# JPhys Photonics

**PAPER • OPEN ACCESS**

# True 3D reconstruction in digital holography

View the article online for updates and enhancements.

# Journal of Physics: Photonics

**PAPER**

# True 3D reconstruction in digital holography

Jasleen Birdi, Sunaina Rajora, Mansi Butola and Kedar Khare (ORCID)

Department of Physics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

**E-mail:** kedark@physics.iitd.ac.in

## Abstract

We examine the nature of the reconstructed 3D image as obtained by replay (or back-propagation) of the object wave from the hologram recording plane to the original object volume. While recording of a hologram involves transferring information from a 3D volume to a 2D detector, the replay of the hologram involves creating information in a set of 3D voxels from a much smaller number of 2D detector pixels, which on a first look appears to be surprising. We point out that the hologram replay process is a Hermitian transpose (and not inverse) of the hologram formation process and therefore only provides an approximation to the original 3D object function. With the knowledge of this Hermitian transpose property, we show how one may realize true 3D image reconstruction via a regularized optimization algorithm. The numerical illustrations of this optimization approach as presented here show excellent slice-by-slice tomographic 3D reconstruction of the original object under the weak scattering approximation. In particular, the reconstructed 3D image field has near-zero numerical values at voxels where the original object did not exist. We note that 3D image reconstruction of this kind cannot be achieved by the traditional physical replay process. In this sense, the proposed methodology for digital holographic image reconstruction goes beyond numerically mimicking the physical process involved in traditional film based holographic replay. The reconstruction approach may find potential applications in a number of digital holographic imaging systems.

## 1. Introduction

Dennis Gabor's work [1] on improvement of the electron microscope led to the remarkable invention of the principle of holographic imaging. Since the hologram is an interference record, it captures information about amplitude as well as phase of the unknown object wave unlike the conventional photography which records intensity-only information. The replay field of the film based hologram has one component that corresponds to the back-propagating object wave. This component focuses back at the location of the original object leading to image formation. The visual inspection of this replay field gives one a perception of the reconstruction of the original three-dimensional (3D) object as our eyes are very good at concentrating on the focused image while ignoring any diffuse background. This 3D imaging aspect has always been an attractive feature of holography and is often highlighted in textbooks [2] as well as in classrooms.

For the last twenty five years, film based holography has slowly been replaced by digital holography [3–5] where the hologram or interference pattern is recorded on a digital sensor (e.g. CCD or CMOS array) and the image reconstruction task is performed numerically. One of the advantages of digital holography is that the interference data is now available in numerical form. Once the complex-valued object wave in the hologram plane is recovered from the interference data, the 3D image reconstruction is usually performed by mimicking the film based hologram replay process that consists of propagating the object wave field at the hologram plane back to the location of the original object volume. While the replay in film based holography was primarily used for applications involving visual observation of images, digital holography offers the possibility of quantitative imaging. This aspect has particularly become important in applications of digital holography to microscopy where the phase images of biological cells are being used for developing diagnostic tools [6, 7]. While a lot of literature has been devoted to digital holography and its various applications over

the past few decades, the nature of the 3D reconstruction offered by numerical back-propagation of the object wave still needs careful examination in our opinion. The terminology '3D imaging' is in fact often used somewhat incorrectly. For example, the typical phase images of cells obtained using a digital holographic microscope are often rendered as surface plots. The nature of 3D image representation in such cases is however not similar to a z-stack of slices as may be obtained from a confocal microscope.

One may begin the probing of this topic by noting that the holographic recording involves storing information about objects in a 3D volume onto a 2D digital array detector in the form of an interference pattern. A number of algorithmic methodologies have evolved over the past few decades that allow recovery of the object field in the hologram recording plane. For example, the off-axis holograms are typically processed using the Fourier transform method where the object field information is recovered by filtering the region near the cross-term peak in 2D Fourier transform of the hologram data frame [8]. The on-axis digital holograms on the other hand are typically processed using the multi-shot phase shifting framework [9, 10]. More recently, optimization methodologies have been developed that allow accurate full resolution object wave recovery from a single hologram frame uniformly for both on-axis as well as off-axis geometries [11–13]. The holographic replay of the object wave is a linear operation on the object wave field. If this operation truly reconstructed the original 3D object, it would mean recovery of voxels in the original object volume using comparatively a much smaller number of pixels in the hologram plane, which appears to be problematic. It is actually known from early days of holography that the back-propagated field is only an approximation to the original 3D object. In Gabor's own description [14], the term three-dimensional was only used 'to designate an image (or the process which produced it) that is substantially isomorphic with the original object'. In a film-based holography setup, there is not much one can do beyond physical replay of the field for image reconstruction, however, this need not be the case when object field in the hologram plane is available numerically. It is therefore important to investigate if a true 3D object reconstruction better than the usual back-propagation is even possible using the 2D object field information at the hologram plane. The question we investigate here is relevant to a number of allied problems such as holographic memory involving data storage in multiple layers [15, 16], Fresnel incoherent holographic imaging [17], integral imaging [18, 19], lightfield imaging [20], and optical metrology [21] where 3D information is typically needed using data recorded on a 2D detector array.

The problem of 3D image reconstruction has received attention particularly in case of in-line particulate holograms. By simple back-propagation of a particle hologram, it is possible to observe locations in space where the particles come into focus and localize them [22]. A more detailed approach involves fitting of Mie scattering models to particulate holograms [23] and this method has been shown to localize particulates in colloids. The problem of 3D field reconstruction has also been cast in the inverse problems framework mainly for in-line hologram data associated with particulates and thin thread-like objects [24–27]. In these approaches, the starting point is the recorded interference (intensity) pattern and the associated optimization problem aims to find a sharp focused object that agrees with the recorded intensity data with image domain constraints such as reality, positivity and sparsity. In these works, the non-linear term corresponding to object wave energy is assumed to be small (and treated as model error), and the data contains linear terms corresponding to both the object wave and its conjugate. The back-propagation of this data vector therefore contains both the focused object wave and its diverging twin. The role of constraints in these works is mainly to enforce object sparsity, however, implicitly the sparsity is also expected to eliminate the twin image. A de-convolution approach for 3D reconstruction has been presented in [28]. More recently, a deep learning based method has been developed for localization of particulates from their in-line holograms and has shown promising results [29]. An iterative method using two in-line intensity measurements for 3D reconstruction has also been demonstrated recently [30]. There is considerable interest in 3D slice-by-slice phase reconstruction of volume objects like biological cells. As per the fundamental theorem of diffraction tomography [31], knowledge of scattered field on two planes on either side of a weakly scattering object is sufficient to provide a low-pass filtered version of its 3D structure. This problem has also been handled using phase tomography approach where phase projections of a transparent 3D object are obtained by recording holograms of the same object from a number of view angles as allowed by tilt or other mechanisms [32, 33]. The phase projections are then used for tomographic image reconstruction.

Our aim in this work is to formulate the 3D reconstruction as an inverse problem using the numerically recovered object wave in the hologram plane as the starting point. First of all, this approach will not require us to worry about the dc or twin image elimination issues [34], and as a result, the nature of back-propagated object field in the original object volume can be understood in a simpler manner. Secondly, we do not make any assumption that the energy in non-linear term in the hologram is small. The treatment presented here is independent of the hologram recording configuration (in-line or off-axis) as our starting point is the numerically recovered object wave field in the hologram plane from single or multiple hologram frames. We find that this framework provides a simple analytical result that the back-propagation of the object field from

hologram plane is the Hermitian transpose of the hologram formation process. Further, this association is very useful in implementation of an iterative 3D image reconstruction algorithm that goes back and forth between the 3D object domain and the 2D hologram domain. The resultant 3D image reconstruction is also observed to be a significantly better quantitative estimate of the original 3D object compared to what is obtained by the traditional holographic replay process.

