



On-axis digital holographic microscopy: Current trends and algorithms

Andreas Erik Gejl Madsen^{a,*}, Mohammad Aryaee Panah^b, Peter Emil Larsen^b, Frank Nielsen^b, Jesper Glückstad^a

^a SDU Centre for Photonics Engineering, Mads Clausen Institute, University of Southern Denmark, DK-5230 Odense M, Denmark

^b Radiometer Medical ApS, Åkandevvej 21, 2700 Brønshøj, Denmark

ARTICLE INFO

Keywords:

Holographic microscopy
Phase retrieval
Quantitative phase imaging
Machine learning

ABSTRACT

Digital holographic microscopy (DHM) is a rapidly developing imaging technique that allows for the reconstruction of three-dimensional images from holographic data. In this review paper, we present an overview of the state-of-the-art in on-axis DHM reconstruction methods. We classify these methods into three categories: direct, iterative, and machine learning based algorithms. In the direct category, we summarize methods that directly apply the holographic data to reconstruct the sample without any iterative optimization. These methods are fast and efficient, but they often suffer from noise and low contrast. The iterative category contains methods that use an iterative optimization process to improve the quality of the reconstructed image. This can lead to higher accuracy and noise reduction, but at the cost of increased computational complexity. Within the iterative category, we also discuss methods that assume sparsity in the reconstruction and other domains, which can further improve the convergence of the reconstruction algorithm. Finally, we describe machine learning methods for DHM reconstruction. These methods use deep neural networks trained on large datasets to learn the relationship between 2D holographic data and their reconstructed fields. Machine learning based methods can achieve similar results to the iterative methods at a fraction of the time, however, they require large amounts of computational resources and training data. We provide a summary of these methods in a table (Table B.1) and a map (Fig. 1), and unify the naming convention of all methods, allowing readers to easily compare and contrast the different approaches.

1. Introduction

The light detectors used in conventional microscopes, as well as human retinas, respond solely on the intensity of incident light. As such, objects that only, or mostly, modulate the phase of light, such as transparent biological samples, cannot effectively be imaged without employing e.g. sample staining, which, aside from adding complexity to a microscopy setup, may affect the biological samples. Additionally, the phase information which is lost in conventional “intensity imaging” may contain vital information about the object, including its three-dimensional thickness and index of refraction profiles, which, in turn are governed by the physical density of the object, chemical composition, and various biological processes. In order to gain back this vital phase information, some phase-to-intensity conversion must take place prior to image capture that enables the reconstruction of the complex object wavefront, e.g. by phase contrast imaging [1–5] or other common-path phase retrieving configurations [6].

Digital holography has in recent years gained popularity as another method for this exact reconstruction problem. In theory, holography allows for full, complex, three-dimensional reconstruction of semi- and completely transparent objects with a single image capture. It has,

however, been mostly limited to research applications, due to the heavy computational load, and the sub-par quality of reconstructions. Ongoing advances in computing power, numerical optimization methods, and machine learning, have however made it viable to utilize digital holographic microscopy (DHM) in commercial settings [7,8].

1.1. Background

Holography was pioneered by D. Gabor in 1948 [9], and resulted in the well-known on-axis holography equation, describing the intensity interference pattern that occurs at a distance z from a highly transmissive object, when it is illuminated normally by a coherent and collimated light source:

$$I(x, y) = |A|^2 + |a(x, y)|^2 + A^*a(x, y) + Aa^*(x, y) \quad (1)$$

where A is the directly transmitted unscattered light resulting from the average transmission of the object, a is the scattered electric field of the object, and $*$ denotes the complex conjugate. The second term, $|a(x, y)|^2$, denoted here as a ‘self-interference’ term, arising solely from the scattered electric field, is often considered negligible due to the

* Corresponding author.

E-mail addresses: gejl@mci.sdu.dk (A.E.G. Madsen), jegl@mci.sdu.dk (J. Glückstad).

Nomenclature

λ	Wavelength
λ_k	Regularization Parameter
I	Hologram Intensity
$\mathcal{L}\{\cdot\}$	Loss Function
μ	Background Correction Term
Ω	Learned Sparsifying Transform Matrix
$\text{SFT}_k(\cdot)$	Soft-Thresholding Function
H	Hologram Amplitude
$H_k(\cdot)$	Hard-Thresholding Function
k	Wavenumber
$P_\mu(\cdot)$	Background Dependent Positive Absorption Constraining Function
$P_s(\cdot)$	Positive Absorption Constraining Function
$T(\cdot)$	Free-space Propagation Operator
W	Retrieved Hologram Phase Exponential
X	Reconstructed Sample
Z	Sparse Representation of Training Data
z	Propagation Distance
z_j	Sparse Representation of Patch from Training Data

highly transmissive nature of the object; $a \ll A$. Thus, only the squared modulus of the unscattered transmission A , and the last two terms containing the complex conjugates of A and a , respectively, are of real significance.

Since the interference pattern contains the scattered wavefront $a(x, y)$ and its complex conjugate, it carries information of both the amplitude and phase of the scattered beam and thus the complex wavefront of the object.

When reconstructing the object by illuminating¹ the hologram with a light source identical to the recording beam, a *real* image, located at a distance z behind the hologram, corresponding to the component with $a(x, y)$ will be generated. However, also due to the presence of the conjugate object wavefront $a^*(x, y)$, a *virtual* image located at a distance z in front of the hologram. These images are often referred to as *twin images*.

The twin images present the distinct disadvantage of this particular holographic method in that the two images overlap when either is brought into focus, severely degrading image contrast and signal-to-noise ratio of reconstructions. Apart from the twin image problem, on-axis holography is also limited by the assumption made to the object's transparency. If the object is not merely a weak scatterer, the second term in Eq. (1) is no longer negligible. It may, in fact, completely drown out the more important phase-information carrying components. As a consequence of this, on-axis holography is primarily suited for objects of high average transmissivity.

off-axis holography, in which a separate coherent reference beam is used and shone on the object at an angle, presents several advantages [10–12]. In this case, the intensity of the light incident on the recording medium is dependent on both the scattered object beam, and an off-axis *recording beam*. Since the recording beam consists of plane waves incident on the recording medium at an angle, the reference waves can be considered a carrier wave of spatial frequency $\alpha = \frac{\sin \theta}{\lambda}$, where θ is the off-axis incidence angle of the reference beam. The complex conjugate pair of object beams in the holography equation are then modulated by this carrier frequency. In effect, this means that at a distance z_0 in front of the exposure and deflected by the recording tilt angle θ , a virtual image will be generated. Inversely, a real image will

be generated at a distance z_0 behind the exposure, again deflected by the tilt angle θ . Since both a real and a virtual image of the object is generated, the twin image problem is still present, but, as opposed to on-axis holography, the two images have been spatially separated by the spatial carrier frequency α due to the off-axis recording beam. As such, the twin-image problem is completely avoided, given that the tilt angle is chosen appropriately.

Despite the undeniable advantages of off-axis holography, the on-axis configuration continues to find uses both in industry and academia. More recent improvements made to computational power, compressed sensing algorithms, machine learning, etc. make it a viable option for phase imaging, since the twin-image aberration can be suppressed to a certain degree. Assuming that the twin-images can be removed, on-axis holography may be favorable for two main reasons; lowered mechanical complexity and increased space-bandwidth product.

1.2. Reconstruction algorithms

In mapping the evolution of holographic reconstruction algorithms and determining the current state-of-the-art, we have identified two major categories in which they fall; direct or iterative. Direct methods (Section 2) obtain the object reconstruction in a single computational step, without the need for gradual refinement. Note that the direct methods also include machine learning (ML) methods, as inference is performed in a single step following training,² however the ML based methods are presented in a separate section (Section 4). Iterative methods (Section 3), on the other hand, perform multiple cycles of optimization in order to converge to a solution that minimizes a particular, user-defined, cost function. Equal for all methods, however, is the use of *a priori* knowledge and assumptions about the object or hologram to affect certain properties of the acquired reconstruction e.g. encourage noise reduction, and to improve the posedness of the optimization problem.

In the following sections, each of the algorithms in the two main categories and their sub-categories will be introduced and explained briefly. For increased understanding, all variables names from the original papers have been unified under a single naming convention i.e. H always denotes the captured hologram, T the propagation operator, etc. An associated illustration of the map of reconstruction algorithms can be seen in Fig. 1. In addition, reported results and parameters of the various methods are summarized in Table B.1 in Appendix B, in order to give the reader a convenient overview of the reported properties and results of the methods. Furthermore, in Figs. 2 and 3, abstracted illustrations of the experimental setups of the reviewed methods, and block diagrams of the ML architectures are presented to supplement the explanatory text.

2. Direct holographic reconstruction algorithms

2.1. Direct propagation method

As stated in Section 1, the simplest method of holographic reconstruction is performed by propagating the amplitude of the captured hologram the appropriate distance z to the object plane(s) utilizing some numerical propagator e.g. the Angular Spectrum;

$$X = T(-z) \otimes I \quad (2)$$

where X is the reconstructed object, \otimes denotes convolution, and I is the captured hologram intensity. The propagation operator T is, in this case, written as the impulse response of free-space propagation, and the propagated field is then given by the convolution of the original field and T . In practice, propagation is often carried out using Fourier transforms. Specifically, the angular spectrum method (ASM)

¹ Either in practice or numerically.

² The Deep Image Prior based method is the single found exception to this.

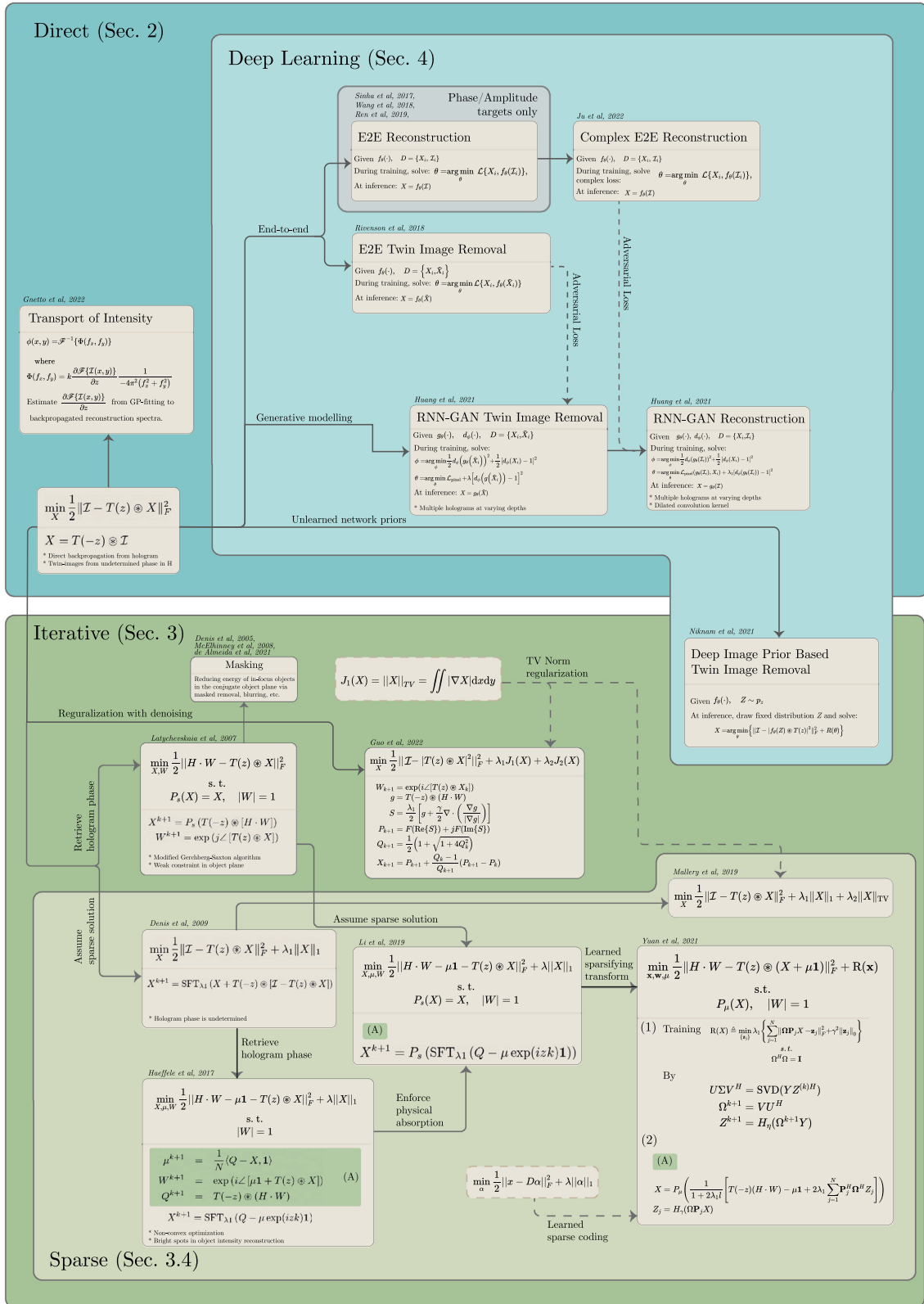


Fig. 1. Map of the methods discussed in this review, categorized into direct, iterative, and machine learning methods.

