose a new Fourier phase retrieval scheme, and show how a proper sensor mask can greatly improve the speed of the iterative algorithm, such as one using the alternating direction method of multipliers (ADMM) [21]. The main contributions are:

- We introduce a sensor mask to decrease the redundancy of the oversampled Fourier diffraction pattern, making it more data-efficient and suitable for the phase retrieval algorithms.
- We build on the ADMM technique to solve the masked Fourier phase retrieval problem, where the learned mask is incorporated as a strong constraint. Both the regularization term and the mask are determined by data-driven methods. Numerical experiments show that our approach can achieve better reconstruction results than other commonly-used phase retrieval algorithms for the Fourier phase retrieval problem.

2. Related work

2.1. Conventional phase retrieval methods

The foundational work of 2D Fourier phase retrieval dates back to Gerchberg and Saxton in the early 1970s [22]. Their algorithm starts with a random initial guess as the estimated phase information attached to the diffraction pattern, and then iterates in the object domain and Fourier domain alternately to update the phase information. This process can ensure that the reconstruction result matches with the captured diffraction pattern and the support constraint at the same time. Subsequent popular alternating phase retrieval algorithms, such as hybrid input-output (HIO) [9,23], mainly improve on the update rules considering the constraint and history information. However, these methods usually require more than 10 initial guesses to achieve a satisfactory result. Furthermore, when the structure of the target object is complicated,

the usual prior knowledge, such as the size and shape of the object, is too simple as a constraint for arriving at a good phase retrieval result.

Recently, the phase retrieval problem is also approached from the perspective of optimization. There are two main approaches. Some methods, such as Phaselift [24] and Wirtinger flow [25], are designed specifically for the phase retrieval problem, and have some restrictions on the imaging model of the coded diffraction pattern (CDP) measurements. For some other methods, phase retrieval is viewed as a special case of inverse imaging, tackled by solving the minimization problem

$$\text{minimize } g(v(x, y); z(x, y)) + R(v(x, y)), \quad (2)$$

where $g(v(x, y); z(x, y))$ is the data fidelity term with respect to the measurement $z(x, y)$, and $R(v(x, y))$ is the regularization term related to the prior knowledge and the constraint on the target object [26]. For phase retrieval problems, $g(v(x, y); z(x, y))$ is often specified as

$$g(v(x, y); z(x, y)) = \frac{1}{2} \|z(x, y) - |Av(x, y)|\|_2^2, \quad (3)$$

where A is the system matrix.

Various first-order algorithms, such as ADMM and its variants [27], fast iterative shrinkage thresholding algorithm (FISTA) [28], etc., can be used to solve this minimization problem. They generally do not require carefully chosen initial guesses, but the regularization term is important to the final reconstruction results. Yet, it is often not easy to find proper handcrafted regularization functions for different applications with complex-valued images.

2.2. Data-driven methods for phase retrieval

With the development of deep learning, data-driven methods are becoming more and more popular for inverse imaging problems [29,30]. For the nonlinear phase retrieval problem, deep neural networks can be used to learn a mapping from the diffraction patterns to the reconstructed images directly [31–33]. This approach is also used in some other related imaging systems, such as holography [4,34], ptychography [35,36], etc. These methods often adopt an end-to-end training, without taking advantage of the physical model. A large training set is typically required to learn the complete inverse mapping.

On the other hand, since the forward imaging model is commonly known in an inverse imaging problem, incorporating the physical model into the neural networks can often simplify the tasks to be learned. Some methods directly incorporate the forward model into the network structure, such as with the generative adversarial network [37,38], where the forward model is used to convert the generated image into the estimated measurements. Some other approaches are based on the conventional phase retrieval methods, where the neural networks are only used to determine the prior knowledge related to the data manifold. Since the regularization function is used to make a projection to the expected image manifold, some trained denoisers can be directly incorporated into the iterative algorithms [39,40]. The plug-and-play methods are based on the assumption that the trained denoiser is an effective projection operator that satisfies the prior knowledge related to the latent image. Unrolled networks have similar structures, but the main difference is that the neural networks are trained end-to-end [41], making them more suitable to a specific data manifold. Recently, a review of unfolding algorithms on phase retrieval is presented in [42]. Specifically-designed masks, such as equi-spaced masks, that satisfy theoretical conditions can improve the quality of phase recovery.

3. Phase retrieval using ADMM with a sensor mask

3.1. Mathematical model for phase retrieval

When the imaging sensor is in the far field, images on the detector plane and the object are related by a Fourier transform. We first consider the noise-free case for mathematical derivation,

then we will show that our method also works in noisy phase retrieval experiments. With the introduction of a random sensor mask m_i to the captured diffraction pattern, the imaging model can be represented as

$$z(x, y) = |m_i(x, y)\mathcal{F}(v(x, y))|, \quad (4)$$

where $v(x, y)$ is the target image in the object plane, $z(x, y)$ is the measurement image in the detector plane, and $\mathcal{F}$ denotes the 2D Fourier transform. For simplicity, we vectorize all the 2D matrices, and the equation can be rewritten as

$$\mathbf{z} = |\mathbf{m}_i \circ F\mathbf{v}|, \quad (5)$$

where the bold variables are column vectors, $\circ$ denotes the Hadamard product, and F is the Fourier matrix. The phase retrieval problem is then solved by optimization. Let ϕ be the phase information attached to the intensity pattern. Then, the optimization problem is

$$\text{minimize } \mathcal{R}(\mathbf{v}) + C(\phi) \quad (6)$$

$$\text{subject to } \mathbf{z} \circ \phi = MF\mathbf{v}, \quad (7)$$

where $\mathcal{R}(\cdot)$ is a regularization function, $C(\cdot)$ is an indicator function defined as

$$C(\phi) = \begin{cases} 0 & |\phi| = 1 \\ \infty & \text{otherwise} \end{cases}, \quad (8)$$

and M is a diagonal matrix with $M = \text{diag}(\mathbf{m}_i)$. In order to separate the undetermined regularization function from other terms, an auxiliary variable $\mathbf{x}$ is introduced as

$$\text{minimize } \mathcal{R}(\mathbf{x}) + C(\phi) \quad (9)$$

$$\text{subject to } \mathbf{z} \circ \phi = MF\mathbf{v}, \quad (10)$$

$$\mathbf{x} = \mathbf{v}. \quad (11)$$

We solve this optimization problem by using the Lagrangian multiplier, where

$$\tilde{L}(\mathbf{v}, \mathbf{x}, \phi, \tilde{\mu}_1, \tilde{\mu}_2) = \mathcal{R}(\mathbf{x}) + C(\phi) + \tilde{\mu}_1^H (\mathbf{z} \circ \phi - MF\mathbf{v}) + \frac{\rho_1}{2} \|\mathbf{z} \circ \phi - MF\mathbf{v}\|_2^2 + \tilde{\mu}_2^H (\mathbf{x} - \mathbf{v}) + \frac{\rho_2}{2} \|\mathbf{x} - \mathbf{v}\|_2^2, \quad (12)$$