This paper is organized as follows. In section 2, we describe the true 3D reconstruction problem, firstly, in an intuitive way that is followed by a mathematical analysis of the problem. Section 3 provides the methodological description of the 3D reconstruction approach adopted by us. Our numerical experimentation results for amplitude and phase test objects are shown in section 4 for reflection and transmission geometries. Finally, in section 5, we provide concluding remarks.

## 2. Problem overview

### 2.1. Physical intuition

For illustrative purpose, let us consider a 3D object volume consisting of multiple planes with four point objects (see figure 1(a)). A spatially coherent monochromatic plane wave is split into two parts. One part reflects off the four point objects generating an object wave and the other part acts as a reference beam. At the 2D sensor plane, the interference pattern due to the total field from the four sources and the reference beam is recorded. We make an assumption that only the primary scattered field from each of the four objects predominantly forms the total object wave at the sensor. This assumption is reasonable and has been used since the early days of holography [35]. In the present work, we assume that the object field at the sensor/hologram plane is known using single-shot or multi-shot methods for complex object wave recovery in digital holography [10, 11, 13]. The task is to reconstruct the 3D object distribution from the numerically recovered 2D object field at the hologram plane. The conventional holographic techniques rely on illuminating the hologram with the conjugate reference beam. This leads to back-propagation of the object wave to the different planes of 3D volume as shown in figure 1(b). For illustration, we have shown the back-propagated field amplitudes corresponding to different particles in the original 3D volume with different colors. It is clear that in any plane where one of the point objects gets focused, there are de-focused versions of fields due to the other objects as well. This physical hologram reconstruction process can be mimicked for 3D image reconstruction in digital holography by numerical back-propagation of the object field. It is then evident that while the four point objects exist at specific locations within the volume as in figure 1(a), the back-propagated field is non-zero everywhere. The operation of back-propagation of the object wave from the hologram plane therefore does not provide a true 3D distribution of the original object. The nature of back-propagated object wave will be explained in more detail in the following discussion.

Before moving on to the mathematical analysis, we show a chart (see figure 2) describing the workflow involved in 3D image reconstruction in (a) film-based holography, (b) digital holography as typically practiced currently, and (c) digital holography via proposed methodology. This chart is intended to help provide a perspective for beginning optics researchers. As a first step, single or multiple (as in phase-shifting method) holograms corresponding to a 3D object may be recorded. In film based holography, a single hologram is physically re-illuminated with the conjugate reference beam. This leads to 3D image reconstruction at the location of the original object but this reconstructed field has dc and twin-image components that may overlap with the desired 3D reconstruction. In digital holography, the recorded hologram intensity pattern(s) may be processed numerically to obtain the object wave field in the hologram plane that we denote in this work by $V(x, y)$. This object wave $V(x, y)$ is then numerically back-propagated to the original 3D volume to generate the 3D image field. While this reconstruction does not have dc or twin image artefacts, the reconstructed field is non-zero in places where the original object did not exit (see figure 1). The proposed method starts with the object field $V(x, y)$ and uses an iterative optimization approach (that we will describe in the following section) for 3D image reconstruction. As mentioned above, the usage of $V(x, y)$ instead of the recorded hologram pattern as a starting point for 3D image reconstruction implies that issues like dc or twin image do not arise. Additionally, the inverse problem framework leads to a reconstruction which is a true representation of the original object as seen in the 3D image reconstruction results in section 4.

### 2.2. Mathematical analysis

Let the 3D complex-valued object function be denoted by $U(x', y'; z')$. Numerical values of $U(x', y'; z')$ represent the reflection (or the transmission) coefficient of the voxel at $(x', y', z')$ depending on the hologram
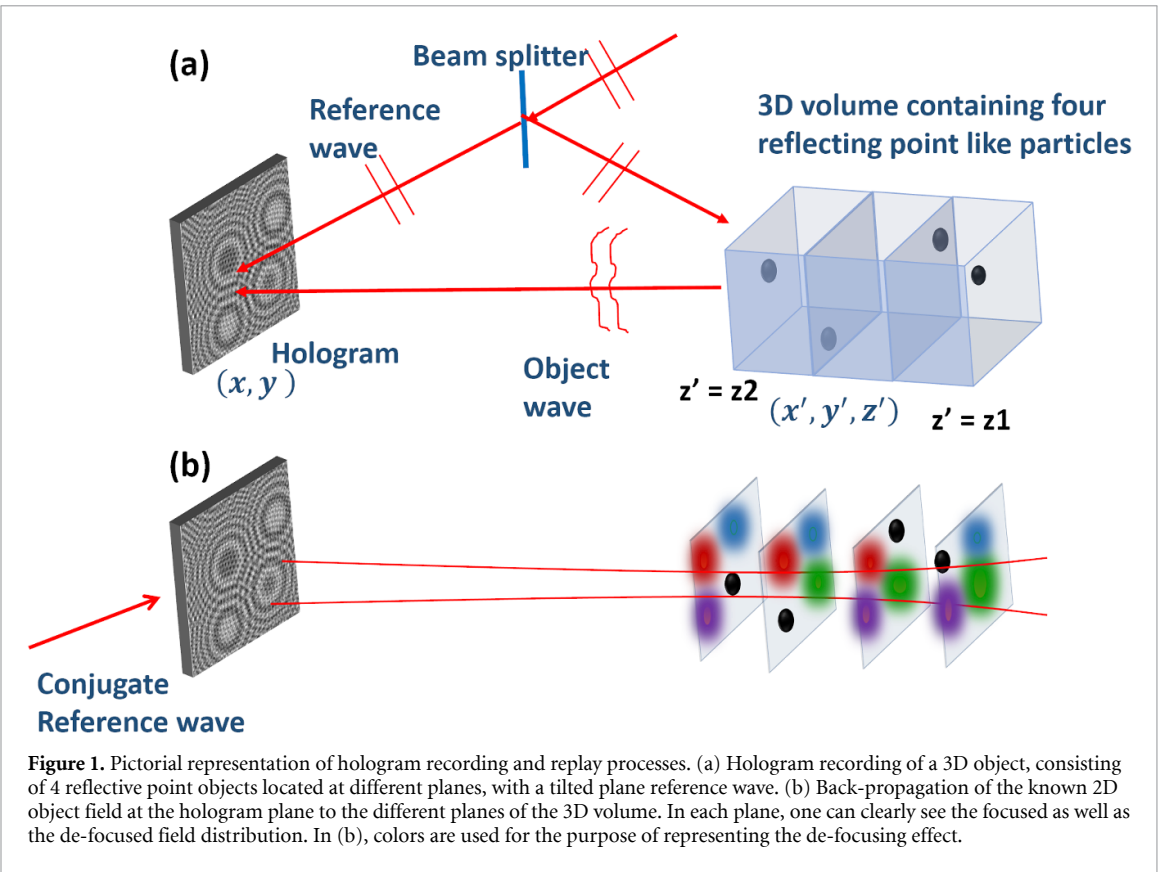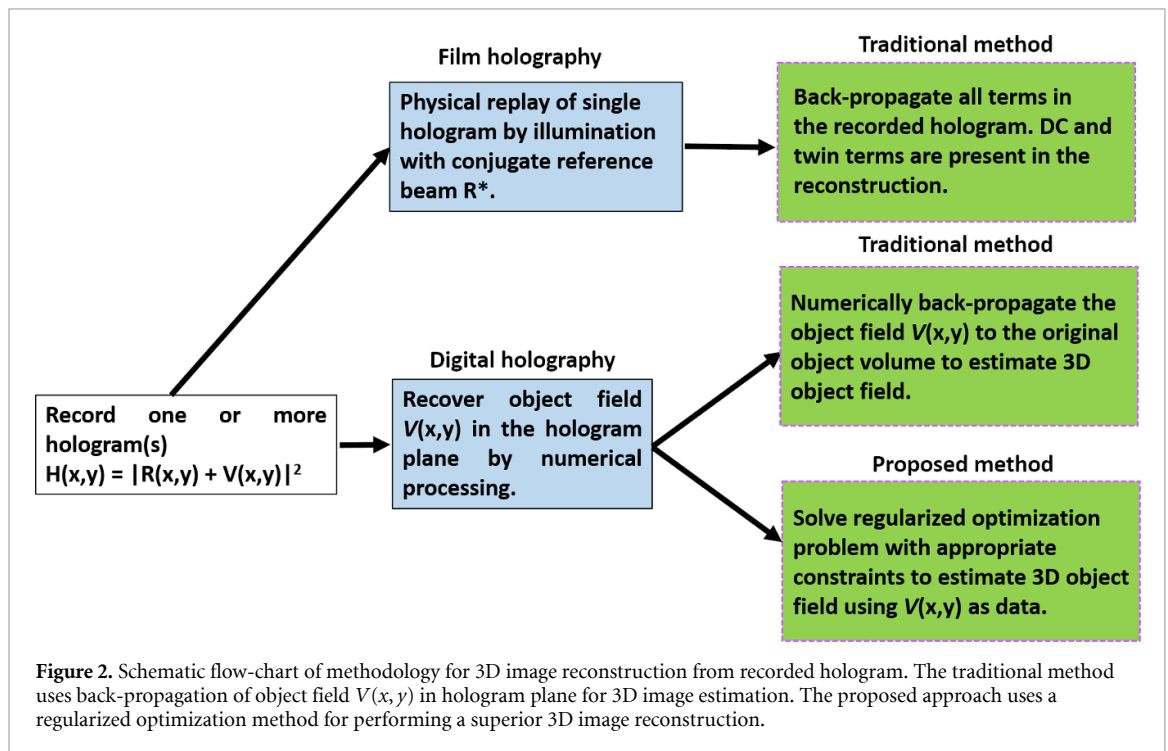
**Figure 1.** Pictorial representation of hologram recording and replay processes. (a) Hologram recording of a 3D object, consisting of 4 reflective point objects located at different planes, with a tilted plane reference wave. (b) Back-propagation of the known 2D object field at the hologram plane to the different planes of the 3D volume. In each plane, one can clearly see the focused as well as the de-focused field distribution. In (b), colors are used for the purpose of representing the de-focusing effect.