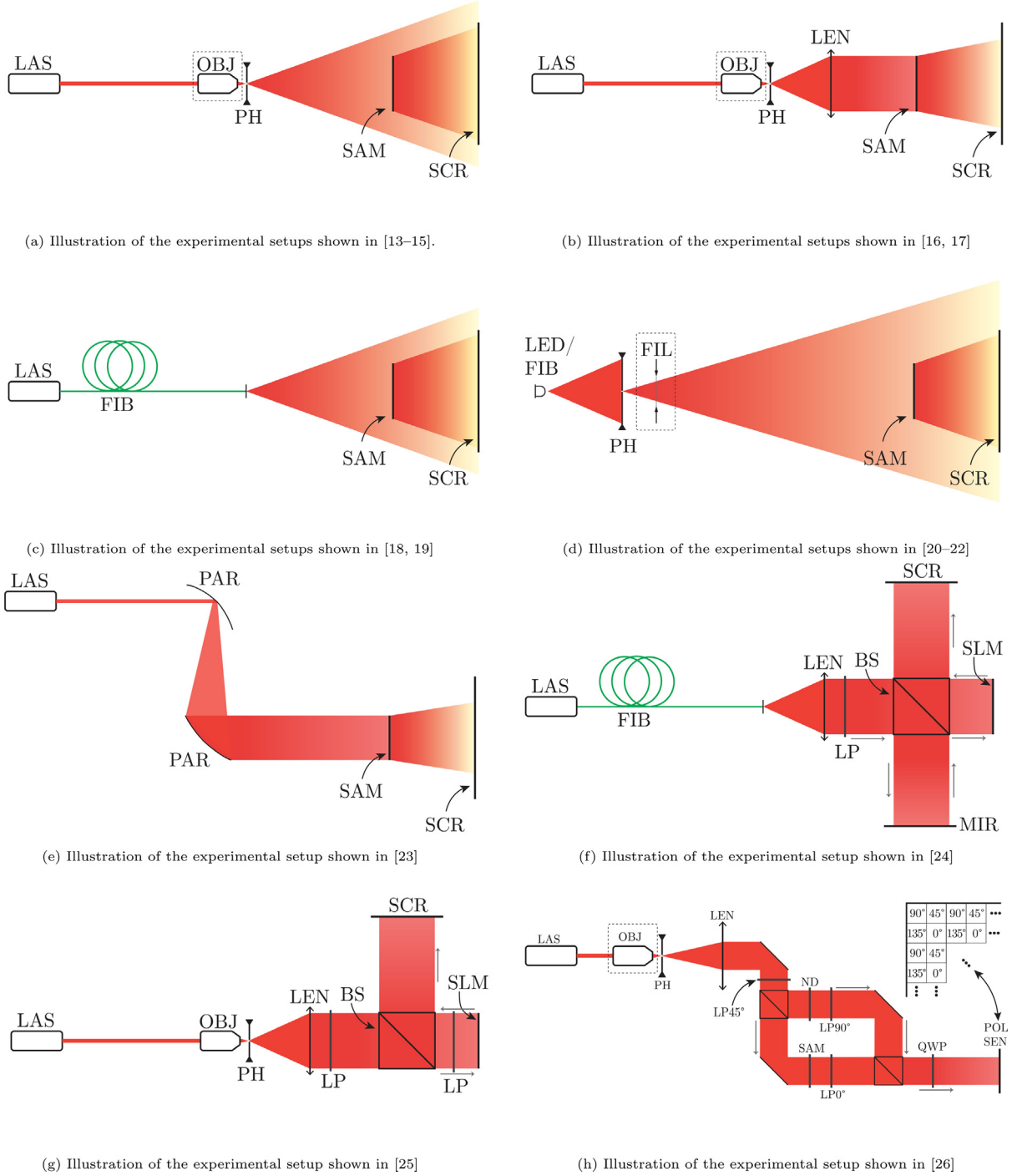


Fig. 2. Abstracted illustrations of the experimental setups utilized in the sources of this review. A dashed outline denotes that the component is not included in all setups of that category. LAS: Laser, OBJ: Microscope objective, PH: Pinhole, SAM: Sample, SCR: Screen/Camera, LENS: Lens, FIB: Fiber optic cable, LED: LED, FIL: Wavelength filter, PAR: Parabolic mirror, LP: Linear polarizer, BS: Beamsplitter, MIR: Flat mirror, SLM: Spatial light modulator, ND: Neutral density filter, QWP: Quarter wave plate, POL SEN: Polarized image sensor (see Refs. [13–26]).

propagation of a field $U(x, y, 0)$ to some axial plane at a distance z is calculated by [28]:

$$U(x, y, z) = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ U(x, y, 0) \} \right. \quad (3)$$

$$\left. \times \exp \left[j 2 \pi \frac{z}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} \right] \right\} \quad (4)$$

The method in Eq. (2) is, in fact, the closed form solution of the most simple formulation of the holographic reconstruction problem,

$$\min_X \frac{1}{2} \| I - T(z) \otimes X \|_F^2 \quad (5)$$

which minimizes the difference between the amplitude of the captured hologram and a numerical hologram estimation, calculated by propagating the obtained object field back to the hologram plane. $\| \cdot \|_F$ denotes the Frobenius matrix norm.

Although this method is both simple to implement and relatively fast in execution using the convolution theorem and Fast Fourier transforms (FFT), it is also the worst performing method in regards to

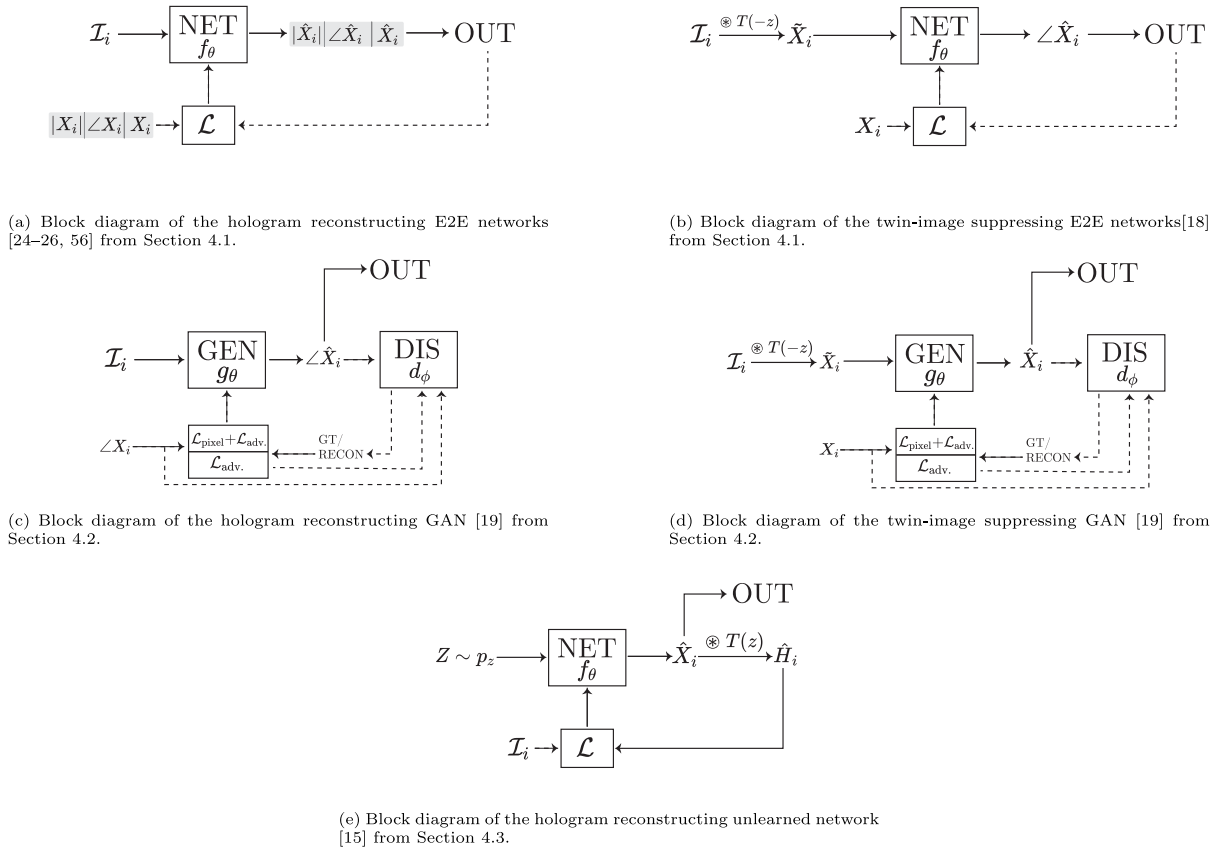


Fig. 3. Block diagrams showing the structure of the DHM neural networks proposed in the sources of this review. Solid lines indicate the data path during inference, and dashed indicate the training data path. (see Refs. [15,18,19,24–27]).

the quality of the reconstructions. Since the minimization in Eq. (5) only considers the *amplitude* of the captured hologram, disregarding its unknown phase, there is no guarantee that the obtained object reconstruction is correct, since it assumes a perfectly flat phase profile in the hologram plane, which is highly unlikely to ever be the case. More practically, since the solution in Eq. (2) is simply the propagated hologram amplitude, the resulting reconstruction will consist of both a real object plane image, and the virtual twin image in the conjugate plane, just as expected from the holography equation in Eq. (1).

2.2. Twin-image masking-based reconstruction

One rather intuitive method of improving the simple propagation reconstruction method is presented by McElhinney et al. [17], in the form of twin-image filtering by masking. The working principle of twin-image filtering is based on the idea of reducing the energy of in-focus objects in the conjugate image plane via masked removal, blurring, etc. In practice, the captured hologram is propagated numerically to the conjugate image plane in which the twin-image is in focus, and the real image is defocused. Here, a support region that covers the focused twin-image is found by way of some segmentation algorithm e.g. thresholding, edge detection, etc. Assuming that the support is well-defined, the twin-image can be removed or suppressed. McElhinney et al. [17] propose a simple negative segmentation mask containing the support region which, when multiplied with the conjugate plane field, sets any pixels within the support region to zero. Following this masking operation, the updated conjugate plane is propagated back to the hologram plane, to form a new, hopefully twin-image-suppressed hologram. With this new hologram, the authors reconstruct the real image, and shows definite suppression of the twin-image. However, since the in-focus twin-image and out-of-focus real image overlap in the conjugate image plane, any suppression filtering performed on the

twin-image will, inevitably, cause some distortion of the real image, for instance in the form of blurring.

2.3. Numerical transport of intensity reconstruction

An innovative direct method based on the simple propagation method in Eq. (2) was published in 2022 by Gnetto et al. [14]. The authors propose a reconstruction method employing a Transport-of-Intensity Equation (TIE) [29–31] modality (Gaussian Process Regression TIE; GP-TIE [32]) to retrieve an accurate estimate of the phase in the object plane with reduced influence of the twin-image artifact.

First, the background of TIE is explained, followed by the modality proposed by Gnetto et al. [14]. It is important to note that while TIE is usually a multi-height phase recovery method, requiring multiple intensity captures, the method proposed by Gnetto et al. [14] maintains the use of only a single hologram capture.