where ρ_1, ρ_2 are the penalty parameters, and $\tilde{\mu}_1, \tilde{\mu}_2$ are the Lagrangian multipliers. To simplify this function, we turn to the scaled form of ADMM, where we define $\mu_1 = \tilde{\mu}_1/\rho_1$ and $\mu_2 = \tilde{\mu}_2/\rho_2$ as scaled Lagrangian multipliers, and the objective function related to the $\mathbf{v}, \mathbf{x}$ and ϕ update steps is [21]

$$L(\mathbf{v}, \mathbf{x}, \phi, \mu_1, \mu_2) = \mathcal{R}(\mathbf{x}) + C(\phi) + \frac{\rho_1}{2} \|\mathbf{z} \circ \phi - MF\mathbf{v} + \mu_1\|_2^2 + \frac{\rho_2}{2} \|\mathbf{x} - \mathbf{v} + \mu_2\|_2^2. \quad (13)$$

Its gradient, with respect to $\mathbf{v}$, is $\nabla_{\mathbf{v}}L = \rho_1 F^H M^H (MF\mathbf{v} - \mathbf{z} \circ \phi - \mu_1) + \rho_2(\mathbf{v} - \mathbf{x} - \mu_2)$. Setting this to 0, we have

$$(\rho_1 F^H M^H MF + \rho_2 I)\mathbf{v} = \rho_1 F^H M^H (\mathbf{z} \circ \phi + \mu_1) + \rho_2(\mathbf{x} + \mu_2). \quad (14)$$

Note that with the Fourier matrix, we have $FF^H = nI$, where n is the dimension of $\mathbf{z}$. Therefore, we can further simplify the equation by left-multiplying F on both sides to obtain

$$(\rho_1 n M^H MF + \rho_2 F)\mathbf{v} = \rho_1 n M^H (\mathbf{z} \circ \phi + \mu_1) + \rho_2 F(\mathbf{x} + \mu_2). \quad (15)$$

This gives the $\mathbf{v}$ -update step as

$$\mathbf{v} = F^{-1}(\rho_1 n M^H M + \rho_2 I)^{-1} [\rho_1 n M^H (\mathbf{z} \circ \phi + \mu_1) + \rho_2 F(\mathbf{x} + \mu_2)]. \quad (16)$$

For the matrix inverse term $(\rho_1 n M^H M + \rho_2 I)^{-1}$, since the matrix is diagonal, it can be easily obtained as the reciprocal of its diagonal values, and F^{-1} can be expressed by F^H as $F^{-1} = F^H/n$.

Similarly, we can obtain all the steps in the ADMM iterations with scaling [21]

$$\mathbf{v}^{(k+1)} = F^{-1}(\rho_1 n M^H M + \rho_2 I)^{-1} \left[\rho_1 n M^H (z \circ \boldsymbol{\phi}^{(k)} + \boldsymbol{\mu}_1^{(k)}) + \rho_2 F(\mathbf{x}^{(k)} + \boldsymbol{\mu}_2^{(k)}) \right], \quad (17)$$

$$\mathbf{x}^{(k+1)} = \text{prox}_{\frac{\mathcal{R}}{\rho_2}}(\mathbf{v}^{(k+1)} - \boldsymbol{\mu}_2^{(k)}) \quad (18)$$

$$\boldsymbol{\phi}^{(k+1)} = \mathcal{P}(M F \mathbf{v}^{(k+1)} - \boldsymbol{\mu}_1^{(k+1)}) \quad (19)$$

$$\boldsymbol{\mu}_1^{(k+1)} = \boldsymbol{\mu}_1^{(k)} + z \circ \boldsymbol{\phi}^{(k+1)} - M F \mathbf{v}^{(k+1)} \quad (20)$$

$$\boldsymbol{\mu}_2^{(k+1)} = \boldsymbol{\mu}_2^{(k)} + \mathbf{x}^{(k+1)} - \mathbf{v}^{(k+1)}, \quad (21)$$

where $\mathcal{P}(\cdot)$ extracts the phase information of a complex-valued image, and $\text{prox}(\cdot)$ is the proximal operator defined as [43]

$$\text{prox}_{\frac{\mathcal{R}}{\rho_2}}(\mathbf{v}) = \arg \min_{\mathbf{x}} \left(\mathcal{R}(\mathbf{x}) + \frac{\rho_2}{2} \|\mathbf{x} - \mathbf{v}\|_2^2 \right). \quad (22)$$

Generally, the regularization function $\mathcal{R}(\cdot)$ can be selected as some commonly-used terms, such as the ℓ_1 -norm, based on the assumption that the latent image is sparse. Its corresponding proximal operator is the soft thresholding, but here we need to extend it to the complex domain. The soft-thresholding operator, for a real-valued u with a threshold a , is defined as [44]

$$\mathcal{S}_a(u) = \text{sign}(u) \max\{|u| - a, 0\}, \quad u, a \in \mathbf{R}. \quad (23)$$

To extend it to the complex domain, we replace $\text{sign}(\cdot)$ with the phase function $\mathcal{P}(\cdot)$, leading to

$$\mathcal{S}_a(u) = \mathcal{P}(u) \max\{|u| - a, 0\}, \quad u \in \mathbf{C}, a \in \mathbf{R}, \quad (24)$$

which is identical with the complex shrinkage operator defined in [45,46]. The $\mathbf{x}$ -update step (Eq. (18)) can then be calculated as

$$\mathbf{x}^{(k+1)} = \mathcal{P}(\mathbf{v}^{(k+1)} - \boldsymbol{\mu}_2^{(k)}) \max \left\{ |\mathbf{v}^{(k+1)} - \boldsymbol{\mu}_2^{(k)}| - \frac{\gamma}{\rho_2}, 0 \right\}, \quad (25)$$

where γ is a scale parameter to control the relative value of the regularization term $\mathcal{R}(\cdot) = \gamma \|\cdot\|_1$.

The termination criterion is designed according to the difference of the target variables between two successive iterations, where the residual is calculated as

$$\epsilon^{(k+1)} = \frac{\|\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)}\|_1}{\|\mathbf{v}^{(k+1)}\|_1}. \quad (26)$$

When $\epsilon^{(k+1)}$ is smaller than a pre-defined error tolerance ϵ_{tol} , ADMM is expected to give a stable result.

The whole algorithm is summarized in Algorithm 1.