**Figure 2.** Schematic flow-chart of methodology for 3D image reconstruction from recorded hologram. The traditional method uses back-propagation of object field $V(x,y)$ in hologram plane for 3D image estimation. The proposed approach uses a regularized optimization method for performing a superior 3D image reconstruction.

recording geometry. The voxels in $U(x',y';z')$ where no object is present are equal to zero. In particular, if the object is a reflective surface represented as $z' = f(x',y')$, then the object distribution may be denoted as:

$$U(x',y';z') = r(x',y') \ \delta[z' - f(x',y')], \tag{1}$$

where $r(x', y')$ denotes the reflection coefficient of the object at location $(x', y', z' = f(x', y'))$. The forward propagation of the 3D object field to the hologram plane can be represented by a linear operator $\hat{A}$:

$$V(x, y) = \hat{A} U(x', y'; z').\tag{2}$$

In accordance with [35] we neglect the secondary scattering, and the operator $\hat{A}$ thus consists of propagating each slice in the function $U(x', y'; z')$ to the detector plane (located at $z = z_d$) and adding the total field to get a complex valued function denoted by $V(x, y)$ as follows:

$$V(x, y) = \int_{z_1}^{z_2} dz' e^{ik(z' - z_2)} \int \int dx' dy' U(x', y'; z') h(x - x', y - y'; z_d - z').\tag{3}$$

The inner integral over $x'$ and $y'$ represents propagation of the slice located at $z = z'$ to the detector plane $z = z_d$. The integral over $z'$ on the other hand represents the sum over all the $z$-slices of the 3D volume of interest. The factor $\exp[ik(z' - z_2)]$ arises in the reflection geometry shown in figure 1(a) as the illuminating beam has different relative phases on different object planes. The limits $z_1$ and $z_2$ on integration over $z'$ denote the $z$-range of the volume of interest. These limits will be omitted in the following discussion for brevity. The impulse response $h(x, y; z)$ in a paraxial geometry may be assumed to be that for Fresnel diffraction [36]:

$$h(x, y; z) = \frac{1}{i\lambda z} \exp(ikz) \exp\left[i\frac{\pi}{\lambda z}(x^2 + y^2)\right].\tag{4}$$

Here $\lambda$ is the wavelength of illumination and $k = 2\pi/\lambda$ is the wave-number. For a more general wide-angle case, the preferred impulse response may be obtained via the angular spectrum approach. Next we proceed to determine the adjoint operation $\hat{A}^\dagger$ corresponding to the forward model in equation (3) by emploing the scalar product relation:

$$< V_1, \hat{A} U_2 >_{2D} = < \hat{A}^\dagger V_1, U_2 >_{3D}.\tag{5}$$

The suffixes 2D and 3D on the scalar products tell us that the scalar product on the LHS is evaluated on the detector coordinates whereas the scalar product on the RHS is evaluated over the 3D volume of interest. In the above equation, the detector field $V_1 = \hat{A} U_1$ and complex-valued functions $U_1, U_2$ represent any two 3D objects. Fully writing out the LHS of the above equation and using equation (3), we have

$$\begin{aligned} &< V_1, \hat{A} U_2 >_{2D} \\ &= \int \int dx dy V_1^*(x, y) \int dz' e^{ik(z' - z_2)} \int \int dx' dy' U_2(x', y'; z') h(x - x', y - y'; z_d - z'). \end{aligned}\tag{6}$$

The $^*$ in the above equation denotes the complex-conjugation that arises out of the usual definition of the scalar product. Rearranging the order of integration we can write the above relation as:
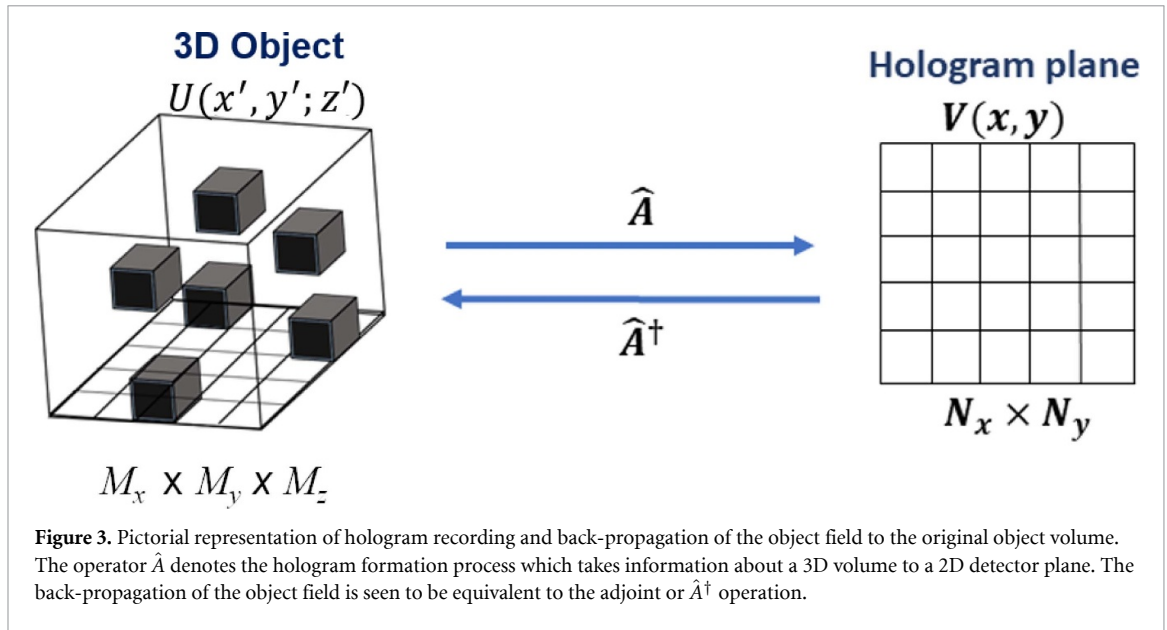
$$\begin{aligned} &< V_1, \hat{A} U_2 >_{2D} \\ &= \int \int \int dz' dx' dy' \left[ \int \int dx dy V_1(x, y) e^{-ik(z' - z_2)} h^*(x - x', y - y'; z_d - z') \right]^* U_2(x', y'; z'). \end{aligned}\tag{7}$$