TIE is an established method of phase retrieval and is based on the inherent relationship that exists between the intensity variations of an optical field along its axis of propagation, and the phase of the field in perpendicular planes. This makes it possible to retrieve the phase of a “plane-of-interest”, by capturing the intensity of the propagating optical field in multiple surrounding planes along the optical axis. More specifically, through an involved derivation [31,33], the following second-order elliptic partial differential equation can be found:

$$-k \frac{\partial^2 I(x, y)}{\partial z^2} = \nabla \cdot [I(x, y) \nabla \phi(x, y)] \quad (6)$$

where z is the optical axis coordinate, k is the wave number of the optical field, ∇ is the 2D gradient operator, and $I(x, y)$ and $\phi(x, y)$ the intensity and phase of the optical field in the plane-of-interest, respectively. By solving Eq. (6), it is possible to obtain an expression of the phase profile $\phi(x, y)$ in terms of the axial spatial derivative of

the optical field. To utilize the TIE approach in practice, the axial derivative is substituted for a numerical estimate e.g. finite differences,³ calculated from multiple intensity measurements captured at different propagation distances (defocused images) surrounding the plane-of-interest:

$$\frac{\partial I(x, y)}{\partial z} \approx \frac{I_{+\Delta z}(x, y) - I_{-\Delta z}(x, y)}{2\Delta z} \quad (7)$$

where Δz is the defocus distance. TIE has been used extensively [18,34–37] in holographic microscopy as a method for retrieving the hologram phase. In these cases, an axial scan of the distance between the sample and the camera sensor is carried out, while capturing intensity images at each propagation distance. With this axial “stack”, the associated phase of a chosen hologram plane can be retrieved by solving Eq. (6), and a complex hologram can be constructed from the square-root of the intensity image captured at the same distance. Since the hologram is now complex, the root cause of the twin-image in Eq. (1) is avoided. As such, the reconstruction of this complex hologram by the simple propagation method should be completely free of twin-images, depending on the accuracy of the axial derivative estimate.

Although this multi-hologram phase recovery method shows excellent results (so good that the results may be used as ground truth training data), a significant advantage of DHM is the ability to obtain 3D reconstructions from a *single* hologram image, without the need for axial scanning. The method presented by Gnetto et al. [14] combines the single hologram reconstruction capabilities of holography with TIE using Gaussian process regression. The presented holographic GP-TIE method creates a stack of axial intensity images by reconstructing a single hologram to multiple, equally spaced, object planes in and around a plane-of-interest; $I_{-n\Delta z}, \dots, I_0, \dots, I_{+n\Delta z}$. From the stack of reconstructions, the authors fit an estimate of the axial derivative in the plane-of-interest $\frac{\partial I(x, y)}{\partial z}$ using Gaussian process regression. More specifically, they actually perform the derivative fitting in the spatial frequency domain, since the evolution of the intensity spectrum while propagating can, locally,⁴ be considered nearly sinusoidal, making fitting more robust.

To reiterate, the authors make use of Gaussian process regression to fit an estimate of the axial derivative of *spectrum* of the intensity in the plane-of-interest $\frac{\partial \mathcal{F}\{I(x, y)\}}{\partial z}$ to the stack of Fourier transformed intensity images, reconstructed from the *single* captured hologram;

$$\mathcal{F}\{I_{-n\Delta z}\}, \dots, \mathcal{F}\{I_0\}, \dots, \mathcal{F}\{I_{+n\Delta z}\}.$$

From the derivative of the intensity spectrum, the phase in the plane-of-interest can, under the assumption of pure phase objects, be retrieved from Eq. (6) via Laplacian inversion as

$$\phi(x, y) = \mathcal{F}^{-1}\{\Phi(f_x, f_y)\} \quad (8)$$

$$\text{where} \quad (9)$$

$$\Phi(f_x, f_y) = k \frac{\frac{\partial \mathcal{F}\{I(x, y)\}}{\partial z}}{-4\pi^2(f_x^2 + f_y^2)} \quad (10)$$

Since the reconstruction stack is created by propagating the hologram amplitude using simply the ASM, each reconstruction contains a prominent twin-image artifact. However, due to the noise suppression ability of the Gaussian process regression algorithm, and the well-behaved evolution of the spectra of propagating optical fields, the spurious twin-image in the plane-of-interest is effectively ignored and filtered out in the fit. The results shown by Gnetto et al. [14] are complex reconstructions of carboxylate beads 10 μm in diameter, diluted in distilled water. The reconstructions display partial removal of the twin-image artifacts, and even though the computational time

complexity of the algorithm is $\mathcal{O}\{N^3\}$, it lends itself well to parallel and GPU computing, which allows for considerable calculation speed. The method relies on the assumption that the objects are purely phase modulating, which limits its applicability. However, as is pointed out by Jingshan et al. [32], two Laplacian inversions can be employed to alleviate this assumption.

3. Iterative methods

3.1. Hologram phase retrieval

As mentioned in Sections 1 and 2.1, the twin-image problem present in holography stems from the fact that the hologram capturing sensor only encodes intensity from incident light, leading to the unwanted terms in the holography equation in Eq. (1). By retrieving the unknown phase in the hologram plane algorithmically, and combining it with the captured hologram amplitude to form a complex hologram, the resulting object reconstruction should consist only of the true image of the object. Many of the methods in this section base their implementations on well-known phase retrieval approaches e.g. Gerchberg–Saxton(GS) algorithm [38] or Fienup algorithms [39], but with application specific constraints.

Latychevskaia and Fink [13] present an iterative method that can be stated as follows,

$$\min_{X, W} \frac{1}{2} \|H \cdot W - T(z) \otimes X\|_F^2 \quad (11)$$

$$\text{s. t.} \quad (12)$$

$$P_s(X) = X, \quad |W| = 1 \quad (13)$$

where \cdot is the element-wise product of the hologram amplitude H and the to-be retrieved phase W , subject to $|W| = 1$ such that it remains entirely phase modulating. In addition, a positive absorption constraint (see Appendix B) is imposed on the object reconstruction X , in the form of the function $P_s(\cdot)$, expressed by:

$$P_s(x) = \begin{cases} x & \text{for } \alpha \geq 0 \\ 1 & \text{otherwise} \end{cases} \quad (14)$$

where $\alpha = -\log(|x|)$ defines absorption., This ensures that no additional energy is introduced following the scattering process. Thus, any detected surplus in energy is the result of either interference between the twin-image and the reference wave or originate from another source of noise. Given these constraints, the method is implemented according to the pseudo-code in Algorithm 1. The algorithm uses the current estimate for the reconstructed object plane X^k in order to update the hologram phase W^{k+1} . This updated phase is multiplied by the captured normalized hologram amplitude H in order to create a complex hologram field, which, again, is propagated back to the object plane where it is imposed with the absorption constraint $P_s(\cdot)$, thus creating a new estimate for the object plane field. Importantly, it can be shown that the implicit error *must* decrease or become stagnant with every iteration [38,40], guaranteeing convergence. However, due to the non-convex phase constraint $|W| = 1$, the algorithm cannot guarantee convergence to a global minimum. As such, it may be prone to converge to local minima [41].

Algorithm 1 Phase Retrieval Reconstruction

```

1:  $n_i$ : Number of iterations
2:  $\angle$ : Argument of complex number
3:  $H \leftarrow \sqrt{I}$ 
4:  $W \leftarrow 0$ 
5: for  $k$  of  $n_i$  do
6:    $X \leftarrow P_s(T(-z) \otimes [H \cdot W])$ 
7:    $W \leftarrow \exp(j\angle[T(z) \otimes X])$ 
8: end for
9: return  $X$ 

```

³ In the simplest case. More sophisticated approaches to this estimate are presented in [31, Sec. 5]

⁴ By observing the intensity of the same spectral frequency (f_x, f_y) through the axial scan.

The effectiveness of the algorithm is demonstrated by Latychevskaia and Fink [13] using both simulated and experimental holograms. In the simulated case, the algorithm reconstructions show good likeness to the ground truth on simulations of a strand of hair, achieving both effectively twin-image-free reconstructions, along with correct absolute phase values. The authors report Mean Square Errors (MSE) for both the absorption and phase of the reconstructions in the order of 10^{-7} . The experimental reconstructions of both a tungsten tip and micro latex beads also show a reduction of the twin-image artifacts. However, as is expected, the reconstructions are more noisy. In both cases, however, it is clear to see the positive absorption constraint at work by the reduction of erroneous negative absorption. Nonetheless, the constraints imposed on the object and hologram in Algorithm 1 are comparatively weak, which could explain the 500–1000 iterations necessary to reach the presented convergence. With stronger constraints and added regularization, the convergence rate can be increased, as will be seen in methods published following Latychevskaia and Fink [13]. Furthermore, from our tests of the algorithm, it appears to be much more effective on targets which exhibit significant absorption. As the targets become more transmissive, to the point of becoming phase-only modulating, the algorithm struggles to achieve convergence to the ground truth.

3.2. Iterative masking-based reconstruction

There have been numerous attempts at eliminating the twin-image artifacts more explicitly, most notably via filtering in the conjugate image plane, as was demonstrated by McElhinney et al. [17] and described in Section 2.2.⁵ Denis et al. [16] and de Almeida et al. [20] present methods based on the same principle, but adapted to iterative algorithms. Both methods follow similar approaches; the captured hologram amplitude is propagated to the *real* image plane. By some segmentation method e.g. thresholding or edge detection, a mask of the in-focus real object is constructed. Then, the current hologram estimate is propagated to the *virtual* image plane, containing the now in-focus twin-image. To reduce the energy of the twin-image, it is either multiplied by a pre-calculated mask [16], or a blurring kernel is applied within the mask [20]. Thus, when this filtered virtual image plane is propagated back to the hologram plane, the real image plane reconstruction of this updated complex hologram should contain a suppressed twin-image artifact. This process is repeated for the desired number of iterations. As the twin-image artifact is suppressed, the mask segmentation becomes more accurate, which, in turn, improves the accuracy of the virtual image filtering. The pseudo-code for the two algorithms [16,20] is laid out in Algorithm 2.

Algorithm 2 Iterative Masking-Based Reconstruction

```

1:  $n_i$ : Number of iterations
2:  $\xi(\cdot)$ : Mask segmentation function
3:  $\Theta(\cdot, \cdot)$ : Filter function using mask
4:  $\tilde{H} \leftarrow I$ : Updated hologram
5: for  $k$  of  $n_i$  do
6:    $X_{\text{real}} \leftarrow T(-z) \otimes \tilde{H}$ 
7:    $M \leftarrow \xi(X_{\text{real}})$ 
8:    $X_{\text{virtual}} \leftarrow T(z) \otimes \tilde{H}$ 
9:    $X_{\text{virtual-filtered}} \leftarrow \Theta(X_{\text{virtual}}, M)$ 
10:   $\tilde{H} \leftarrow T(-z) \otimes X_{\text{virtual-filtered}}$ 
11: end for
12:  $X_{\text{real}} \leftarrow T(-z) \otimes \tilde{H}$ 
13: return  $X_{\text{real}}$ 

```

While the results presented in [16,20] show suppression of the twin-image to a greater extent than the direct masking method shown in

Section 2.2, the algorithms come at the same cost of losing information in the real image. In [20], this is apparent in the experimental reconstruction of some biological sample, in which it is clear that the real image of the samples has become slightly blurred as a result of the masking operation. Furthermore, the effectiveness of these methods rely heavily on the successful segmentation of the virtual objects. This fact imposes limitations on how well the methods can generalize to all types of samples.

3.3. Denoising prior phase retrieval

Guo et al. [22] presents an iterative reconstruction method, set on reducing the influence of twin-images in the reconstruction problem of equation Eq. (5) by way of dual regularization. The authors propose the use of both a Total Variation (TV) norm regularizer in combination with regularization by denoising [42]. The authors present solutions to the following minimization problem;

$$\min_X \frac{1}{2} \|T(z) \otimes X\|^2 - I\|_F^2 + \lambda_1 J_1(X) + \lambda_2 J_2(T) \quad (15)$$

where I is the captured hologram intensity, $J_1(X)$ and J_2 are the TV norm and denoising regularizer, respectively, and λ_1 and λ_2 are corresponding parameters affecting the strength of the given regularization. The first regularizing term, the TV norm, expressed as $\|X\|_{\text{TV}} = \iint |\nabla X| dx dy$ functions as an effective edge-preserving noise remover [43], smoothing signals in flat regions while maintaining important high-frequency information. The second regularizer, $J_2(X)$, is presented as a denoising prior, a placeholder for interchangeable and sophisticated denoising algorithms that process the real and imaginary part of the current reconstruction estimate separately, in an attempt to filter out the erroneous twin-image aberrations. The authors test a collection of popular denoisers, and deem the Trainable Nonlinear Reaction Diffusion (TNRD) [44] filter to be the best performing for the task at hand. The pseudo-code for the algorithm is laid out in Algorithm 3. Guo et al. [22] validate their algorithm experimentally on both amplitude (negative and positive) and phase samples, including both synthetic targets as well as biological samples.