3.2. Unrolled network design using ADMM

Since hand-crafted or pre-trained regularization functions may not always be suitable as priors for different tasks, we further adopt unrolled networks for end-to-end training. Given an iterative algorithm such as ADMM, the unrolled network can be generated by cascading its iterations [47]. This also makes it possible to replace the relevant components in the algorithm with neural networks to incorporate the knowledge embedded in the training data. Specifically, in Algorithm 1, M is a random mask, which is not always a proper choice for a specific diffraction pattern. Instead, we make use of a neural network $\mathcal{N}_m(\cdot)$ to generate a mask according to the diffraction pattern. Similarly, the hand-crafted proximal operator is not always suitable for the target objects.

Algorithm 1. ADMM for sensor-masked Fourier phase retrieval with ℓ_1 -regularization**Algorithm 1** ADMM for sensor-masked Fourier phase retrieval with ℓ_1 -regularization

Input: z : measurement image; $\mathbf{v}^{(0)}, \mathbf{x}^{(0)}, \boldsymbol{\phi}^{(0)}, \boldsymbol{\mu}_1^{(0)}, \boldsymbol{\mu}_2^{(0)}$: initial values; ρ_1, ρ_2 : penalty parameters; γ : scale parameter; M : random mask; $iter$: number of maximum iterations, ϵ_{tol} : error tolerance.

Output: $\mathbf{v}$: reconstructed image.

```

1: for  $i = 1 : iter$  do
2:    $\mathbf{v}^{(k+1)} = F^{-1}(\rho_1 n M^H M + \rho_2 I)^{-1} [\rho_1 n M^H (z \circ \boldsymbol{\phi}^{(k)} + \boldsymbol{\mu}_1^{(k)}) + \rho_2 F(\mathbf{x}^{(k)} + \boldsymbol{\mu}_2^{(k)})]$ ,
3:    $\mathbf{x}^{(k+1)} = \mathcal{P}(\mathbf{v}^{(k+1)} - \boldsymbol{\mu}_2^{(k)}) \max \left\{ |\mathbf{v}^{(k+1)} - \boldsymbol{\mu}_2^{(k)}| - \frac{\gamma}{\rho_2}, 0 \right\}$ ,
4:    $\boldsymbol{\phi}^{(k+1)} = \mathcal{P}(M F \mathbf{v}^{(k+1)} - \boldsymbol{\mu}_1^{(k+1)})$ ,
5:    $\boldsymbol{\mu}_1^{(k+1)} = \boldsymbol{\mu}_1^{(k)} + z \circ \boldsymbol{\phi}^{(k+1)} - M F \mathbf{v}^{(k+1)}$ ,
6:    $\boldsymbol{\mu}_2^{(k+1)} = \boldsymbol{\mu}_2^{(k)} + \mathbf{x}^{(k+1)} - \mathbf{v}^{(k+1)}$ .
7:   if  $\epsilon^{(k+1)} < \epsilon_{tol}$  then
8:     break.
9:   end if
10: end for

```

To deal with this, we design another neural network $\mathcal{N}_x(\cdot)$ to learn the operator according to the latent data distribution. By learning data distribution from the trained data, the empirically designed unrolled masked ADMM architecture is exploited to build globally converged learning networks for a specific phase retrieval setup [42]. Following the idea of the unrolled network, the number of iterations should be fixed as n_{iter} . Details of this unrolled network are summarized in Algorithm 2, and also illustrated in Fig. 1.

Algorithm 2. Unrolled ADMM network for Fourier phase retrieval with a sensor mask (MADMM)**Algorithm 2** Unrolled ADMM network for Fourier phase retrieval with a sensor mask (MADMM)

Input: z : measurement image; $\mathbf{v}^{(0)}, \mathbf{x}^{(0)}, \boldsymbol{\phi}^{(0)}, \boldsymbol{\mu}_1^{(0)}, \boldsymbol{\mu}_2^{(0)}$: initial values; ρ_1, ρ_2 : learnable penalty parameters; $\mathcal{N}_m(\cdot), \mathcal{N}_x(\cdot)$: neural networks; n_{iter} : number of iterations.

Output: $\mathbf{v}$: reconstructed image.

```

1:  $M = \mathcal{N}_m(z)$ ,
2: for  $i = 1 : n_{iter}$  do
3:    $\mathbf{v}^{(k+1)} = F^{-1}(\rho_1 n M^H M + \rho_2 I)^{-1} [\rho_1 n M^H (z \circ \boldsymbol{\phi}^{(k)} + \boldsymbol{\mu}_1^{(k)}) + \rho_2 F(\mathbf{x}^{(k)} + \boldsymbol{\mu}_2^{(k)})]$ ,
4:    $\mathbf{x}^{(k+1)} = \mathcal{N}_x(\mathbf{v}^{(k+1)} - \boldsymbol{\mu}_2^{(k)})$ ,
5:    $\boldsymbol{\phi}^{(k+1)} = \mathcal{P}(M F \mathbf{v}^{(k+1)} - \boldsymbol{\mu}_1^{(k+1)})$ ,
6:    $\boldsymbol{\mu}_1^{(k+1)} = \boldsymbol{\mu}_1^{(k)} + z \circ \boldsymbol{\phi}^{(k+1)} - M F \mathbf{v}^{(k+1)}$ ,
7:    $\boldsymbol{\mu}_2^{(k+1)} = \boldsymbol{\mu}_2^{(k)} + \mathbf{x}^{(k+1)} - \mathbf{v}^{(k+1)}$ .
8: end for

```

In Algorithm 2, we use two neural networks: $\mathcal{N}_m$ for the mask generation and $\mathcal{N}_x$ for the $\mathbf{x}$ -update, respectively. As is shown in Fig. 1, the generated mask has the same size as the diffraction pattern, and the pixel values in the mask show their weights with the same locations in the diffraction patterns. Hence, for $\mathcal{N}_m$, we adopt a modified U-net architecture with an additional self-attention module [48] in the center, as shown in Fig. 2. The U-net $\mathcal{N}_m$ is composed of three stages, each consisting of a 3×3 convolution, batch normalization [49], and the rectified linear unit (ReLU). The number of convolution layers is set from 16 to 256 in different stages. The 2×2 maximum pooling and bilinear interpolation are used for the pooling and unpooling, respectively.