Using equations (5) and (7), we observe that the adjoint operator $\hat{A}^\dagger$ acting on the detector field $V_1(x, y)$ may be identified as:

$$[\hat{A}^\dagger V_1](x', y'; z') = e^{-ik(z' - z_2)} \int \int dx dy\, V_1(x, y) h^*(x - x', y - y'; z_d - z').\tag{8}$$

Apart from the phase factor $e^{-ik(z' - z_2)}$ which nullifies the initial phase on different object planes due to illumination, we observe that the $\hat{A}^\dagger$ operation simply involves the back-propagation (as denoted by $h^*$) from detector plane $z = z_d$ to all the planes $z = z'$ in the volume of interest. The operations corresponding to the hologram formation and the back-propagation of the object field are shown pictorially in figure 3 for clarity.

The traditional holographic replay of the object field back to the original 3D object volume is thus equivalent to the Hermitian transpose of the forward hologram formation operation denoted by operator $\hat{A}$. In fact since operator $\hat{A}$ takes voxels in 3D object space to a much smaller number of pixels in 2D, the formal inverse operation $\hat{A}^{-1}$ does not exist. Obtaining a true 3D solution which is non-zero only at the voxels where the original object existed may however be possible using a regularized optimization framework where appropriate constraints representing the desirable properties of the object function are modelled. We

**Figure 3.** Pictorial representation of hologram recording and back-propagation of the object field to the original object volume. The operator $\hat{A}$ denotes the hologram formation process which takes information about a 3D volume to a 2D detector plane. The back-propagation of the object field is seen to be equivalent to the adjoint or $\hat{A}^{\dagger}$ operation.

proceed to do this in the next section. It is important to note that such true 3D image reconstruction cannot be generated by any physical replay. Applying diffraction theory to the back-propagating field, we can clearly observe that the replay field cannot be identically zero in some intermediate planes (where the original object did not exist), and suddenly reappear at voxels where the original 3D object existed. In this sense the optimization based solution we construct in the following section takes us away from the traditional picture of physical hologram replay.

## 3. True 3D recovery: methodology

From the previous section, firstly, it is clear that the back-propagated object wave constitutes the $\hat{A}^{\dagger}$ or Hermitian transpose operation. Secondly, due to the nature of the reconstruction problem where number of detector pixels is typically much smaller compared to the number of voxels in the 3D volume of interest, it may not be possible to formally invert the operation $\hat{A}$. The solution may however be feasible in case the object function has sufficient sparsity in the voxel domain or some transform domain. Such constraints if applicable to the object under consideration, will allow us to reconstruct the true 3D solution despite the fact that this is an incomplete data problem. For example, if the object consists of a collection of point sources as in case of particulate holograms, $\ell_1$ norm penalty may be suitable for imposition of solution sparsity. For a variable $\overline{U}$, the $\ell_1$ norm penalty is defined as:

$$\|\overline{U}\|_1 = \sum_{a=1}^{M_x} \sum_{b=1}^{M_y} \sum_{c=1}^{M_z} |\overline{U}_{a,\,b,\,c}|, \tag{9}$$

where indices $(a, b, c)$ run over $x, y$ and $z-$ axes, respectively, and $M = M_x \times M_y \times M_z$ denotes the total number of voxels in the 3D object. On the other hand, if the object is sparse in the image gradient domain, its sparsity can be enforced by the total variation (TV) penalty [37]. For simplicity, we consider a TV penalty for each of the constituting 2D planes, which for every $c^{th}$ plane, denoted by $\overline{U}_{:,\,:,\,c}$, is given by

$$\|\overline{U}_{:,\,:,\,c}\|_{TV} = \sum_{a=1}^{M_x} \sum_{b=1}^{M_y} \sqrt{\left|\overline{U}_{a,\,b,\,c} - \overline{U}_{a-1,\,b,\,c}\right|^2 + \left|\overline{U}_{a,\,b,\,c} - \overline{U}_{a,\,b-1,\,c}\right|^2}. \tag{10}$$

More general forms of gradient based penalties such as Huber penalty [7] may also be applicable. Depending upon the object of interest, the gradient based penalty may also be applied in all three directions. In addition to the sparsity constraint, the operation $\hat{A}$ applied to the object $\overline{U}$ is expected to be consistent with the given object field at the hologram plane. Keeping this in mind, we propose to solve the following minimization problem to get an estimate $\overline{U}$ of the 3D field:

$$\underset{\overline{U}}{\text{minimize}} \; C_1(\overline{U}, \overline{U}^*) + \alpha \, C_2(\overline{U}, \overline{U}^*). \tag{11}$$

The first term $C_1$ is given by

$$C_1(\overline{U}, \overline{U}^*) = \frac{1}{2} \|V - \hat{A}\,\overline{U}\|_F^2 \tag{12}$$

that ensures the data consistency of the estimated 3D field. As explained before, $V$ is the data representing the 2D object field at the hologram plane (as may be derived from the recorded hologram), $\hat{A}$ is the operator defined in equation (3), and $\|.\|_F$ denotes the Frobenius norm of its argument. The second term $C_2$ is the chosen sparsity regularization term. The real valued quantity $\alpha > 0$ is the regularization parameter that determines the relative weight between the two terms of the cost function.