Algorithm 3 Denoising Prior Phase Retrieval

```

1:  $n_i$ : Number of iterations
2:  $X_0 \leftarrow T(-z) \otimes I$ 
3:  $H \leftarrow \sqrt{I}$ 
4:  $\gamma$ : Gradient descent step size
5: for  $k$  of  $n_i$  do
6:    $W \leftarrow \exp(i\angle[T(z) \otimes X_{k-1}])$  ▷ Retrieved phase in the
hologram plane
7:    $g \leftarrow T(-z) \otimes (H \cdot W)$  ▷ Temporary
reconstruction estimate
8:    $S \leftarrow \frac{\lambda_1}{2} \left[ g + \frac{\gamma}{2} \nabla \cdot \left( \frac{\nabla g}{|\nabla g|} \right) \right]$  ▷ Gradient descent step for
TV regularization
9:    $P_k \leftarrow F(\text{Re}\{S\}) + jF(\text{Im}\{S\})$  ▷ Denoising by algorithm F
10:   $Q_k \leftarrow \frac{1}{2} \left( 1 + \sqrt{1 + 4Q_{k-1}^2} \right)$ 
11:   $X_k \leftarrow P_k + \frac{Q_{k-1}-1}{Q_k} (P_k - P_{k-1})$ 
12: end for
13: return  $X$ 

```

3.4. Sparsity constrained methods

The reconstruction of complex objects from intensity-only holograms is, inherently, a severely under-determined problem. The methods in Section 3 attempt to mitigate this problem by the inclusion of certain assumptions and constraints on the reconstructions such as positive absorption and de-noising. Another constraint is found in the world of Compressed Sensing (CS); sparsity. Sparsity encouraging methods for holographic reconstruction have become a fruitful research

⁵ Added method in direct section

area in and of themselves, and in this section, the evolution of their development and how they relate to the methods presented previously will be explored.

Sparse holographic reconstruction relies on the assumption that the object plane be sparse, ie. the objects only occupy a fraction of the pixels in the reconstruction, while the majority simply constitutes background illumination. While this assumption is not far-fetched, it does limit the density of samples that can be reconstructed successfully with twin-images removed, depending on the sophistication of the sparsity promoting model. A properly sparsified object reconstruction will have its twin-image removed and replaced by the background illumination, while the minimized number of active pixels will be designated for objects of interest.

3.4.1. Sparse amplitude-hologram reconstruction

Taking again the starting point of the unconstrained holographic reconstruction problem of Eq. (5), Denis et al. [45] propose an object sparsifying method that solves the optimization problem

$$\min_X \frac{1}{2} \|I - T(z) \otimes X\|_F^2 + \lambda_1 \|X\|_1 \quad (16)$$

s.t.

$$\forall i, X_i \geq 0, X_i = |X_i|$$

where λ_1 is a regularization parameter and $\|\cdot\|_1$ is the ℓ_1 norm. As can be seen by the absolute positivity constraint, the problem in question exclusively regards amplitude-only objects. Such minimization problems have been shown to be solvable [46] by iteratively employing the proximal operator of the ℓ_1 norm considering only real and non-negative inputs, commonly named the Soft-Thresholding (SFT) operation, which is given by

$$\begin{aligned} \text{SFT}_{\lambda_1}(\hat{\xi}) &= \text{prox}_{\lambda_1 \|\cdot\|_1}(\hat{\xi}) = \arg \min_{\xi \geq 0} \xi + \frac{1}{2\lambda_1} (\xi - \hat{\xi})^2 \\ &= \begin{cases} \hat{\xi} - \lambda_1 & \text{if } \hat{\xi} \geq \lambda_1 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (17)$$

The proximal operator and thus the SFT in Eq. (17) serves to minimize the ℓ_1 norm of an input, but is penalized for straying too far from the value of the original input. The iterative solution to problems on the form in Eq. (16), as stated by Daubechies et al. [46, p. 1427], is summarized in Algorithm 4. As can be seen, the algorithm consists of a single update step. Here, the estimate for the reconstruction X is updated with the soft-thresholded Landweber iterate [47], which is given by the sum of the current estimate and its backpropagated residual error $H - T(z) \otimes X$.

Algorithm 4 Amplitude-only Sparse Reconstruction

```

1:  $n_i$ : Number of iterations
2:  $X \leftarrow X_0$  ▷ The initial value of  $X$  not stated in original paper
3: for  $k$  of  $n_i$  do
4:    $X \leftarrow \text{SFT}_{\lambda_1}(X + T(-z) \otimes [I - T(z) \otimes X])$ 
5: end for
6: return  $X$ 

```

While the optimization problem in Eq. (16) is convex, and the ℓ_1 norm regularizer does encourage sparsity in the reconstructions, the amplitude-only hologram used by Denis et al. [45] can only go so far as to reducing twin-images [48]. Due to the missing phase information in the hologram, this method is applicable only for both sufficiently small and sparsely distributed objects in the object plane, such that “cross-interference” between their holograms can be neglected. In addition, for this method to work optimally, samples should be significantly absorptive.

Promoting sparsity in object reconstructions via regularization is not limited to the use of the ℓ_1 norm. Other non-phase-retrieval algorithms with alternative regularization schemes have been proposed, including

using the Total Variation (TV) norm [43,49,50]. The TV norm promotes small gradients, and therefore smooth solutions. However, TV regularization is substantially more computationally complex than ℓ_1 , since its proximal operator does not have a closed form solution.

3.4.2. Sparse phase retrieval reconstruction

The missing hologram phase of the algorithm in Algorithm 4 inhibits the convergence of the method. To improve upon the method, Haeffele et al. [48] explore a phase retrieving iterative reconstruction method, also based on the sparse assumption. By considering the missing phase of the hologram, Haeffele et al. [48] attempt to solve a modified Eq. (16), given by

$$\min_{X, W, \mu} \frac{1}{2} \|H \cdot W - \mu \mathbf{1} - T(z) \otimes X\|_F^2 + \lambda_1 \|X\|_1 \quad (18)$$

s.t.

$$|W| = 1$$

Where, as in [13], W is the to-be retrieved hologram phase, subject to $|W| = 1$, such that it is phase-only modulating, and $\mu \in \mathbb{C}$ is an adaptive background correcting term. The authors present closed-form solutions of all free variables X , W , and μ in Eq. (18), given by

$$\arg \min_{\mu} (18) = \frac{1}{N} \langle T(-z) \otimes H \cdot W - X, \mathbf{1} \rangle \quad (19)$$

$$\arg \min_W (18) = \exp(i \angle [\mu \mathbf{1} + T(z) \otimes X]) \quad (20)$$

$$\arg \min_X (18) = \text{SFT}_{\lambda_1}(T(-z) \otimes (H \cdot W) - \mu \exp(izk) \mathbf{1}) \quad (21)$$

Where N is the total number of pixels in the hologram, $\langle \cdot, \cdot \rangle$ is the Frobenius inner product, and SFT_{λ_1} is the now complex soft-threshold operator:

$$\text{SFT}_{\lambda_1}(\hat{\xi}) = \begin{cases} (\hat{\xi} - \lambda) \frac{\hat{\xi}}{|\hat{\xi}|} & \text{if } |\hat{\xi}| \geq \lambda_1 \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

These individual problems are solved by iteratively alternating updates for each until convergence is reached, as is laid out in Algorithm 5. The optimal regularization parameter λ_1 can either be found manually for each hologram, searched for by running multiple trials, or estimated from the found value for μ [23]. This phase retrieval algorithm based on object plane sparsity shows greater twin-image reduction and better cross-interference stability despite the fact that the optimization problem has become non-convex ($|W| = 1$), due to the increased solution space by including the missing hologram phase.

Algorithm 5 Sparse Phase Retrieval Reconstruction

```

1:  $n_i$ : Number of iterations
2:  $N$ : Number of pixels in hologram
3:  $\lambda_1$ : Regularization parameter
4:  $X \leftarrow \mathbf{0}$ 
5:  $W \leftarrow \mathbf{1}$ 
6:  $\mu \leftarrow 0$ 
7:  $Q \leftarrow T(-z) \otimes I$ 
8:  $H \leftarrow \sqrt{I}$ 
9: for  $k$  of  $n_i$  do
10:    $\mu \leftarrow \frac{1}{N} \langle Q - X, \mathbf{1} \rangle$ 
11:    $W \leftarrow \exp(i \angle [\mu \mathbf{1} + T(z) \otimes X])$ 
12:    $Q \leftarrow T(-z) \otimes (H \cdot W)$ 
13:    $X \leftarrow \text{SFT}_{\lambda_1}(Q - \mu \exp(izk) \mathbf{1})$ 
14: end for
15: return  $X$ 

```

3.4.3. Positive absorption constrained sparse phase retrieval

Li et al. [23] further constrains the sparsity-based optimization problem in Eq. (18), by imposing positive absorption, as was introduced

to holography by Latychevskaia and Fink [13]. As such, the authors identify and attempt to solve the problem of bright-spot imperfections in object reconstructions, which are the result of twin-image interference. Therefore, the updated optimization problem stated in [23] is given by

$$\min_{X, W, \mu} \frac{1}{2} \|H \cdot W - \mu \mathbf{1} - T(z) \odot X\|_F^2 + \lambda_1 \|X\|_1 \quad (23)$$

s.t.

$$|W| = 1, \quad P_s(X) = X$$

where $P_s(\cdot)$ is the positive absorption constraint function [13] in Eq. (14). The algorithm with which Eq. (23) is solved differs only from Algorithm 5 in the update step for X ,

$$\arg \min_X (23) = P_s \left(\text{SFT}_{\lambda_1} (T(-z) \odot (H \cdot W) - \mu \mathbf{1}) \right) \quad (24)$$

resulting in the pseudo-code in Algorithm 6. With the added positive absorption constraint, the algorithm presented by Li et al. [23] reaches both better and faster convergence than the algorithm by Haeffele et al. [48], with significant suppression of both twin-images and the aforementioned bright-spots.

Algorithm 6 Absorption Constrained Sparse Phase Retrieval Reconstruction

```

1:  $n_i$ : Number of iterations
2:  $N$ : Number of pixels in hologram
3:  $\lambda_1$ : Regularization parameter
4:  $X \leftarrow \mathbf{0}$ 
5:  $W \leftarrow \mathbf{1}$ 
6:  $\mu \leftarrow 0$ 
7:  $Q \leftarrow T(-z) \odot I$ 
8:  $H \leftarrow \sqrt{I}$ 
9: for  $k$  of  $n_i$  do
10:    $\mu \leftarrow \frac{1}{N} \langle Q - X, \mathbf{1} \rangle$ 
11:    $W \leftarrow \exp(i \angle [\mu \mathbf{1} + T(z) \odot X])$ 
12:    $Q \leftarrow T(-z) \odot (H \cdot W)$ 
13:    $X \leftarrow P_s \left( \text{SFT}_{\lambda_1} (Q - \mu \exp(izk) \mathbf{1}) \right)$ 
14: end for
15: return  $X$ 

```

3.4.4. Learned sparsifying transform reconstruction

The methods in Sections 3.4.1–3.4.3 reduce spurious noise and the twin-image in reconstructions under the assumption that the sample is sparse i.e. the object plane is mainly constant (zero-valued or a calculated background) with few active pixels comprising the objects in the sample. However, as this assumption is challenged when examining more complex samples such as biological tissue, the sparsity constraints yield less and less impressive results. In addition, the thresholding operation associated with the ℓ_1 norm regularizer may lead to unrealistic staircase edges in reconstructions.

To facilitate the reconstruction of non-sparse samples with good accuracy, Yuan et al. [21] propose a learned sparsifying transform [51, 52] (LST) based reconstruction algorithm. The learned sparsifying transform algorithm builds on the assumption that a non-sparse signal may be represented sparsely in some unknown domain, and that the transform operation necessary to convert to and from that domain can be learned from training data, as opposed to e.g. the known discrete cosine transform (DCT) used in JPEG compression.