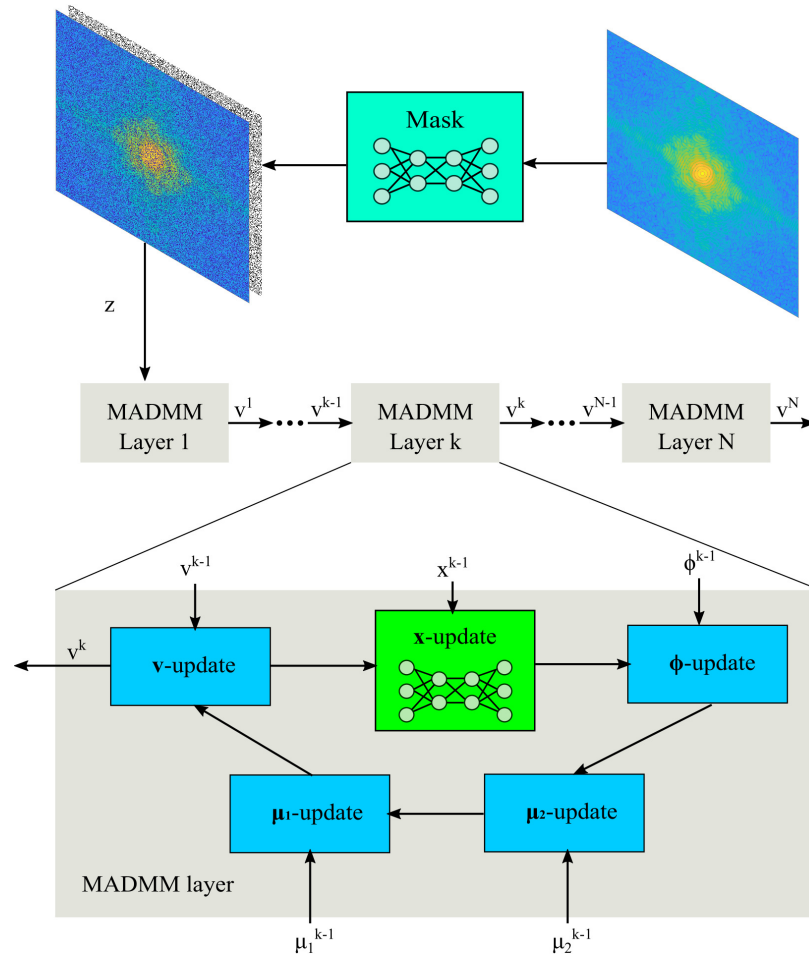


Fig. 1. Structure of the unrolled ADMM network for phase retrieval with a sensor mask (MADMM). The “Mask” block and the “x-update” block are implemented with two neural networks separately. The number of iterations $N = n_{\text{iter}}$ is pre-defined, and the final output is the value of the variable v in the last iteration.

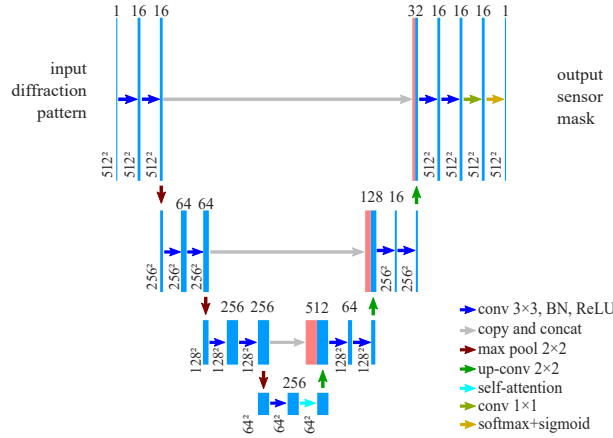


Fig. 2. Structure of the mask generation network $\mathcal{N}_m$ (the “Mask” block in Fig. 1).

Besides, inspired by the extension of the binary masks to the real-valued masks, we use a combination of the softmax and sigmoid functions to derive the final sensor mask. For the binary masks, the pixel values can only be 0 and 1, hence it is regarded as a binary classification problem using the softmax function. If the first channel is larger than 0.5, then the mask should be set as 1. While a step function can be used to convert the probability to the final binary mask, it is non-differentiable. Instead, the pixel values are allowed to be any number in $[0, 1]$, so that we can directly use the softmax function to obtain the probability, and the sigmoid function to compute the final pixel values in the mask.

The $\mathbf{x}$ -update network $\mathcal{N}_x$ is also a U-net, but with only one stage, which is the same as the setup in [50]. As shown in Eq. (18), the $\mathbf{x}$ -update step is also related to the penalty parameter ρ_2 , hence the input is designed to be the concatenation of $\mathbf{v}^{(k+1)} - \mu_2^{(k)}$ and $\rho_2 \mathbf{1}$. The number of 3×3 convolution filters is set to be 16. The same concatenation, max pooling and unpooling operators, as shown in Fig. 2, are also used in $\mathcal{N}_x$. Finally, a fully connected layer can provide the output with two channels, which are the real and imaginary parts, respectively.

Compared with the existing model-based learning algorithms for phase retrieval, such as UpR [51], GEC-SR [52] and DualPRNet [53], the main contribution is that we introduce a sensor mask into the unrolled structure, and an additional neural network is used to determine the mask to further improve the phase retrieval results.

There are three hyperparameters in Algorithm 2, namely, the penalty parameters ρ_1 , ρ_2 , and the number of iterations n_{iter} . The penalty parameters are determined automatically in the training process. We treat these parameters as learnable in the training process, and then their values are calculated by back propagation using the training data as the general network training procedure [47]. The number of iterations is determined by the trial and error strategy, and the constraints imposed by the computational resources.

The loss function $l(\mathbf{v}, \mathbf{v}_0)$ is designed to be a shift-invariant ℓ_1 norm

$$l(\mathbf{v}, \mathbf{v}_0) = \min\{\|\mathbf{v} - \mathbf{v}_0\|_1, \|\mathbf{v}_{\text{rot}180} - \mathbf{v}_0\|_1\}, \quad (27)$$

where $\mathbf{v}_{\text{rot}180}$ is the image of $\mathbf{v}$ after rotating 180° with the center point in the 2D space, and $\mathbf{v}_0$ is the ground truth latent image. The loss function is used for the final output of the unrolled MADMM network. As is summarized in Algorithm 2, the input is the measurement image $\mathbf{z}$ and the output is the reconstructed latent image $\mathbf{v}$. We can evaluate the quality of the output $\mathbf{v}$ with the ground truth image $\mathbf{v}_0$ using this loss function.

4. Experiments

4.1. Phase retrieval with a random binary sensor mask

We first demonstrate the advantage of introducing a sensor mask in ADMM for Fourier phase retrieval. In this experiment, We use the images from the MNIST dataset as the ground truth. These images are 28×28 pixels, and the imaging model is simulated by an oversampled Fourier transform, where the size of the measurement image is 128×128 . The sensor mask is a random binary mask, which is proved to be efficient when sparsity is assumed [16]. The number of blocked pixels is determined by the sampling rate, while their locations are random. The sampling rate is calculated as the number of blocked pixels over the total number of pixels times 100%. For each sampling rate, 1000 experiments with different random masks are conducted. In the phase retrieval, we use the method described in Algorithm 1. The maximum number of iterations is set as 2000, while $\mathbf{v}^{(0)}, \mathbf{x}^{(0)}, \boldsymbol{\mu}_1^{(0)}, \boldsymbol{\mu}_2^{(0)}$ are all zero vectors, and $\boldsymbol{\phi}^{(0)}$ is initialized as vector $\mathbf{1}$. The penalty parameters are set as $\rho_1 = \rho_2 = 1$, and the scale parameter is $\gamma = 0.2$. The error tolerance is $\epsilon = 5 \times 10^{-3}$.