In problem (11), the term $C_1$ is Lipschitz-differentiable, whereas $C_2$, whether chosen to be $\ell_1$ norm or TV penalty, is non-differentiable. In the optimization literature, this type of problem has been dealt with in an iterative manner. In each iteration, a gradient descent step is applied on $C_1$, followed by a proximity step on $C_2$, which is referred to as the forward-backward structure. Based on this, we solve problem in equation (11) using an accelerated optimization algorithm encompassing forward-backward like iterations, namely fast iterative shrinkage-thresholding algorithm (FISTA) [38]. The main steps involved in this algorithm are described in the following. We initialize $\overline{U}$ with the guess solution $\overline{U}^{(0)}$ and consider $Z^{(1)} = \overline{U}^{(0)}$. At $n^{\text{th}}$ iteration, gradient descent scheme consists of updating the solution by going in the descent direction obtained by the gradient of the first term $C_1$ of the cost function to be minimized, i.e.

$$\widetilde{U}^{(n)} = Z^{(n)} - \tau [\nabla C_1]_{Z^{(n)}}, \tag{13}$$

where $-[\nabla C_1]_{Z^{(n)}} = -\hat{A}^\dagger (\hat{A} Z^{(n)} - V)$ corresponds to the descent term evaluated at $Z^{(n)}$. Note that the forward operator $\hat{A}$ and its Hermitian transpose $\hat{A}^\dagger$ as derived in section 2 are explicitly present in this expression. The step size $\tau$ is chosen as $\tau = 1/\kappa$, where $\kappa = \|\hat{A}^\dagger \hat{A}\|_S$ with $\|.\|_S$ denoting the spectral norm of its argument. With this choice, $\kappa$ is essentially the Lipschitz constant associated with the Lipschitz-differentiable function $C_1$ and it may be evaluated using the power iteration method. Once the gradient-updated solution $\widetilde{U}^{(n)}$ is available, the proximity operator of the second term in problem (11) needs to be evaluated. By definition, proximity operator of $\alpha C_2$ at $\widetilde{U}^{(n)}$ is given by
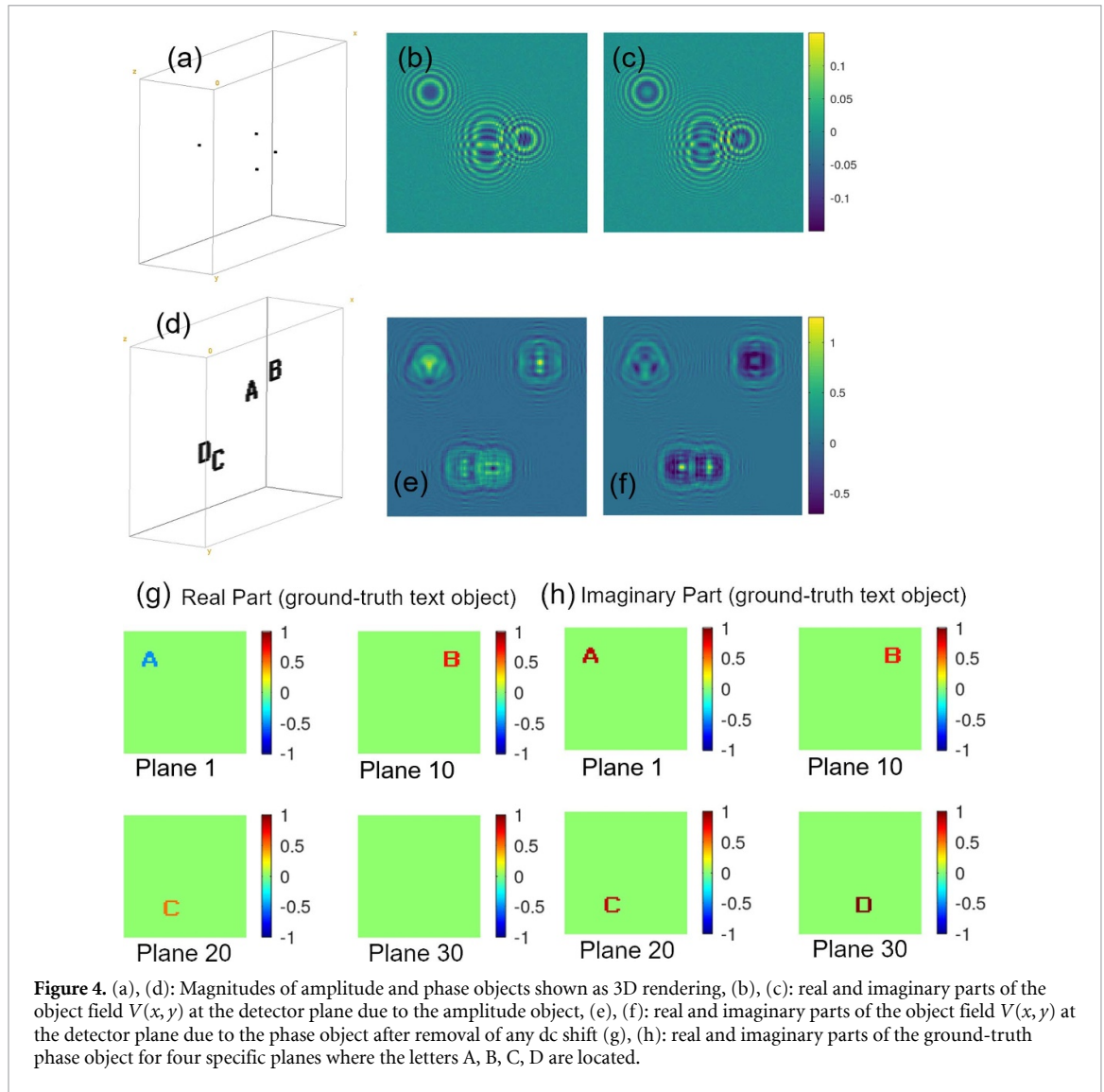
$$\text{prox}_{\alpha C_2}(\widetilde{U}^{(n)}) = \underset{L}{\arg\min}\; \alpha C_2(L) + \frac{1}{2} \|L - \widetilde{U}^{(n)}\|_F^2. \tag{14}$$

Initially introduced by Moreau [39], this operator is widely used to tackle non-smooth functions in a variety of signal and image processing problems. To get an idea behind this operator, note that the quadratic term in equation (14) restricts the solution from going too far from $\widetilde{U}^{(n)}$ while minimizing for $C_2$ as well. Thus, the gradient step together with the proximity step aims to find the minimizer of problem of the form (11). As previously discussed, the chosen prior for a sparse object consisting of point sources is the $\ell_1$ norm penalty. In addition, if the object is an amplitude object, the reality and positivity constraints may be imposed. This can be enforced by the use of indicator function $1_{\mathbb{R}_+^N}(.)$ of real and positive orthant. Thus, the regularization term in this case is given by: $C_2(.) = \|.\|_1 + 1_{\mathbb{R}_+^N}(.)$. The associated proximity operator reduces to the positive soft-thresholding operator, $\Gamma$ [40]. In particular, we set $\Gamma_\mu(\widetilde{U}^{(n)}) = \overline{U}^{(n+1)}$ such that for all $(a, b, c) \in (M_x, M_y, M_z)$,

$$[\overline{U}^{(n+1)}]_{a,b,c} = \begin{cases} \{\text{Re}([\widetilde{U}^{(n)}]_{a,b,c}) - \mu & \text{for } \text{Re}([\widetilde{U}^{(n)}]_{a,b,c}) \geq \mu \\ 0 & \text{otherwise}, \end{cases} \tag{15}$$

where $\text{Re}(.)$ denotes the real part operator and $\mu = \alpha \tau$ in the considered case. Intuitively, the positive soft-thresholding operator sets all values less than the thresholding parameter $\mu$ to zero, while shrinking the rest of the values by an amount $\mu$. As a result, iteration-by-iteration, this operation leads to promoting sparsity and positivity of its argument. For the second considered case in our numerical experiments, $C_2(.) = (.)_{TV}$, and the proximity operator does not have a closed form solution. In fact, it needs to be computed using sub-iterations, for which we use the fast gradient projection (FGP) method proposed in [41]. Once we get an updated $\overline{U}^{(n+1)}$, the peculiarity of FISTA is that in every iteration, the gradient scheme is employed on a linear combination of last two updated values $\overline{U}^{(n+1)}$ and $\overline{U}^{(n)}$, for faster convergence. This step is shown as follows:

$$Z^{(n+1)} = \overline{U}^{(n+1)} + \frac{t^{(n)} - 1}{t^{(n+1)}} \left( \overline{U}^{(n+1)} - \overline{U}^{(n)} \right), \tag{16}$$

**Figure 4.** (a), (d): Magnitudes of amplitude and phase objects shown as 3D rendering, (b), (c): real and imaginary parts of the object field $V(x, y)$ at the detector plane due to the amplitude object, (e), (f): real and imaginary parts of the object field $V(x, y)$ at the detector plane due to the phase object after removal of any dc shift (g), (h): real and imaginary parts of the ground-truth phase object for four specific planes where the letters A, B, C, D are located.

where

$$t^{(n+1)} = \frac{1 + \sqrt{1 + 4\left(t^{(n)}\right)^2}}{2}, \tag{17}$$

and $t^{(1)} = 1$. The updated $Z^{(n+1)}$ is then used in the next iteration to compute gradient step (13) and the process is repeated until convergence.