More precisely, Ravishankar and Bresler [51] state that a square sparsifying transform matrix $\Omega \in \mathbb{C}^{l \times l}$ can be learned from N_p image patches of l pixels extracted from the training images, such that the following objective function is minimized,

$$\min_{\Omega, Z} \|\Omega Y - Z\|_F^2 + \eta^2 \|Z\|_0 \quad (25)$$

s.t.

$$\Omega^H \Omega = \mathbf{I}$$

where \cdot^H denotes the Hermetian transpose, η is a real positive coefficient controlling the degree of sparsity in the Z -domain, $Y \in \mathbb{C}^{l \times N_p}$ is the training data, and $Z \in \mathbb{C}^{l \times N_p}$ is the sparse representation of Y . Both Y and Z consists of N_p columns of l pixels, where each column corresponds to a patch extracted from the training data. Each column in Y can then be transformed to its sparse representation in Z by multiplication with Ω following learning. The learning of this transform matrix Ω comprises of two steps; a transform update step, and a sparse coding step. With Z fixed, the transform update step refines the learned transform by the full singular value decomposition (SVD) of $Y Z^H$:

$$\Omega^{k+1} = V U^H \quad (26)$$

where $U \Sigma V^H = \text{SVD}(Y Z^H)$. With the updated transform matrix, Ω^{k+1} , the sparse coding step simply consists of hard-thresholding the current sparse representation Z , due to the ℓ_0 norm minimization in Eq. (25):

$$Z^{k+1} = H_\eta(\Omega^{k+1} Y) \quad (27)$$

where H_η is the hard-thresholding operation, setting elements with magnitudes smaller than η to zero. Thus, as training progresses, the transform matrix Ω is tuned to form the bridge between the training data Y and its representation in the Z -domain. In addition, the thresholding of Z enforces sparsity in this domain, and prevents Ω from reaching its trivial solutions $\Omega = \mathbf{0}$ and $\Omega = \mathbf{I}$. Now, the LST algorithm mainly finds its uses in image processing and compression applications, as the tailoring of a sparsifying transform to a particular use case may yield better and more efficient results than the more general transforms such as wavelet, DCT and discrete Fourier transform (DFT).

Yuan et al. [21] describe a method of incorporating the notion of a learned sparsifying transform into the reconstruction problem of DHM. The holographic problem is stated, again, as a minimization problem,

$$\min_{X, W, \mu} \frac{1}{2} \|H \cdot W - T(z) \odot (X + \mu \mathbf{1})\|_F^2 + R(X) \quad (28)$$

s.t.

$$P_\mu(X), \quad |W| = 1$$

where $P_\mu(\cdot)$ enforces positive absorption, while also considering the background μ , and where the regularizer R is defined using the pre-learned transform Ω ,

$$R(X) \triangleq \min_{\{z_j\}} \lambda_1 \left\{ \sum_{j=1}^N \|\Omega \mathbf{P}_j X - \mathbf{z}_j\|_F^2 + \gamma^2 \|\mathbf{z}_j\|_0 \right\} \quad (29)$$

s.t.

$$\Omega^H \Omega = \mathbf{I}$$

Here, $\mathbf{P}_j \in \mathbb{R}^{l \times N_p}$ is a patch extraction operator, that extracts the j th patch of l pixels from the reconstructed field X , \mathbf{z}_j is the sparse representation of the j th patch of X , λ_1 is a real positive coefficient controlling the influence of the regularizer on the reconstruction, and γ is a real positive coefficient specifying the degree of sparsity of \mathbf{z} .

Since Ω has been pre-learned, and is not updated in the reconstruction algorithm, the regularizer attempts to converge on a sparse representation Z of the reconstruction X , following transformation by Ω . The better X can be represented in this sparse domain, the better the minimization of the regularizer in Eq. (29).

This means that if any part of the reconstructed image X cannot be represented in the sparse domain by the learned transformation Ω , which has been conditioned to transform proper reconstructions,⁶ these parts will cause an increased error, and will thus be suppressed by the regularizer. Due to the training of Ω on twin-image-free images, the regularizer is minimized if the reconstruction X also approaches a twin-image-free form where the Ω transformation of each patch $\Omega \mathbf{P}_j X$ can be well-represented sparsely by \mathbf{z}_j .

⁶ Amplitude-only microscope images in the original paper

The problem in Eq. (28) is solved iteratively and resembles the solution in Algorithm 6 closely. However, the reconstruction update step on Line 12 of Algorithm 6 is now given with the updated regularizer by

$$\min_{\mathbf{X}} \frac{1}{2} \|\mathbf{H} \cdot \mathbf{W} - T(z)(\mathbf{X} + \mu \mathbf{1})\|_F^2 + \lambda_1 \sum_{j=1}^N \|\Omega \mathbf{P}_j \mathbf{X} - \mathbf{z}_j\|_2^2 \quad (30)$$

s.t

$$\mathbf{P}_\mu(\mathbf{X})$$

Which can be solved with the closed-form solution

$$\mathbf{X} = \mathbf{P}_\mu \left(\frac{1}{1 + 2\lambda_1 l} \left[T(-z)(\mathbf{H} \cdot \mathbf{W}) - \mu \mathbf{1} \right. \right. \quad (31)$$

$$\left. \left. + 2\lambda_1 \sum_{j=1}^N \mathbf{P}_j^H \Omega^H \mathbf{Z}_j \right] \right) \quad (32)$$

Following the reconstruction update step, the sparse representation \mathbf{Z} is updated by hard-thresholding the current sparse transformation of \mathbf{X} :

$$\mathbf{Z}_j = H_\gamma(\Omega \mathbf{P}_j \mathbf{X}) \quad (33)$$

with the coefficient γ from the regularizer in Eq. (29). The complete algorithm is shown in Algorithm 7.

Algorithm 7 Learned Sparsifying Transform based Holographic Reconstruction

```

1:  $n_t$ : Number of iterations of Sparsifying Transform learning
2:  $n_r$ : Number of iterations of holographic reconstruction
3:  $N$ : Number of pixels in hologram
4:  $N_p$ : Number of training patches
5: for  $i$  of  $n_t$  do                                ▷ Learning Sparsifying Transform
6:    $\mathbf{U} \Sigma \mathbf{V}^H \leftarrow \text{SVD}(\mathbf{Y} \mathbf{Z}^H)$ 
7:    $\Omega \leftarrow \mathbf{V} \mathbf{U}^H$ 
8:    $\mathbf{Z}_j \leftarrow H_\gamma(\Omega \mathbf{Y})$ 
9: end for
10:  $\mathbf{X}, \mathbf{W}, \mu \leftarrow \text{ASC}(\mathbf{H})$                         ▷ Initialize using Algorithm 6
11:  $\mathbf{Q} \leftarrow T(-z) \odot (\mathbf{H} \cdot \mathbf{W})$ 
12:  $\mathbf{Z}_j \leftarrow H_\gamma(\Omega \mathbf{P}_j \mathbf{X}) \forall j$ 
13: for  $k$  of  $n_r$  do                                ▷ Holographic Reconstruction
14:    $\mu \leftarrow \frac{1}{N} \langle \mathbf{Q} - \mathbf{X}, \mathbf{1} \rangle$ 
15:    $\mathbf{W} \leftarrow \exp(i \angle [\mu \mathbf{1} + T(z) \odot \mathbf{X}])$ 
16:    $\mathbf{Q} \leftarrow T(-z) \odot (\mathbf{H} \cdot \mathbf{W})$ 
17:
18:    $\mathbf{Z}_j \leftarrow H_\gamma(\Omega \mathbf{P}_j \mathbf{X}) \forall j$ 
19: end for
20: return  $\mathbf{X}$ 

```

Yuan et al. [21] show the results of the application on experimental holograms of their reconstruction algorithm using two samples; a 1951 USAF target and a microscope slide of a pumpkin stem, and compares them to both the Angular Spectrum backpropagation in Section 2.1 and the Absorption Constrained Sparse Phase Retrieval method in Algorithm 6. The results rightly show reduced twin-image artifacts and staircase edges in the reconstructions of the highly non-sparse samples as hypothesized, with resulting increased reconstruction contrast. The authors specify that the images on which the learned transformation Ω is trained on are not necessarily required to resemble the hologram samples for which it will be used. The aim is for Ω to learn a transformation that is specific enough to be a more efficient sparsifying transform than the DCT or other *a priori* known transforms, but generalizable enough to be able to sparsify reconstructions within a similar

area of application. One might imagine that a transformation trained on the pumpkin stem microscope slide can be applicable to other plant microscope slides and, perhaps to a lesser degree, to general biological samples. In addition, though the sparsifying transform is trained on an amplitude-only microscope image for the pumpkin stem example, the algorithm is able to improve upon the phase reconstruction when compared to Algorithm 6. However, the authors speculate that using phase-retrieved training images may improve results even further.

4. Machine learning methods

The use of Machine learning (ML) for solving the holographic reconstruction problem presents several advantages. First, following the initial and computationally expensive training phase, unseen holograms can (in most cases) be reconstructed non-iteratively by a single pass through the trained network. This yields an inherent computation speed increase that is seen in practically all areas where iterative methods have previously dominated. Furthermore, many of the networks are trained to be “system agnostic”, such that the influence of experimental parameters (sensor pixel size, non-uniform illumination, sample/sensor shift, etc.) are baked into the networks, allowing for inherent aberration compensation. The drawback of many of the ML methods are, of course, the black-box nature of neural networks. In order to use the networks for holographic imaging efficiently and wisely, an understanding of the foundational physics is necessary.

In this review, the landscape of ML and Deep Neural Network (DNN) based holographic imaging algorithms has here been divided into three more or less broad categories; End-to-end networks, Generative Adversarial networks, and Untrained Deep Decoder networks, which are also shown in Fig. 1.

4.1. End-to-end networks

Although the networks included in this section deviate significantly in design, implementation, and application, the end-to-end (E2E) nature (needing no human feature design or extraction) for solving inverse problems can be abstracted to any task as follows. Given a chosen network architecture $f_\theta(\cdot)$, where θ are the parameters of the network (weights, biases, etc.), and a training dataset $D = \{X_i, y_i\}$, where X_i are ground truth images and y_i are the associated “raw” training images,⁷ the network is trained by optimizing θ such that,

$$\theta = \arg \min_{\theta} \mathcal{L} \{X_i, f_\theta(y_i)\}, \quad (34)$$

where \mathcal{L} is a chosen loss function. At inference, the network will, when presented with an unseen test image, y , approximate its reconstruction by a single pass through the network:

$$\mathbf{X} = f_\theta(y). \quad (35)$$

Now, as is shown in Fig. 1, the E2E category splits into both reconstruction and twin-image-removal networks. The E2E reconstruction networks accept “raw” holograms as input and, by training on associated “perfect” reconstructions, output either amplitude [53], phase [24, 25, 53], or full complex [26] reconstructions. Sinha et al. [25], Wang et al. [24], and Ren et al. [53] all propose convolutional neural networks (CNN) that share many characteristics; all three employ an encoding/decoding-path structure, similar to the popular UNet architecture [55], in order to facilitate the learning of the structure of the input holograms at multiple spatial levels. In addition, they all take advantage of the residual blocks of ResNet [56] in place of the more conventional convolutional block in certain stages of the networks. Sinha et al. [25] and Wang et al. [24] both train and test their networks on experimental on-axis holography setups utilizing reflective Spatial

⁷ Raw holograms for [24–26, 53, 54], noisy reconstructions for [18], and either for [19].

Light Modulators (SLM) on which phase-only images X_i can be displayed. The interference patterns H_i from the SLMs are then captured on CMOS sensors, from which their respective datasets $D = \{X_i, I_i\}$ are constructed. As such, their networks are trained to reconstruct pure phase objects. However, there is no reason that their networks could not, with a slight change to the experimental setup, be trained on, and applied to, pure amplitude objects, as is done by Ren et al. [53], who train identical but separate networks (named HRNet) for either object type. In addition to their E2E hologram reconstruction network, Ren et al. [53] also present network modalities capable of extended-focus-imaging (EFI) and depth map (DM) reconstruction. These network modalities are trained on samples consisting of two separate target planes placed at different axial distances to the detector and to each other. Then, the EFI network is trained to produce an all-in-focus reconstruction of the two target planes, while the DM reconstruction is able to distinguish between the two planes and categorize their respective axial focus distances.