Experimental results with different sampling rates are summarized in Table 1, while some visual results are shown the Fig. 3. In the table, “success” means that the ADMM method can provide a reasonable phase retrieval result, and terminate before the maximum number of iterations. The quality of the phase retrieval results can be evaluated by some commonly-used metrics, such as the structural similarity index (SSIM) [54], while the number of required iterations before termination shows the efficiency of this method, and the effectiveness is demonstrated by the success rate of our method. W-SSIM shows the average of SSIM over all cases if we regard the SSIM of unsuccessful cases as 0. We can find that when the sampling rate is 70% in this experiment, our method achieves the highest success rate and SSIM, leading to the highest W-SSIM as overall evaluations. For the visual results, when the sampling rate is high (over 90%), in most cases the quality of the results are not satisfactory. When the sampling rate is 90% ~ 65%, the phase retrieval results of our method have good visual quality with fewer number of iterations, if they succeed, as we show in Code 1 (Ref. [55]). When the sampling rate is 60%, no useful information is available in the phase retrieval results for most cases.

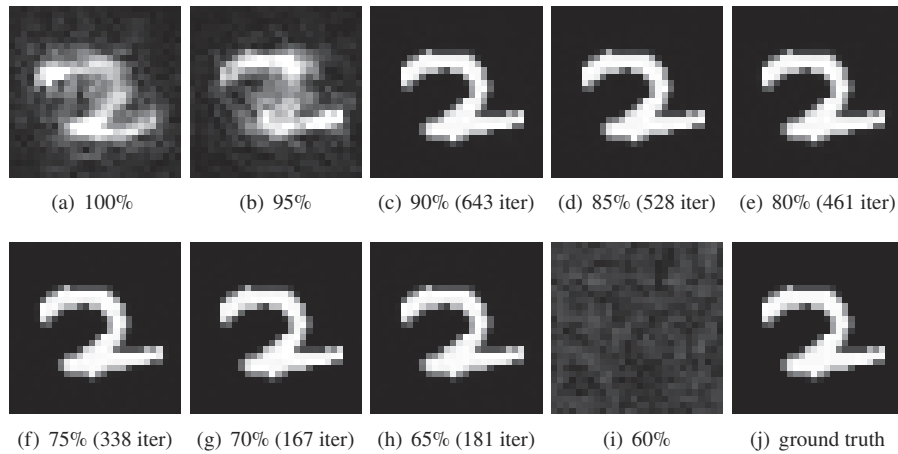


Fig. 3. Phase retrieval results for CDI with different sampling rates, where (a), (b) and (i) terminate at the maximum allowed iterations (2000) in this experiment.

Furthermore, Fig. 4 and 5 show the influences of the sampling rates on the number of required iterations and SSIM of the reconstructed images. When the sampling rate is high ($\geq 90\%$), in most cases ADMM cannot terminate within 2000 iterations, given the pre-defined error tolerance,

Table 1. The rate of success and SSIM of the ADMM results with different sampling ratios for the noise-free diffraction patterns. “Success” means that the ADMM method can provide a reasonable phase retrieval result, and terminate before the maximum number of iterations. SSIM is the average over successful cases. Weighted SSIM (W-SSIM): success rate $\times$ SSIM.

Sampling	95%	90%	85%	80%	75%	70%	65%	60%
Success(%)	16.2	53.0	74.7	82.3	89.5	91.1	71.6	24.1
SSIM	0.4762	0.6812	0.8115	0.9003	0.9538	0.9662	0.8519	0.3166
W-SSIM	0.0771	0.3610	0.6062	0.7409	0.8537	0.8802	0.6100	0.0763

but they can achieve SSIM near 0.4, with no failed case (SSIM < 0.1). A lower sampling rate (80% ~ 90%) will lead to fewer required iterations and more successful (SSIM > 0.9) cases. Furthermore, with a proper sampling rate (70% ~ 80% in this experiment), more than 80% cases can terminate within 600 iterations, and their final results have SSIM larger than 0.9. However, when the sampling rate is too low ($< 70\%$), the number of failed cases will increase, i.e., no useful information is reconstructed, and more cases can not terminate within 2000 iterations with lower sampling rates.

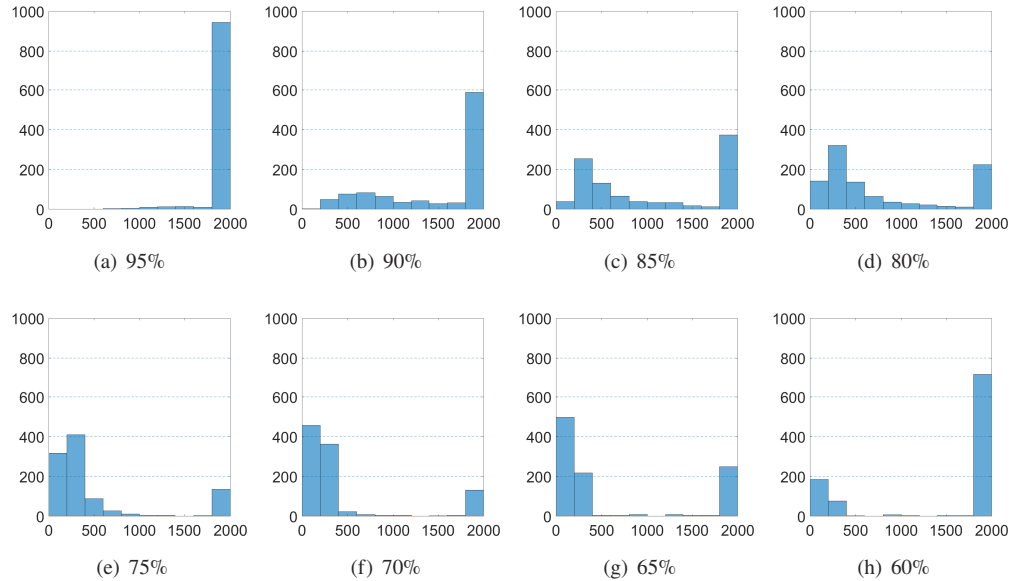


Fig. 4. Histogram of the number of iterations before termination with different sampling rates. x-axis: the number of required iterations; y-axis: the number of cases.