## 4. Simulations and results

In simulations, we have considered a 3D computational box containing $128 \times 128 \times 30$ voxels which has equivalent physical dimension of $128\,\mu\mathrm{m} \times 128\,\mu\mathrm{m} \times 60\,\mu\mathrm{m}$, and the illumination wavelength is assumed to be $\lambda = 0.5\,\mu\mathrm{m}$. The distance between the detector and the center of the 3D object volume is set to be $|(z_d - (z_1 + z_2)/2)| = 110.85\,\mu\mathrm{m}$. The detector is assumed to be of size $128 \times 128$ pixels with pixel size of $1\,\mu\mathrm{m}$. It may be noted that the sensor pixel sizes in commercially available CMOS sensors (particularly those in mobile phone cameras) are already close to $1\,\mu\mathrm{m}$. Further if the proposed 3D reconstruction methodology is to be applied to small objects like individual biological cells whose hologram is recorded with a slight de-focus using a digital holographic microscope, it is conceivable that a small detector region-of-interest as used here may be practically sufficient to be used as data space. The numerical aperture of the detector at the center of the object volume is thus given by:

$$NA = \sin\left[\arctan\left(\frac{64 \times 1\,\mu\mathrm{m}}{110.85\,\mu\mathrm{m}}\right)\right] = 0.5. \tag{18}$$

Note that the voxel size in the 3D object volume has been kept equal to $\lambda/(NA)$ in the lateral direction and $\lambda/(NA)^2$ in the depth direction. Each voxel in the 3D object volume thus has dimensions nominally equal to the diffraction-limited resolution allowed by the geometry of the setup. For the illustration we take two 3D objects: the first is a reflective amplitude object consisting of four small square reflectors located at four different planes (3, 11, 20, 28) of the volume, as shown in figure 4(a). Each square object on a particular z-plane occupies a $2 \times 2$ pixel region. The second object is a transmissive phase object with four letters A, B, C, D placed in four different planes (1, 10, 20, 30), as shown in figure 4(d). The phase values for the four objects A, B, C, D are $2\pi/3$, $\pi/4$, $\pi/3$ and $\pi/2$ respectively and they all have unit amplitude. While the forward model[1] in (2) and (3) is described for a reflection geometry, the main difference for transmission geometry is that the detector field $V(x, y)$ now contains an additional illumination field term. Assuming that the illumination in transmission geometry is a plane wave, this will mean that the detector field will now have an additional complex-valued constant term [42]. Prior to using the $V(x, y)$ in the reconstruction algorithm we therefore subtract any constant dc offset from it. The real and imaginary parts of the object field $V(x, y)$ corresponding to the two objects as obtained by application of the operator $\hat{A}$ (equation (2)) are shown in figure 4(b)-(c) and (e)-(f) respectively. In figure 4(g)-(h) we show the real and imaginary parts of the ground-truth text phase object in four specific planes where the letters A, B, C, D are located. We want to emphasize here that we are not considering the case of highly scattering objects (e.g. turbid tissues) in the transmission geometry as the Born approximation will not be valid for them. Simpler objects like unstained biological cells have high light transmission and the object wave signal is not necessarily weak. In a laboratory experiment, the object field $V(x, y)$ is not directly measured, but as explained before it is recoverable from the recorded interference pattern $H(x, y) = |R(x, y) + V(x, y)|^2$ where $R(x, y)$ denotes the reference beam. For a low-noise cooled array sensor the accuracy in the measurement of $H(x, y)$ is ultimately limited by photon noise which is a Poisson random process. It is known that the numerical recovery of object wave $V(x, y)$ is at least as accurate as permitted by shot-noise [11, 13, 43]. If the hologram is recorded with an average light level of $N_0$ photons per pixel, the relative shot-noise is of the order of $1/\sqrt{N_0}$ which will be distributed equally in real and imaginary parts of $V(x, y)$. The average number of photons per pixel for noise level estimation was assumed to be $N_0 = 3 \times 10^4$ which is approximately half of the full-well capacity of typical cooled scientific CMOS array sensors available commercially. For simplicity we added independent Gaussian random noise realizations, with standard deviation of $1/\sqrt{2}$ times the shot noise level corresponding to the above $N_0$, to real and imaginary parts of the computed $V(x, y)$. The noise level plays an important role in selecting the regularization constant $\alpha$ in equation (11) as explained below.

Given the noisy object field $V(x, y)$ at the detector plane, we solve the optimization problem in equation (11) as explained in section 3 to get an estimate $\overline{U}(x, y; z)$ of the true 3D object. For the amplitude object in this illustration, we used $\ell_1$ penalty with reality and positivity constraints. On the other hand for the text phase object, we used the total variation penalty. The regularization parameter $\alpha$ is chosen empirically to be $1.5 \times 10^{-2}$ for the amplitude object case and $5 \times 10^{-3}$ for the text phase object case. The parameter $\alpha$ has been selected such that on convergence, the L2-norm of the data-domain error $||V - \hat{A}\overline{U}||_2$ is approximately equal to the L2-norm of the random noise realization added to $V(x, y)$. Smaller values of $\alpha$ will lead to over-fitting the noise and larger values of $\alpha$ will result in over regularization of the solution. This methodology for selection of the regularization parameter is known as Morozov's discrepancy principle [44]. Automated tuning of $\alpha$ may be possible (e.g. see [13]), however, this topic is not central to the main theme of the present work and will be explored in future.

The 3D recovery results for the amplitude object are shown in figure 5 using both the simple back-propagation of the object field and the proposed iterative approach. Figures 5 (a),(b) show the 3D rendering of the absolute magnitude of these two solutions. Figures 5 (c),(d) show reconstruction of the slices where the reflective objects existed. In order to assess the lateral and depth resolution of the proposed method, in figure 5(e), (f) we have drawn line profiles along the dotted lines shown in figure 5(b). The dotted lines are parallel to z-axis and x-axis respectively. From these line profile plots we observe that the reconstructed numerical values are negligible at voxels where the original reflective object did not exist. Since the lateral and depth dimensions of the voxels in the 3D volume of interest were selected based on the numerical aperture ($\lambda/(NA)$ and $\lambda/(NA)^2$ respectively), the proposed iterative method may be considered to yield 3D reconstruction with diffraction-limited resolution. Similarly for the text phase object, we show 3D rendering of magnitude of the solutions recovered using simple back-propagation and the iterative approach in figures 6(a) and (b) respectively. The real and imaginary parts of the recoveries in the planes where the original letters A, B, C, D were located are shown in figures 6(b) and (d). The 3D renderings shown in all the