Now, due to the lateral extent of raw holograms in combination with the small pixel pitches necessary to capture high frequency information in the holograms, the number of trainable parameters in these networks grow rapidly, leading to increased training and inference time, in addition to very large memory requirements. As such, it may be impractical, if not impossible, to continue to extend these networks further in depth or input size without the computational cost exploding.

The aforementioned networks all reconstruct either pure phase or pure amplitude samples. An instance of complex reconstruction via E2E networks is presented by Ju et al. [26], whose UNet based network is trained to accurately retrieve the phase of the hologram plane, allowing for later numerical reconstruction to arbitrary focus planes without twin-images. The ground truth complex holograms are gathered by use of a polarized image sensor, whose pixels consist of four sub-pixels, each containing linear polarizers configured at angles 0° , 45° , 90° , and 135° respectively. By circularly polarizing the object (and reference) beam prior to the image sensor, the four polarizing sub-pixels result in distinct phase shifts between the object and reference beam, allowing for the calculation of the complex hologram by

$$U_h = (I_{0^\circ} - I_{90^\circ}) + i(I_{45^\circ} - I_{135^\circ}). \quad (36)$$

During training, the proposed network then attempts to recover this complex field from a single polarization (I_{0°) image as input, without the information of the three other polarizations. Ju et al. [26] show reconstructions of both simulated and experimental samples. As expected, the results are mainly focused on the *a posteriori* propagation capabilities of the method. It is shown how the recovered complex hologram can be propagated to multiple focus planes, each containing separate simulated samples. The interaxial distances between samples are well-maintained by the recovered complex holograms of the network. It is also shown to function unimpeded by overlapping samples positioned at different axial distances.

The other branch of the E2E section in Fig. 1 concerns twin-image suppression. Rivenson et al. [18] propose a network capable of recovering the correct complex object field of microscope samples from “simple” (ASM) numerical reconstructions contaminated by significant twin-image and interferometric aberrations. The network takes as input the complex object field that results from numerical propagation of some “raw” input hologram to a desired object plane, located at a desired axial distance. The complex input field is split into its real and imaginary component images and stacked to form a 2-channel input layer. The input layer then serves as input for four separate information paths of the network, each of which processing the complex field at separate spatial dimensions via downsampling by factors of $\times 1$, $\times 2$, $\times 4$, and $\times 8$. As mentioned before, the processing of inputs at multiple scales increases the receptive field of the network, allowing it to better learn the structure of the input. Each path contains four residual blocks, followed by upsampling by the respective factor, in order to regain the original reconstruction resolution. The paths are

finally concatenated depth-wise and processed in a convolutional layer in order to achieve the final 2-channel complex reconstruction (real and imaginary). The recovered complex fields are compared to ground truth fields – obtained via multi-height hologram reconstruction [35] – in a loss function comparing real and imaginary components separately. The proposed network is trained on pap smears, blood smears, and breast tissue separately resulting in three distinct trained networks. In addition, a network with twice the feature maps in its convolutional layers was trained on the complete dataset containing all sample types to create a “Universal” network. Rivenson et al. [18] show impressive twin-image suppression on all sample types and report on the reconstruction SSIM of both the sample-type-specific (STS) and universal networks, after testing on unseen microscope images. As expected, the STS trained networks perform well on the sample types on which they were trained. However, when presented with sample types not included in their training data, artifacts appear in the outputs. The deeper, universally trained network performed similarly to the STS networks for all sample types. Additionally, the authors calculate an *effective refractive volume* [fL] (volume of a cell based on the *average* refractive index) and a cell phase integral [$\text{rad} \cdot \mu\text{m}^2$] (the relative phase shift due to a cell w.r.t the background, integrated over the area of the cell) for each red blood cell (RBC) in a whole blood smear. These values are compared to the same calculations based on the ground truth images, as well as the unfiltered network inputs. The network outputs show far better alignment to the ground truth effective refractive volumes and cell phase integrals than the raw inputs, demonstrating that these single hologram based network outputs may be used to evaluate biological samples quantitatively. While computationally expensive, the networks significantly reduce the computation time when compared to the multi-height phase recovery of the ground truth images.

4.2. Generative adversarial networks

Generative Adversarial Networks (GAN) [57] differ fundamentally from E2E networks by their process of learning. Where E2E networks learn by updating their weights based solely on direct comparisons (pixel wise, statistical, etc.) between their outputs and ground truth images, GANs learn by the including *adversarial* loss. The term adversarial stems from the fact that, during training, GANs usually consist of two competing networks (adversaries); the generator and discriminator, that attempt to outperform each other. Through training, the generator learns to replicate the ground truth training data, based on its “raw” input data.⁸ At the same time, the competing network – the discriminator – learns to distinguish between ground truth images and the generator outputs. Each competing network is updated based on individual loss functions, each responding differently to the result of the discriminator’s classification. The misclassification of ground truth images and generator outputs by the discriminator results in low loss for the generator, as it successfully replicated the ground truth data to such an extent that the discriminator was fooled. With the same misclassification, the discriminator loss is high, as it did not succeed in its task. The combined parallel training of the two networks work to improve the abilities of both. The addition of adversarial loss to the generator loss function (which still includes some pixel based loss), amounts to the quantification of more abstract differences between the network outputs and the ground truths. Mathematically, this GAN structure can be abstracted as follows. Given generator and discriminator network architectures $g_\theta(\cdot)$ and $d_\phi(\cdot)$, respectively, and a training dataset $D = \{X_i, y_i\}$, where X_i are ground truth images and y_i are “raw” training images, the following minimization problems are solved during training, updating the parameters of both adversarial networks:

$$\phi = \arg \min_{\phi} \frac{1}{2} d_\phi(g_\theta(y_i))^2 + \frac{1}{2} [d_\phi(X_i) - 1]^2 \quad (37)$$

⁸ Typically a latent random variable. In this application, however, the GANs function as image-to-image translators, and thus accept raw holograms/reconstructions as inputs.

$$\theta = \arg \min_{\theta} \mathcal{L}_{\text{pixel}}(g_{\theta}(y_i), X_i) + \lambda_1 [d_{\phi}(g_{\theta}(y_i)) - 1]^2 \quad (38)$$

In Eq. (37), the discriminator loss consists of two terms evaluating both the discrimination of the network output and the ground truth. The discriminator d_{ϕ} outputs a value between 0 and 1, based on its assuredness that its input is the ground truth data. Thus, if the discriminator is well-trained, and is able to easily distinguish the network output from the ground truth, then $d_{\phi}(g_{\theta}(y_i)) \rightarrow 0$ and $d_{\phi}(X_i) \rightarrow 1$, which lowers the discriminator loss correctly.

In Eq. (38), the generator loss consists of a pixel-wise loss function (mean-square-error, SSIM, etc.), and an adversarial term. If the generator is well-trained and is able to replicate the training data perfectly, then $d_{\phi}(g_{\theta}(y_i)) \rightarrow 1$, which yields a low loss value.

After training of the two networks, only the generator network is used for inference. When presented with a new “raw” input y , the generator outputs its approximate for a corresponding ground truth image X , by

$$X = g_{\theta}(y). \quad (39)$$

Now, there exists many implementations of GANs, but the ones described here have all been specialized to the task of holographic microscopy, in both reconstruction and twin-image suppression. Huang et al. [19] present two UNet architecture Recurrent Neural Network (RNN) based GANs, capable of each task individually. RNNs are specially designed networks for processing *sequential* data.⁹ The efficacy of RNNs stems from the cyclic nature of recurrent blocks that allow the output from a block to influence subsequent input to the same block. This makes it possible for the network to “remember” and utilize prior information to shape its output. Huang et al. [19] take advantage of this fact by treating the holographic reconstruction problem sequentially in axial distance, much like the multi-height phase recovery methods [35, 58, 59]. As such, to reconstruct the complex field U at the sample plane, the authors capture M holograms in a z-stack with increasing sample-to-sensor distances. Since the structure of these holograms are interconnected due to the free-space propagation that exists between them, the RNN blocks of the network can take advantage of its outputs from hologram planes I_0, I_1 , etc. when evaluating hologram I_M . The sequential input of M holograms results, after a pass through the network, in a two-channel layer, constituting the real and imaginary part of the field at the sample plane. This complex field is then evaluated in a loss function similar to Eq. (38), where the pixel-wise and adversarial loss is calculated from multi-height phase retrieval reconstructed ground truth images. From the same result, the discriminator is updated by a loss function similar to Eq. (37).

In the same paper, Huang et al. [19] propose to use an identical network structure to perform twin-image suppression. To accomplish this, a z-stack of holograms is again captured, each of which is propagated to the same sample plane using some numerical propagator. Then, the RNN-GAN is trained on this stack of noisy reconstructions, whose influence of twin-images is proportional to their individual propagation distances, to replicate the associated ground truth complex sample field.

The networks presented by Huang et al. [19] are better compared to the iterative multi-height phase retrieval methods than the single hologram methods of the previous sections, since they require the capture of two or more holograms.¹⁰ As such, although they show excellent reconstruction quality, twin-image removal, and depth of field, there exists a trade-off with the mechanical complexity and capture time of the experimental setup, when compared to the single hologram methods.

⁹ Typically temporal data applications, e.g., language modeling, speech recognition, time-series anomaly detection, etc.

¹⁰ Reconstruction and twin-image suppression is significantly improved by including more holograms.

4.3. Unlearned neural networks as image priors

Niknam et al. [15] propose a holographic reconstruction network architecture based on Deep Image Prior [60, 61] (DIP), which they dub a Deep Compressed Object Decoder (DCOD). DCOD is an implementation a physics informed generative untrained deep neural network, capable of holographic reconstruction without the need for training data. In fact, the deep optimization of the network occurs at inference, hence the overlapping location of the method in Fig. 1. Using such unlearned networks as image priors is fundamentally different from both E2E networks and GANs. Given some randomly initialized deep decoder network (DDN) architecture $f_{\theta}(\cdot)$, which transforms input data (fixed random tensor from some distribution $Z \sim p_z$) from a low-dimensional latent space \mathbb{R}^k to a higher dimensional space \mathbb{R}^n where $n \gg k$, the inverse holographic reconstruction problem can be solved by minimizing

$$X = \arg \min_{\theta} \{ \|I - |f_{\theta}(Z) \otimes T(z)|^2\|_F^2 + R(\theta) \} \quad (40)$$

for any given captured hologram H . Here R is a regularizer aimed at reducing the proneness of the network to overfit to aberrations, twin-images, and other noise. More clearly, Niknam et al. [15] propose an algorithm that consists of a randomly initialized deep decoder network f_{θ} , which is implemented as a stack of convolutional and upsampling layers that accepts a fixed random tensor Z , and outputs a two-channel layer, corresponding to the estimated sample amplitude and sample phase. The complex sample field is constructed from estimated amplitude and phase, and is then propagated to the hologram plane using numerical angular spectrum propagation where the intensity is calculated. The estimated hologram intensity is then compared to the ground truth captured hologram in the loss function, as shown in Eq. (40), and the parameters of the DDN, θ , are updated accordingly. Niknam et al. [15] present impressive amplitude and phase reconstructions of lensless holograms of smeared cheek and blood cells, a diffraction grating, and a 1951 USAF target. The advantage of using these unlearned networks for holographic reconstruction lies in the fact that no training data is necessary. As such, it resembles the iterative methods of Section 3, while still being able to take advantage of ML practices. Unfortunately, it presumably also carries with it the accompanying computation time of the iterative methods, which, when compared to the single inference pass of a trained network must be much greater, despite the optimization taking place on an efficient GPU.

5. Conclusion

In this review, we have presented an overview of state-of-the-art reconstruction methods for on-axis DHM. These methods have been broadly categorized and mapped (Fig. 1) into three main categories: direct, iterative, and machine learning methods. Direct methods are the most straightforward approaches to holographic reconstruction. These methods directly compute the complex object field from the recorded hologram. These methods are computationally efficient and do not require any additional information about the sample. However, their accuracy and fidelity is limited by the quality of the recorded hologram and twin-image artifacts.

Iterative methods, on the other hand, use an iterative approach to improve the reconstructed field. These methods typically involve the optimization of some objective function, which may incorporate additional information about the sample or the imaging system. For example, some iterative methods assume that the sample is sparse, or sparse in a certain domain. This assumption allows for the use of sparsity-promoting regularization, which can significantly improve the quality of the reconstructed image. However, iterative methods are generally more computationally intensive than direct methods.