In summary, a sensor mask with proper sampling rate can reduce the number of iterations required by the ADMM phase retrieval method. Therefore, if the number of iterations is fixed, a proper mask can lead to higher-quality reconstructed results, compared with using the full diffraction pattern. The reason is that although the full diffraction pattern contains more information, many pixels inside are highly correlated. Removing some of them does not reduce the information content, but decreases the scale of the whole phase retrieval problem, i.e., fewer pixels in the phase map need to be retrieved.

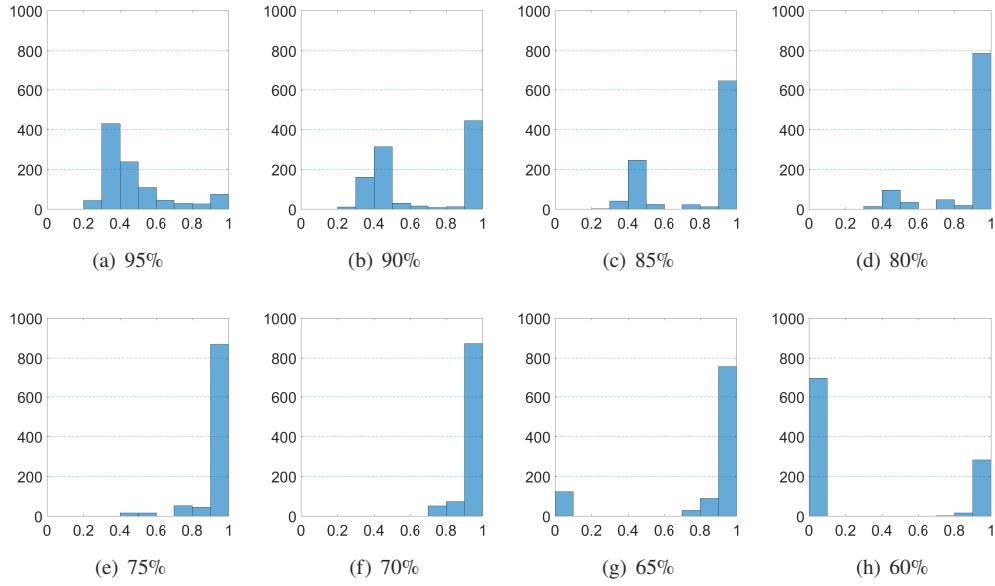


Fig. 5. Histogram of SSIM in different cases with different sampling rates. x-axis: SSIM; y-axis: the number of cases.

4.2. Phase retrieval with a learned sensor mask and unrolled ADMM network

4.2.1. Datasets and imaging model

This experiment aims at solving the phase retrieval problem in biological imaging, where the ground truth images are histopathologic scans of lymph node sections from the PatchCamelyon benchmark [56]. The ground truth images are resized to 128×128 , and zero-padding is used to guarantee uniqueness up to the trivial ambiguities described in [5] for Fourier measurements, leading to a size of 512×512 for the diffraction patterns after the phaseless Fourier transform. Since color image is rarely considered in the Fourier phase retrieval problem with reference to the diffraction optics [57], we change the RGB ground truth to grayscale images first. A test set with 1000 images is used to verify the performance of the unrolled network.

The intensity range of the ground truth images is set to $[0, 1]$. In this experiment, we consider the existence of additive noise, and the model can be described as

$$z^2 = |Fv|^2 + e, \quad (28)$$

where $e \sim N(0, \alpha^2 |Fv|^2)$ is a Gaussian random variable. In this experiment, the noise level α is set to be 3. The unrolled network described in Algorithm 2 is used for phase retrieval. The number of iterations is set to be $n_{\text{iter}} = 120$, which is slightly lower than the conventional ADMM iterations in the last section. The learning rate is set to be 1×10^{-4} in this experiment. The Adam optimizer with parameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$ is used in the training process. All the experiments in this section are conducted on the machine with 2.2 GHz Intel Xeon 4210 CPU and NVIDIA Tesla V100 (32Gb memory).

4.2.2. Experimental results

The experimental results of our unrolled network, as well as some other commonly-used phase retrieval algorithms, are presented in this section. For comparison, different kinds of phase retrieval methods are used for the same test images. The support area is known to all of these methods. HIO [9] and oversampling smoothness (OSS) [58] are two conventional alternating

methods for phase retrieval. HIO runs for 1000 iterations, while OSS runs for 2000 iterations, according to their default setups. The hyper-parameter β is set to be 0.9 in both methods. Wirtinger flow (WF) [25] is an optimization method designed for the CDP measurements. Here, we adopt it for the Fourier phase retrieval problem directly with the error tolerance equal to 1×10^{-3} . The prDnCNN is a specific realization of the prDeep [39] method. It is a plug-and-play iterative phase retrieval technique, and the trained denoising network DnCNN is used for regularization. Since HIO is already listed as another method, prDnCNN is given the same random initialization instead of the HIO result for fair comparison. The number of iterations is set to be 200, which is the same as the default setup. We also train an unrolled ADMM network without the sensor mask for comparison (named UADMM). In this network, $\mathcal{N}_m(\cdot)$ is used for the $\mathbf{x}$ -update step. All the other training details are the same as our unrolled network.

Some visual results are shown in Fig. 6, while quantitative results are summarized in Table 2. Our method outperforms all the other methods in this experiment. Conventional alternating methods, such as HIO and OSS, only consider the support as prior knowledge in the imaging domain, hence the measurement noise has a large impact on the final results. WF is designed for the CDP measurement, and has some restrictions on the system matrix, whose rows are required to be i.i.d. Gaussian, hence it does not perform well on this problem. The plug-and-play method, such as prDnCNN, is training-free, but the trained denoising network may not be suitable for all the images. Besides, the initialization of the original prDnCNN is required to be the output of other conventional methods, such as HIO. Its performance is worse with random initialization. Compared with the UADMM method, our MADMM method can give reconstructed images with more textures within the given number of iterations, due to the existence of a learned sensor mask.

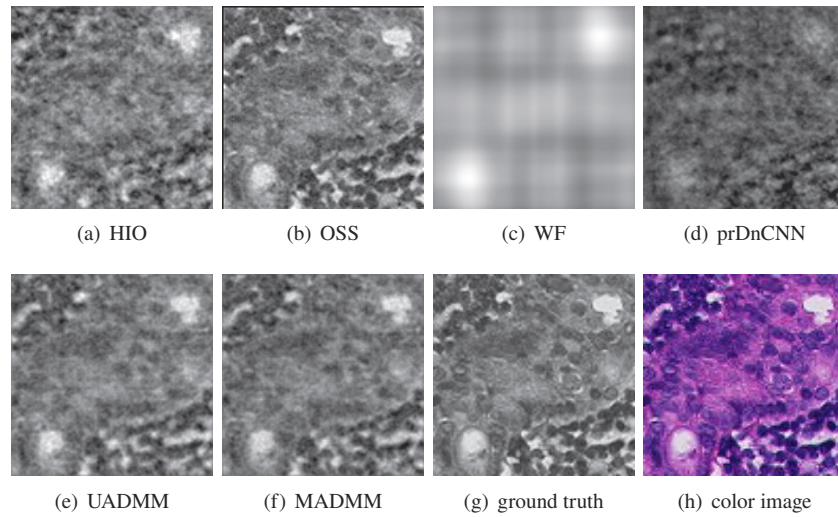


Fig. 6. Visual phase retrieval results with different kinds of methods.