---

[1] the wording 'forward model' is typically used in the context of inverse problems, which here refers to back-scattering for reflection geometry and forward scattering for transmission geometry
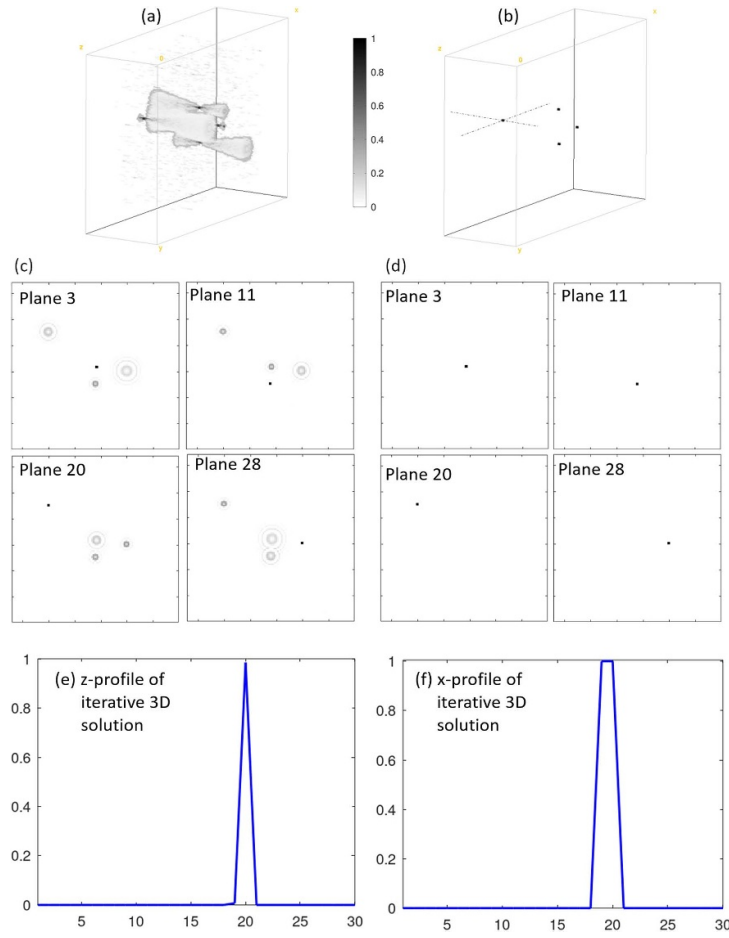
**Figure 5.** 3D reconstruction of the amplitude object with four small reflectors. (a), (b): 3D rendering of the amplitude of the recovered four-reflector object obtained by simple back-projection of the object field and by using the iterative reconstruction respectively. (c), (d): Image of the planes where the point objects originally existed in the 3D object box shown in (a) and (b) respectively. (e), (f): Line profiles along dotted lines (parallel to *z*-axis and *x*-axis) as shown in (b) for one of the reflective object. In (e) the horizontal axis labels refer to voxel indices along *z*-axis (1 unit = 2 $\mu$m). In (f) the horizontal axis labels refer to the voxel indices along *x*-axis (1 unit = 1 $\mu$m). Note that the peak in plot (f) appears flat since the reflective object is 2x2 pixel wide in transverse (x-y) direction. The quantity plotted on the vertical axis is real part of the solution $\overline{U}$, along z- and x-dimensions for (e) and (f) respectively.

illustrations were generated using the Volume Viewer plug-in in the open-source software ImageJ. On visual inspection it may be observed that the real and imaginary parts of the recovered text object using the proposed iterative formalism closely match with those of the ground truth 3D objects shown in figures 4(g) and (h). The simple back-projection reconstructions over the original object volume as shown in figure 5(a), (c) and figures 6(a) and (c) however show non-zero de-focused fields in places where no object originally existed. In a MATLAB/Octave implementation of the code, both illustrations used 2000 iterations of the proposed solution. The relative reconstruction error in the recovered 3D solutions with respect to the ground truth objects (as shown in figure 4) is given by:

$$E^{(\text{object domain})} = \frac{||U_{(\text{ground truth})} - \overline{U}||_2}{||U_{(\text{ground truth})}||_2}, \tag{19}$$

and is plotted for the two cases in figure 7(a),(b) as a function of iteration number. The relative error for the amplitude object was down to approximately 0.023 and for the text phase object was seen to be close to 0.095. The relative 3D object domain errors for the simple back-projection reconstructions shown in figure 5(a), (c) and figure 6(a), (c) are however quite large and are equal to 4.32 and 5.39 for the amplitude and phase objects respectively. This large error for simple back-projection solution is expected since this solution is non-zero even in the slices of the 3D volume where no object originally existed.

The true 3D reconstruction using a single object field data frame $V(x, y)$ in the hologram plane shown here is feasible since both the 3D objects were sparse. Further, as long as the secondary scattering can be assumed to be small and neglected, the overlapping of diffraction patterns for the test objects as shown in
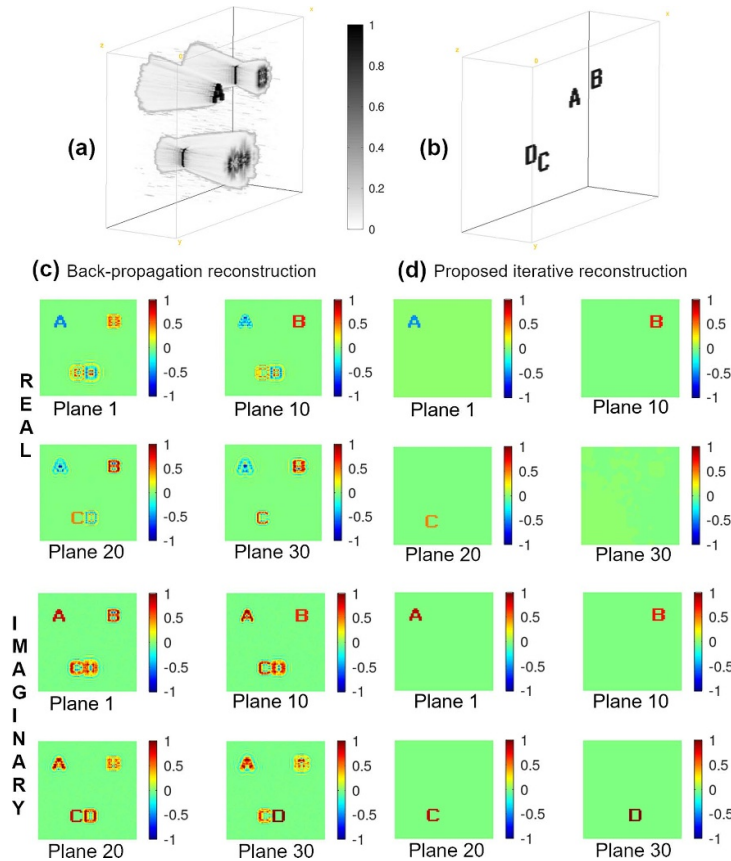
**Figure 6.** 3D reconstruction of the text phase object. (a),(b): 3D rendering of amplitude of the reconstruction using simple back-projection and using proposed iterative reconstruction. (c),(d): Real and imaginary parts for four different planes where the letters were located in the true 3D object box are shown for the reconstructions in (a), (b) respectively.
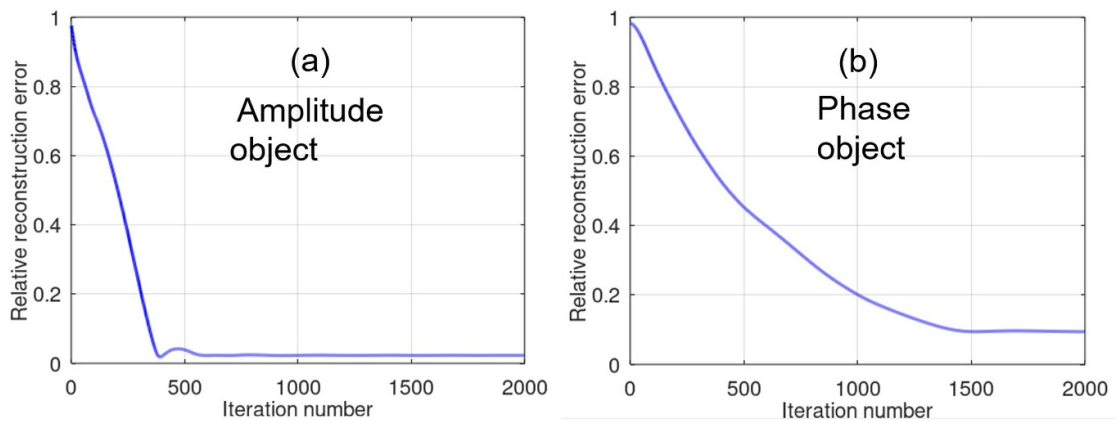