Finally, machine learning methods have emerged as a promising approach to holographic reconstruction. These methods use machine learning algorithms, such as CNNs, GANs, etc., to learn the mapping

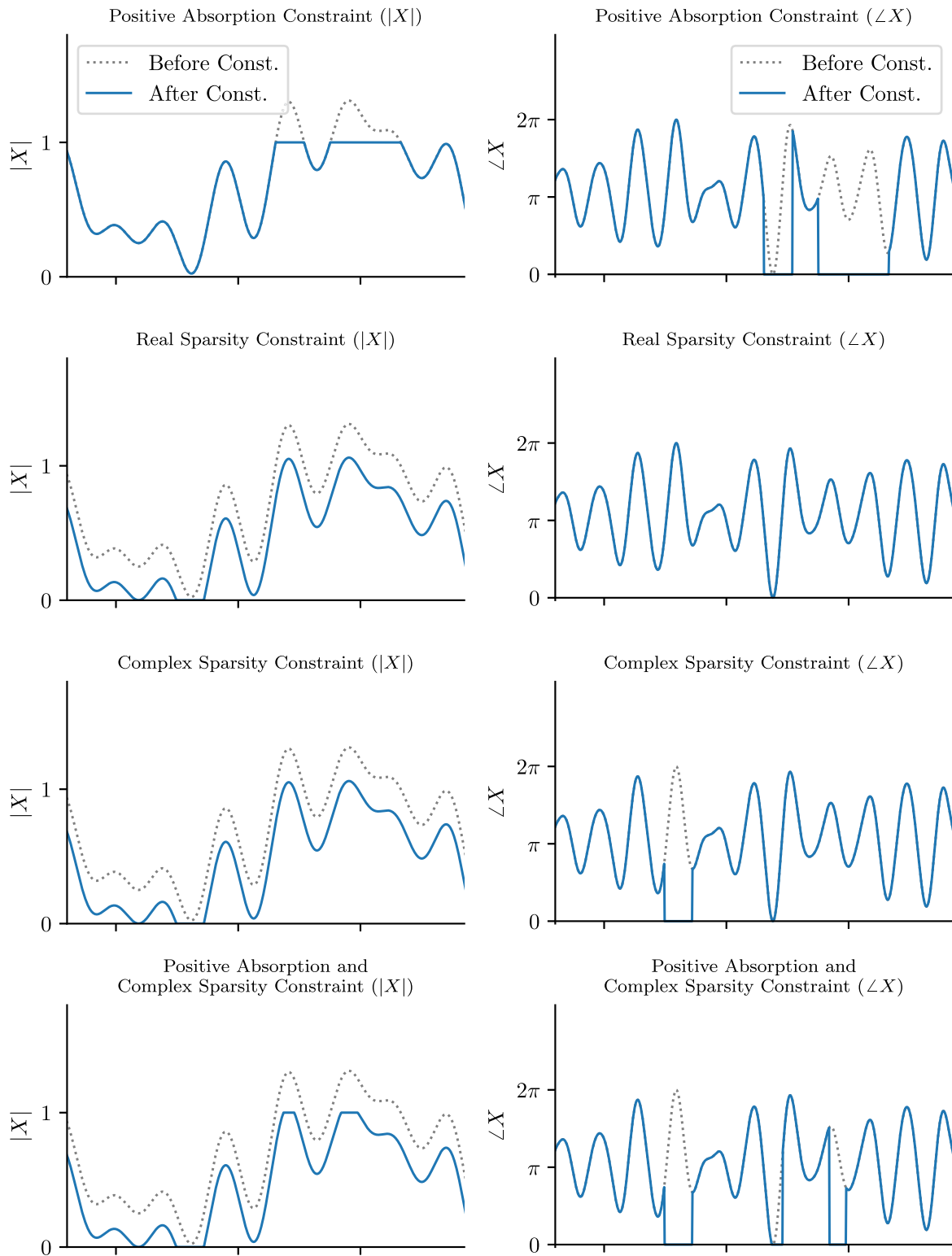


Fig. A.4. Illustrations showing the effect of applying the constraints of the iterative methods on a complex signal (dummy data) once. In the left column, the magnitude of the signals are shown and on the left, the phases are shown. From the top, the constraints are defined in [13,45,48], and [23].

from the recorded hologram to the complex amplitude of the object wave. These methods exploit large amounts of data, allowing them to learn the underlying patterns and features of the data to improve the accuracy of the reconstructed image. However, the performance of machine learning methods can vary greatly depending on the quality of the training data and the design of the machine learning algorithm.

Overall, it is clear that the approaches presented in this review each have their own advantages and limitations, and the choice of reconstruction method will depend on the specific application and imaging conditions. Further research is needed to fully understand the underlying physics and optimize these methods, such that the advantages of on-axis DHM can be exploited fully.

Table B.1

Reported values extracted from papers described in this review. Where no value was reported, a dash (–) has been written. Reported errors include the amplitude $|X|$, angle $\angle X$, real part $\text{Re}\{X\}$, and imaginary part $\text{Im}\{X\}$ of the reconstructions. The error measurements include peak signal-to-noise ratio (PSNR), (multi-scale) structural similarity index measurement ((MS)-SSIM), mean-square error (MSE), mean absolute error (MAE), root-mean-square error (RMSE), standard deviation (STD), and a thresholding-and-summation metric (TSM). Truth column describes whether a ground truth image was used for evaluating the method, and how such an image was created; experimentally (Exp.) or in simulation (Sim.).

Author(s)	Category	Hol. size [px]	Comp. time [s]	Error [$ X $, $\angle X$, $\text{Re}\{X\}$, $\text{Im}\{X\}$, I]	Error meas.	Truth
Gnetto et al. [14]	Direct	–	–	–	–	–
McElhinney et al. [17]	Direct	4.19M	8.4	–	–	–
de Almeida et al. [20]	Iterative	10.55M	–	$[11.8 - 29.47]/[0.59 - 0.94]$	PSNR/MSSIM	Sim.
Latychevskaia and Fink [13]	Iterative	.25M	–	$1.4E - 7$; $4.6E - 7$	MSE	Sim.
Rong et al. [62]	Iterative	1.05M	–	$\sim 3E - 14$; $\sim 2E - 14$	MSE	Sim.
Denis et al. [45]	Sparse iterative	1.31M	–	–	–	–
Haeffele et al. [48]	Sparse iterative	8.39M	0.7	–	–	–
Li et al. [23]	Sparse iterative	0.33M	–	$\sim 0.5E - 3$; $\sim 1.8E - 3$	MAE	Sim.
Yuan et al. [21]	Sparse iterative	4.00M	–	–	–	–
Guo et al. [22]	Iterative	1.05M	775	0.851	SSIM	Exp.
Sinha et al. [25]	E2E	1.05M	–	$\sim 0.05\pi$	MAE	Exp.
Wang et al. [24]	E2E	0.59M	–	$[21 - 6]/[0.92 - 0.99]/[0.16\pi - 0.05\pi]$	RMSE/SSIM/MAE	Exp.
Ren et al. [53]	E2E	0.64M	1.21	$24.62/0.91$; $30.49/0.96$	PSNR/SSIM	Exp.
Ju et al. [26]	E2E	1.27M	–	$35.35/0.94$; $33.28/0.899$	PSNR/SSIM	Exp.
Rivenson et al. [18]	E2E twin supp.	$\sim 7.30M$	6.45 ^a	0.942; 0.93	SSIM	Exp.
Huang et al. [19]	GAN	–	0.154 (/mm ²)	0.7; 0.69	SSIM	Exp.
Huang et al. [19]	GAN twin supp.	–	0.162 (/mm ²)	$\sim 0.15/0.69$; $\sim 0.19/0.69$	RMSE/SSIM	Exp.
Niknam et al. [15]	Unlearned network	0.26M	–	–	–	–

^aIncluding pre-processing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Andreas Erik Gejl Madsen reports financial support was provided by Novo Nordisk Foundation. Andreas Erik Gejl Madsen reports financial support was provided by Innovation Fund Denmark. Andreas Erik Gejl Madsen reports a relationship with Radiometer Medical ApS that includes: employment. Mr. Madsen is employed as an Industrial PhD student by Radiometer Medical ApS, in a collaboration with the University of Southern Denmark.

Data availability

No data was used for the research described in the article.

Acknowledgments

This work has been supported by the Novo Nordisk Foundation, Denmark (Grand Challenge Program; NNF16OC0021948), the Innovation Fund Denmark, and by Radiometer Medical ApS, Denmark.

Appendix A. Constraint comparison

In Fig. A.4, illustrations of the effect of the various regularizing constraints applied to the reconstruction plane in Section 3 are shown. The positive absorption constraint introduced by Latychevskaia and Fink [13] forces the parts of the signals that contain too much energy to unity. In the same locations, the phase is set to zero. In the real sparsity constraint set by Denis et al. [45], the SFT operator lowers the entire signal area by the threshold value, while setting any values below the threshold to zero. The complex sparsity constraint by Haeffele et al. [48] is nearly identical, but with the addition of phase suppression in areas below the threshold. The final constraint by Li et al. [23] combines the positive absorption and complex sparsity constraint.

Appendix B. Comparison of reported values

In Table B.1, reported values from the sources are presented. In the table, we have included the category of the method (direct, iterative, sparse iterative, E2E, E2E twin-suppression, GAN, GAN twin suppression, unlearned network), the size of the holograms which are tested in total pixels, the computation time*, error values (with text color corresponding to magnitude, phase, real part, imaginary part, and intensity, respectively) and their corresponding measure of error*, and the type of data the method was tested on in the source. The purpose of the table is to give the reader an overview of the properties and quantitative results provided by each source, and to serve as a convenient indicator of whether a source is relevant for the specific interest of the reader.

References

- [1] F. Zernike, How i discovered phase contrast, *Science* 121 (3141) (1955) 345–349, <http://dx.doi.org/10.1126/science.121.3141.345>.
- [2] P.J. Rodrigo, D. Palima, J. Glückstad, Accurate quantitative phase imaging using generalized phase contrast, *Opt. Express* 16 (4) (2008) 2740, <http://dx.doi.org/10.1364/OE.16.002740>.
- [3] D. Palima, J. Glückstad, Diffractive generalized phase contrast for adaptive phase imaging and optical security, *Opt. Express* 20 (2) (2012) 1370, <http://dx.doi.org/10.1364/OE.20.001370>.
- [4] J. Glückstad, P.C. Mogensen, Optimal phase contrast in common-path interferometry, *Appl. Opt.* 40 (2) (2001) 268, <http://dx.doi.org/10.1364/AO.40.000268>.
- [5] E. Engay, A. Bañas, A.-I. Bunea, S.D. Separa, J. Glückstad, Interferometric detection of OAM-carrying helico-conical beams, *Opt. Commun.* 433 (2019) 247–251, <http://dx.doi.org/10.1016/j.optcom.2018.10.019>.
- [6] P.F. Almoró, J. Glückstad, S.G. Hanson, Single-plane multiple speckle pattern phase retrieval using a deformable mirror, *Opt. Express* 18 (18) (2010) 19304, <http://dx.doi.org/10.1364/OE.18.019304>.
- [7] A.E.G. Madsen, R.L. Eriksen, J. Glückstad, Comparison of state-of-the-art computer generated holography algorithms and a machine learning approach, *Opt. Commun.* 505 (2022) 127590, <http://dx.doi.org/10.1016/j.optcom.2021.127590>.
- [8] A.G. Madsen, J. Glückstad, HoloTile: rapid and speckle-suppressed digital holography by matched sub-hologram tiling and point spread function shaping, *Opt. Commun.* 525 (2022) 128876, <http://dx.doi.org/10.1016/j.optcom.2022.128876>.
- [9] D. Gabor, et al., A new microscopic principle, *Nature* 161 (1948) 777–778.
- [10] E.N. Leith, J. Upatnieks, Reconstructed wavefronts and communication theory*, *J. Opt. Soc. Amer.* 52 (10) (1962) 1123, <http://dx.doi.org/10.1364/JOSA.52.001123>.

* If reported.