Table 2. Quantitative results of different algorithms for the Fourier phase retrieval problem.

Method	HIO	OSS	WF	prDnCNN	UADMM	MADMM
PSNR(dB)	18.75	16.35	11.24	11.13	20.11	21.11
SSIM	0.482	0.594	0.201	0.210	0.668	0.876

5. Conclusion

In this paper, we propose a new phase retrieval scheme by introducing a sensor mask to the Fourier intensity pattern directly. A proper sensor mask can improve the performance of the iterative phase retrieval algorithms. Specifically, we show that ADMM can give better phase retrieval results using fewer iterations with some binary sensor masks, compared with using the original complete Fourier measurements. Since random masks can lead to instability of the ADMM iterations, we refer to the learning-based method to determine a proper sensor mask. A modified U-net is designed for this task. The mask generation network and the unrolled ADMM are jointly trained. Comparisons with other methods show that our MADMM method can achieve higher-quality phase retrieval results. This paves the way of designing new optical systems, such as CDI, with a sensor mask to improve the quality and speed of the phase retrieval process.

Funding. University Grants Committee (17200019, 17200321, 17201620); University Research Committee, University of Hong Kong (104005864).

Acknowledgement. The authors would like to thank Mr. Zhou Ge for useful discussions about this paper.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. R. P. Millane, "Phase retrieval in crystallography and optics," *J. Opt. Soc. Am. A* **7**(3), 394–411 (1990).
2. M. M. Seibert, T. Ekeberg, F. R. N. C. Maia, M. Svenda, J. Andreasson, O. Jönsson, D. Odić, B. Iwan, A. Rocker, D. Westphal, M. Hantke, D. P. DePonte, A. Barty, J. Schulz, L. Gumprecht, N. Coppola, A. Aquila, M. Liang, T. A. White, A. Martin, C. Caleman, S. Stern, C. Abergel, V. Seltzer, J.-M. Claverie, C. Bostedt, J. D. Bozek, S. Boutet, A. A. Miahnahri, M. Messerschmidt, J. Krzywinski, G. Williams, K. O. Hodgson, M. J. Bogan, C. Y. Hampton, R. G. Sierra, D. Starodub, I. Andersson, S. Bajt, M. Barthelmess, J. C. H. Spence, P. Fromme, U. Weierstall, R. Kirian, M. Hunter, R. B. Doak, S. Marchesini, S. P. Hau-Riege, M. Frank, R. L. Shoeman, L. Lomb, S. W. Epp, R. Hartmann, D. Rolles, A. Rudenko, C. Schmidt, L. Foucar, N. Kimmel, P. Holl, B. Rudek, B. Erk, A. Hömke, C. Reich, D. Pietschner, G. Weidenspointner, L. Strüder, G. Hauser, H. Gorke, J. Ullrich, I. Schlichting, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.-U. Kühnel, R. Andritschke, C.-D. Schröter, F. Krasniqi, M. Bott, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, H. Hirsemann, G. Potdevin, H. Graafsma, B. Nilsson, H. N. Chapman, and J. Hajdu, "Single mimivirus particles intercepted and imaged with an X-ray laser," *Nature* **470**(7332), 78–81 (2011).
3. T. Zeng, Y. Zhu, and E. Y. Lam, "Deep learning for digital holography: a review," *Opt. Express* **29**(24), 40572–40593 (2021).
4. Z. Ren, Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," *Optica* **5**(4), 337–344 (2018).
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 Process. Mag.* **32**(3), 87–109 (2015).
6. G. Popescu, T. Ikeda, R. R. Dasari, and M. S. Feld, "Diffraction phase microscopy for quantifying cell structure and dynamics," *Opt. Lett.* **31**(6), 775–777 (2006).
7. J. W. Goodman, *Introduction to Fourier Optics* (4th Edition) (Roberts and Company Publishers, 2017).
8. E. Y. Lam, X. Zhang, H. Vo, T.-C. Poon, and G. Indebetouw, "Three-dimensional microscopy and sectional image reconstruction using optical scanning holography," *Appl. Opt.* **48**(34), H113–H119 (2009).
9. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**(15), 2758–2769 (1982).
10. E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Appl. Comput. Harmon. Analysis* **39**(2), 277–299 (2015).
11. K. Hirata, K. Shinzawa-Itoh, N. Yano, S. Takemura, K. Kato, M. Hatanaka, K. Muramoto, T. Kawahara, T. Tsukihara, E. Yamashita, K. Tono, G. Ueno, T. Hikima, H. Murakami, Y. Inubushi, M. Yabashi, T. Ishikawa, M. Yamamoto, T. Ogura, H. Sugimoto, J.-R. Shen, S. Yoshikawa, and H. Ago, "Determination of damage-free crystal structure of an X-ray-sensitive protein using an XFEL," *Nat. Methods* **11**(7), 734–736 (2014).
12. 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).
13. L. Strüder, S. Epp, D. Rolles, R. Hartmann, P. Holl, G. Lutz, H. Soltau, R. Eckart, C. Reich, K. Heinzinger, C. Thamm, A. Rudenko, F. Krasniqi, K.-U. Kühnel, C. Bauer, C.-D. Schröter, R. Moshhammer, S. Techert, D. Miessner, M. Porro, O. Hälker, N. Meidinger, N. Kimmel, R. Andritschke, F. Schopper, G. Weidenspointner, A. Ziegler, D. Pietschner, S. Herrmann, U. Pietsch, A. Walenta, W. Leitenberger, C. Bostedt, T. Möller, D. Rupp, M. Adolph, H. Graafsma, H. Hirsemann, K. Gärtner, R. Richter, L. Foucar, R. L. Shoeman, I. Schlichting, and J. Ullrich, "Large-format, high-speed, X-ray pnCCDs combined with electron and ion imaging spectrometers in a multipurpose chamber for experiments at 4th generation light sources," *Nucl. Instrum. Methods Phys. Res., Sect. A* **614**(3), 483–496 (2010).

14. K. He, X. Chen, S. Xie, Y. Li, P. Dollar, and R. Girshick, "Masked autoencoders are scalable vision learners," arXiv preprint arXiv:2111.06377 (2021).
15. L. Jiang, A. G. Hauptmann, and G. Xiang, "Leveraging high-level and low-level features for multimedia event detection," in *Proceedings of the 20th ACM International Conference on Multimedia*, (ACM, 2012), pp. 449–458.
16. A. Guerrero, S. Pinilla, and H. Arguello, "Phase recovery guarantees from designed coded diffraction patterns in optical imaging," *IEEE Trans. on Image Process.* **29**, 5687–5697 (2020).
17. P. Kocsis, I. Shevkunov, V. Katkovnik, H. Rekola, and K. Egiazarian, "Single-shot pixel super-resolution phase imaging by wavefront separation approach," *Opt. Express* **29**(26), 43662–43678 (2021).
18. 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," *Nat. Commun.* **9**(1), 1826 (2018).
19. L. Song and E. Y. Lam, "Fast and robust phase retrieval for masked coherent diffractive imaging," *Photonics Res.* **10**(3), 758–768 (2022).
20. L. Song and E. Y. Lam, "Phase retrieval with data-driven dual alternating direction method of multipliers for coherent diffraction imaging," in *Novel Techniques in Microscopy*, (Optical Society of America, 2021), pp. NTu1C–2.
21. S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers* (Now Publishers Inc, 2011).
22. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35**, 237–246 (1972).
23. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Opt. Lett.* **3**(1), 27–29 (1978).
24. E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Comm. Pure Appl. Math.* **66**(8), 1241–1274 (2013).
25. E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Trans. Inf. Theory* **61**(4), 1985–2007 (2015).
26. Z. Xu and E. Y. Lam, "Maximum a posteriori blind image deconvolution with huber–markov random-field regularization," *Opt. Lett.* **34**(9), 1453–1455 (2009).
27. L. Song, Z. Ge, and E. Y. Lam, "Dual alternating direction method of multipliers for inverse imaging," *IEEE Trans. on Image Process.* **31**, 3295–3308 (2022).
28. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.* **2**(1), 183–202 (2009).
29. P. Zhang and E. Y. Lam, "From local to global: Efficient dual attention mechanism for single image super-resolution," *IEEE Access* **9**, 114957–114964 (2021).
30. N. Meng, H. K.-H. So, X. Sun, and E. Y. Lam, "High-dimensional dense residual convolutional neural network for light field reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.* **43**(3), 873–886 (2021).
31. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**(9), 1117–1125 (2017).
32. S. Li and G. Barbastathis, "Spectral pre-modulation of training examples enhances the spatial resolution of the phase extraction neural network (PhENN)," *Opt. Express* **26**(22), 29340–29352 (2018).
33. A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.* **121**(24), 243902 (2018).
34. T. Zeng, H. K.-H. So, and E. Y. Lam, "Redcap: Residual encoder-decoder capsule network for holographic image reconstruction," *Opt. Express* **28**(4), 4876–4887 (2020).
35. T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for Fourier ptychography microscopy," *Opt. Express* **26**(20), 26470–26484 (2018).
36. Y. F. Cheng, M. Strachan, Z. Weiss, M. Deb, D. Carone, and V. Ganapati, "Illumination pattern design with deep learning for single-shot Fourier ptychographic microscopy," *Opt. Express* **27**(2), 644–656 (2019).
37. Y. Zhang, M. A. Noack, P. Vagovic, K. Fezzaa, F. Garcia-Moreno, T. Ritschel, and P. Villanueva-Perez, "PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets," *Opt. Express* **29**(13), 19593–19604 (2021).
38. 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).
39. C. Metzler, P. Schniter, A. Veeraraghavan, and R. Baraniuk, "prDeep: robust phase retrieval with a flexible deep network," in *International Conference on Machine Learning*, (PMLR, 2018), pp. 3501–3510.
40. Y. Wang, X. Sun, and J. Fleischer, "When deep denoising meets iterative phase retrieval," in *International Conference on Machine Learning*, (PMLR, 2020), pp. 10007–10017.
41. 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*, (IEEE, 2021), pp. 10523–10531.
42. S. Pinilla, K. V. Mishra, I. Shevkunov, M. Soltanalian, V. Katkovnik, and K. Egiazarian, "Unfolding-aided bootstrapped phase retrieval in optical imaging," arXiv preprint arXiv:2203.01695 (2022).
43. N. Parikh and S. Boyd, "Proximal algorithms," *FNT in Optim.* **1**(3), 127–239 (2014).
44. S. J. Wright, R. D. Nowak, and M. A. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.* **57**(7), 2479–2493 (2009).
45. L. Li, X. Wang, and G. Wang, "Alternating direction method of multipliers for separable convex optimization of real functions in complex variables," *Math. Probl. Eng.* **2015**, 1–14 (2015).

46. S. Pinilla, K. V. Mishra, and B. Sadler, "WaveMax: FrFT-based convex phase retrieval for radar waveform design," in *2021 IEEE International Symposium on Information Theory (ISIT)*, (IEEE, 2021), pp. 2387–2392.
47. V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.* **38**(2), 18–44 (2021).
48. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing* 30, pp. 6000–6010 (Curran Associates, 2017).
49. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International Conference on Machine Learning*, (PMLR, 2015), pp. 448–456.
50. K. Monakhova, J. Yurtsever, G. Kuo, N. Antipa, K. Yanny, and L. Waller, "Learned reconstructions for practical mask-based lensless imaging," *Opt. Express* **27**(20), 28075–28090 (2019).
51. N. Naimipour, S. Khobahi, and M. Soltanalian, "Upr: A model-driven architecture for deep phase retrieval," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, (IEEE, 2020), pp. 205–209.
52. C.-J. Wang, C.-K. Wen, S.-H. Tsai, S. Jin, and G. Y. Li, "Phase retrieval using expectation consistent signal recovery algorithm based on hypernetwork," *IEEE Trans. Signal Process.* **69**, 5770–5783 (2021).
53. B. Shi and Q. Lian, "DualPRNet: Deep shrinkage dual frame network for deep unrolled phase retrieval," *IEEE Signal Process. Lett.* **29**, 1177–1181 (2022).
54. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
55. L. Song, "Matlab tools for iterative phase retrieval with a sensor mask," figshare (2022), <https://doi.org/10.6084/m9.figshare.19613886>.
56. B. S. Veeling, J. Linmans, J. Winkens, T. Cohen, and M. Welling, "Rotation equivariant CNNs for digital pathology," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (Springer, 2018), pp. 210–218.
57. E. Cha, C. Lee, M. Jang, and J. C. Ye, "DeepPhaseCut: Deep relaxation in phase for unsupervised Fourier phase retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence* p. 1 (2021).
58. 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," *J. Appl. Crystallogr.* **46**(2), 312–318 (2013).