- [11] E.N. Leith, J. Upatnieks, Wavefront reconstruction with continuous-tone objects*, *J. Opt. Soc. Amer.* 53 (12) (1963) 1377, <http://dx.doi.org/10.1364/JOSA.53.001377>.
- [12] E.N. Leith, J. Upatnieks, Wavefront reconstruction with diffused illumination and three-dimensional objects*, *J. Opt. Soc. Amer.* 54 (11) (1964) 1295, <http://dx.doi.org/10.1364/JOSA.54.001295>.
- [13] T. Latychevskaia, H.-W. Fink, Solution to the twin image problem in holography, *Phys. Rev. Lett.* 98 (23) (2007) 233901, <http://dx.doi.org/10.1103/PhysRevLett.98.233901>.
- [14] M.I. Gnetto, Y.T.A. Kossonou, Y. Koffi, K.A. Kaduki, J.T. Zoueu, Solving the twin image problem in in-line holography by using multiple defocused intensity images reconstructed from a single hologram, *J. Modern Opt.* 69 (3) (2022) 121–129, <http://dx.doi.org/10.1080/09500340.2021.2011450>.
- [15] F. Niknam, H. Qazvini, H. Latifi, Holographic optical field recovery using a regularized untrained deep decoder network, *Sci. Rep.* 11 (1) (2021) 10903, <http://dx.doi.org/10.1038/s41598-021-90312-5>.
- [16] L. Denis, C. Fournier, T. Fournel, C. Ducottet, Twin-image noise reduction by phase retrieval in in-line digital holography, in: M. Papadakis, A.F. Laine, M.A. Unser (Eds.), *Optics & Photonics 2005*, San Diego, California, USA, 2005, p. 59140J, <http://dx.doi.org/10.1117/12.617405>.
- [17] C. McElhinney, B.M. Hennelly, L. Ahrenberg, T.J. Naughton, Removing the twin image in digital holography by segmented filtering of in-focus twin image, in: A.A.S. Awwal, K.M. Iftikharuddin, B. Javidi (Eds.), *Optical Engineering + Applications*, San Diego, California, USA, 2008, 707208, <http://dx.doi.org/10.1117/12.795894>.
- [18] Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, A. Ozcan, Phase recovery and holographic image reconstruction using deep learning in neural networks, *Light. Sci. Appl.* 7 (2) (2018) 17141, <http://dx.doi.org/10.1038/lsa.2017.141>.
- [19] L. Huang, T. Liu, X. Yang, Y. Luo, Y. Rivenson, A. Ozcan, Holographic image reconstruction with phase recovery and autofocusing using recurrent neural networks, 2021, p. 18.
- [20] J.L. de Almeida, E. Comunello, A. Sobieranski, A.M. da Rocha Fernandes, G.S. Cardoso, Twin-image suppression in digital in-line holography based on wavefront filtering, *Pattern Anal. Appl.* 24 (3) (2021) 907–914, <http://dx.doi.org/10.1007/s10044-020-00949-7>.
- [21] S. Yuan, H. Cui, Y. Long, J. Wu, Digital inline holographic reconstruction with learned sparsifying transform, *Opt. Commun.* 498 (2021) 127220, <http://dx.doi.org/10.1016/j.optcom.2021.127220>.
- [22] C. Guo, X. Liu, F. Zhang, Y. Du, S. Zheng, Z. Wang, X. Zhang, X. Kan, Z. Liu, W. Wang, Lensfree on-chip microscopy based on single-plane phase retrieval, *Opt. Express* 30 (11) (2022) 19855, <http://dx.doi.org/10.1364/OE.458400>.
- [23] Z. Li, Q. Yan, Y. Qin, W. Kong, G. Li, M. Zou, D. Wang, Z. You, X. Zhou, Sparsity-based continuous wave terahertz lens-free on-chip holography with sub-wavelength resolution, *Opt. Express* 27 (2) (2019) 702, <http://dx.doi.org/10.1364/OE.27.000702>.
- [24] H. Wang, M. Lyu, G. Situ, eHoloNet: A learning-based end-to-end approach for in-line digital holographic reconstruction, *Opt. Express* 26 (18) (2018) 22603, <http://dx.doi.org/10.1364/OE.26.022603>.
- [25] A. Sinha, J. Lee, S. Li, G. Barbastathis, Lensless computational imaging through deep learning, *Optica* 4 (9) (2017) 1117, <http://dx.doi.org/10.1364/OPTICA.4.001117>.
- [26] Y.-G. Ju, H.-G. Choo, J.-H. Park, Learning-based complex field recovery from digital hologram with various depth objects, *Opt. Express* 30 (15) (2022) 26149, <http://dx.doi.org/10.1364/OE.461782>.
- [27] Z. Ren, J. Zhao, E.Y. Lam, Automatic compensation of phase aberrations in digital holographic microscopy based on sparse optimization, *APL Photonics* 4 (11) (2019) 110808, <http://dx.doi.org/10.1063/1.5115079>.
- [28] J.W. Goodman, *Introduction To Fourier Optics*, fourth ed., W.H. Freeman, Macmillan Learning, New York, 2017.
- [29] M.R. Teague, Deterministic phase retrieval: A Green's function solution, *J. Opt. Soc. Amer.* 73 (11) (1983) 1434, <http://dx.doi.org/10.1364/JOSA.73.001434>.
- [30] M. Beleggia, M. Schofield, V. Volkov, Y. Zhu, On the transport of intensity technique for phase retrieval, *Ultramicroscopy* 102 (1) (2004) 37–49, <http://dx.doi.org/10.1016/j.ultramicro.2004.08.004>.
- [31] C. Zuo, J. Li, J. Sun, Y. Fan, J. Zhang, L. Lu, R. Zhang, B. Wang, L. Huang, Q. Chen, Transport of intensity equation: A tutorial, *Opt. Lasers Eng.* 135 (2020) 106187, <http://dx.doi.org/10.1016/j.optlaseng.2020.106187>.
- [32] Z. Jingshan, R.A. Claus, J. Dauwels, L. Tian, L. Waller, Transport of intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes, *Opt. Express* 22 (9) (2014) 10661, <http://dx.doi.org/10.1364/OE.22.010661>.
- [33] M.R. Teague, Irradiance moments: Their propagation and use for unique retrieval of phase, *J. Opt. Soc. Amer.* 72 (9) (1982) 1199, <http://dx.doi.org/10.1364/JOSA.72.001199>.
- [34] L. Waller, Y. Luo, S.Y. Yang, G. Barbastathis, Transport of intensity phase imaging in a volume holographic microscope, *Opt. Lett.* 35 (17) (2010) 2961, <http://dx.doi.org/10.1364/OL.35.002961>.
- [35] A. Greenbaum, Y. Zhang, A. Feizi, P.-L. Chung, W. Luo, S.R. Kandukuri, A. Ozcan, Wide-field computational imaging of pathology slides using lens-free on-chip microscopy, *Sci. Transl. Med.* 6 (267) (2014) <http://dx.doi.org/10.1126/scitranslmed.3009850>.
- [36] D. Chen, L. Wang, X. Luo, H. Xie, X. Chen, Resolution and contrast enhancement for lensless digital holographic microscopy and its application in biomedicine, *Photonics* 9 (5) (2022) 358, <http://dx.doi.org/10.3390/photonics9050358>.
- [37] J. Li, N. Zhou, J. Sun, S. Zhou, Z. Bai, L. Lu, Q. Chen, C. Zuo, Transport of intensity diffraction tomography with non-interferometric synthetic aperture for three-dimensional label-free microscopy, *Light. Sci. Appl.* 11 (1) (2022) 154, <http://dx.doi.org/10.1038/s41377-022-00815-7>.
- [38] R.W. Gerchberg, W.O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures, *Optik* 35 (2) (1972) 237–246.
- [39] J.R. Fienup, Phase retrieval algorithms: A comparison, *Appl. Opt.* 21 (15) (1982) 2758, <http://dx.doi.org/10.1364/AO.21.002758>.
- [40] P.-A. Blanche (Ed.), *Optical holography: Materials, theory and applications*, Elsevier, St. Louis, Missouri, 2020.
- [41] E. Osherovich, Numerical methods for phase retrieval, 2012, *Astro-Ph, Physics: Physics*, arXiv:1203.4756.
- [42] Y. Romano, M. Elad, P. Milanfar, The little engine that could: regularization by denoising (RED), 2016, <http://dx.doi.org/10.48550/ARXIV.1611.02862>.
- [43] L.L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1–4) (1992) 259–268, [http://dx.doi.org/10.1016/0167-2789\(92\)90242-F](http://dx.doi.org/10.1016/0167-2789(92)90242-F).
- [44] Y. Chen, T. Pock, Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration, *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (6) (2017) 1256–1272, <http://dx.doi.org/10.1109/TPAMI.2016.2596743>.
- [45] L. Denis, D. Lorenz, E. Thiébaud, C. Fournier, D. Trede, Inline hologram reconstruction with sparsity constraints, *Opt. Lett.* 34 (22) (2009) 3475, <http://dx.doi.org/10.1364/OL.34.003475>.
- [46] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Comm. Pure Appl. Math.* 57 (11) (2004) 1413–1457, <http://dx.doi.org/10.1002/cpa.20042>.
- [47] L. Landweber, An iteration formula for fredholm integral equations of the first kind, *Amer. J. Math.* 73 (3) (1951) 615, <http://dx.doi.org/10.2307/2372313>, arXiv:2372313.
- [48] B.D. Haeffele, R. Stahl, G. Vanmeerbeeck, R. Vidal, Efficient reconstruction of holographic lens-free images by sparse phase recovery, in: M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D.L. Collins, S. Duchesne (Eds.), *Medical Image Computing and Computer-Assisted Intervention*, Vol. 10434, MICCAI 2017, Springer International Publishing, Cham, 2017, pp. 109–117, http://dx.doi.org/10.1007/978-3-319-66185-8_13.
- [49] D.J. Brady, K. Choi, D.L. Marks, R. Horisaki, S. Lim, Compressive holography, *Opt. Express* 17 (15) (2009) 13040, <http://dx.doi.org/10.1364/OE.17.013040>.
- [50] Y. Endo, T. Shimobaba, T. Kakue, T. Ito, GPU-accelerated compressive holography, *Opt. Express* 24 (8) (2016) 8437, <http://dx.doi.org/10.1364/OE.24.008437>.
- [51] S. Ravishanker, Y. Bresler, Learning sparsifying transforms, *IEEE Trans. Signal Process.* 61 (5) (2013) 1072–1086, <http://dx.doi.org/10.1109/TSP.2012.2226449>.
- [52] S. Ravishanker, Y. Bresler, Sparsifying transform learning with efficient optimal updates and convergence guarantees, *IEEE Trans. Signal Process.* 63 (9) (2015) 2389–2404, <http://dx.doi.org/10.1109/TSP.2015.2405503>.
- [53] Z. Ren, Z. Xu, E.Y. Lam, End-to-end deep learning framework for digital holographic reconstruction, *Adv. Photonics* 1 (1) (2019) 1, <http://dx.doi.org/10.1117/1.AP.1.1.016004>.
- [54] I. Moon, K. Jaferzadeh, Y. Kim, B. Javidi, Noise-free quantitative phase imaging in Gabor holography with conditional generative adversarial network, *Opt. Express* 28 (18) (2020) 26284, <http://dx.doi.org/10.1364/OE.398528>.
- [55] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, 2015, <http://dx.doi.org/10.48550/ARXIV.1505.04597>.
- [56] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2015, <http://dx.doi.org/10.48550/ARXIV.1512.03385>.
- [57] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, *Generative adversarial nets*, 2014, p. 9.
- [58] Y. Park, C. Depeursinge, G. Popescu, Quantitative phase imaging in biomedicine, *Nat. Photonics* 12 (10) (2018) 578–589, <http://dx.doi.org/10.1038/s41566-018-0253-x>.
- [59] A. Ozcan, E. McLeod, Lensless imaging and sensing, *Annu. Rev. Biomed. Eng.* 18 (1) (2016) 77–102, <http://dx.doi.org/10.1146/annurev-bioeng-092515-010849>.
- [60] D. Ulyanov, A. Vedaldi, V. Lempitsky, Deep image prior, 2017, <http://dx.doi.org/10.48550/ARXIV.1711.10925>.
- [61] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, G. Situ, Phase imaging with an untrained neural network, *Light. Sci. Appl.* 9 (1) (2020) 77, <http://dx.doi.org/10.1038/s41377-020-0302-3>.
- [62] L. Rong, Y. Li, S. Liu, W. Xiao, F. Pan, D. Wang, Iterative solution to twin image problem in in-line digital holography, *Opt. Lasers Eng.* 51 (5) (2013) 553–559, <http://dx.doi.org/10.1016/j.optlaseng.2012.12.007>.