**Figure 7.** (a), (b): Relative reconstruction error for the two illustrations shown in figure 5 and figure 6 as a function of iteration number respectively. The relative error is computed between the reconstructed 3D solution $\overline{U}(x,y,z)$ and the ground truth 3D objects depicted in figure 4.

figure 3(b), (c), (e), (f) does not lead to undesirable artefacts as our reconstruction results show. Such 3D reconstructions may not be possible for non-sparse objects from degrees-of-freedom considerations. In such cases, additional diversity mechanisms may be required for measuring non-redundant data. The question of whether true 3D reconstruction of a general object is possible is however not just dependent on degrees-of-freedom considerations - the assumption about small secondary scattering that has been used here may not be valid as the 3D object gets more and more complex in its structure. Further study is required before making any quantitative statement regarding possibility of reconstructing a 3D object of given complexity. Nevertheless the iterative approach presented here may be valid for 3D reconstruction of low complexity objects such as transparent cells or optical elements like lenses, etc from their 2D

holograms/interferograms. We will report on experiments using the proposed methodology for such simple objects in future.

# 5. Conclusion

In conclusion, we have examined the nature of 3D image reconstruction as obtained by the replay process in holography. The traditional holographic replay was identified to be the Hermitian transpose operation corresponding to the forward hologram formation process. Using this knowledge, we formulated an iterative solution for the 3D reconstruction problem which used object sparsity as a constraint. Note that in this iterative approach, we do not search for a best focus plane associated with individual objects. The best focus is instead determined implicitly by the sparsity constraint. Our methodology of using the numerically reconstructed object field in the hologram plane as the starting point for the inverse problem ensures that this approach automatically eliminates any effect of dc or twin image terms. The resultant 3D reconstruction is seen to be a significantly better representation of the original 3D object as compared to the simple back-projection based image reconstruction. The 3D reconstruction using the constrained optimization strategy as described here cannot be achieved by any form of physical holographic replay, since the back-propagating waves from a hologram cannot suddenly vanish over some intermediate z-planes and reappear again at locations where the original 3D object existed. The methodology therefore truly takes full advantage of the fact that complex object wave at the hologram plane is available in numerical form in digital holography. The technique may find important applications in a number of digital holographic or interferometric imaging systems as we will explore in future.

# Acknowledgments

# ORCID iD

Kedar Khare https://orcid.org/0000-0002-3104-3850

# References

[1] Gabor D 1948 A new microscopic principle *Nature* **161** 777–8
[2] Hecht E 2002 *Optics* (Boston, MA: Addison Wesley)
[3] Schnars U, Falldorf C, Watson J and Jüptner W 2015 Digital holography *Digital Holography and Wavefront Sensing* (Berlin: Springer) pp 39–68
[4] Kim M K, Yu L and Mann C J 2006 Interference techniques in digital holography *J. Opt. A: Pure Appl. Opt.* **8** S518
[5] Kim M K 2010 Principles and techniques of digital holographic microscopy *SPIE Rev.* **1** 018005
[6] Depeursinge C and Propescu G 2018 Quantitative phase imaging in biomedicine *Nat. Photonics* **12** 578
[7] Mangal J, Monga R, Mathur S R, Dinda A K, Joseph J, Ahlawat S and Khare K 2019 Unsupervised organization of cervical cells using bright-field and single-shot digital holographic microscopy *J. Biophotonics* **12** e201800409
[8] Takeda M, Ina H and Kobayashi S 1982 Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry *J. Opt. Soc. Am.* A **72** 156–60
[9] Creath K 1988 Phase-measurement interferometry techniques *Prog. Opt.* **26** 349–93
[10] Yamaguchi I and Zhang T 1997 Phase-shifting digital holography *Opt. Lett.* **22** 1268–70
[11] Singh M, Khare K, Jha A K, Prabhakar S and Singh R P 2015 Accurate multipixel phase measurement with classical light interferometry *Phys. Rev.* A **91** 021802
[12] Singh M and Khare K 2017 Single-shot interferogram analysis for accurate reconstruction of step phase objects *J. Opt. Soc. Am.* A **34** 349–55
[13] Rajora S, Butola M and Khare K 2019 Mean gradient descent: an optimization approach for single-shot interferogram analysis *J. Opt. Soc. Am.* A **36** D7–D13
[14] Gabor D, Kock W E and Stroke G W 1971 Holography *Science* **173** 11–23
[15] Hong J H, McMichael I C, Chang T Y, Christian W R and Paek E G 1995 Volume holographic memory systems: techniques and architectures *Opt. Eng.* **34** 2193–204
[16] Li H, Guo C, Muniraj I, Schroeder B C, Sheridan J T and Jia S 2017 Volumetric lightfield encryption at the microscopic scale *Sci. Rep.* **7** 40113
[17] Rosen J and Brooker G 2007 Digital spatially incoherent fresnel holography *Opt. Lett.* **32** 912–14
[18] Park J-H, Hong K and Lee B 2009 Recent progress in three-dimensional information processing based on integral imaging *Appl. Opt.* **48** H77–H94
[19] Xiao X, Javidi B, Martinez-Corral M and Stern A 2013 Advances in three-dimensional integral imaging: sensing, display and applications *Appl. Opt.* **52** 546–60
[20] Ihrke I, Restrepo J and Mignard-Debise L 2016 Principles of light field imaging: Briefly revisiting 25 years of research *IEEE Sig. Process. Mag.* **33** 59–69

[21] Malacara D 2007 *Optical Shop Testing* vol 59 (New York: Wiley)

[22] Adams M, Kreis T M and Jueptner W P O 1997 Particle size and position measurement with digital holography *Optical Inspection and Micromeasurements II* ed C Gorecki vol **3098** (SPIE) pp 234–40

[23] Fung J, Perry R W, Dimiduk T G and Manoharan V N 2012 Imaging multiple colloidal particles by fitting electromagnetic scattering solutions to digital holograms *J. Quant. Spectrosc. Radiat. Transfer* **113** 2482–9

[24] Soulez F, Denis L, Fournier C, Thiébaut É and Goepfert C. 2007 Inverse-problem approach for particle digital holography: accurate location based on local optimization *J. Opt. Soc. Am.* A **24** 1164–71

[25] Brady D J, Choi K, Marks D L, Horisaki R and Lim S 2009 Compressive holography *Opt. Express* **17** 13040–9

[26] Denis L, Lorenz D, Thiébaut E, Fournier C and Trede D 2009 Inline hologram reconstruction with sparsity constraints *Opt. Lett.* **34** 3475–7

[27] Mallery K and Hong J 2019 Regularized inverse holographic volume reconstruction for 3d particle tracking *Opt. Express* **27** 18069–84

[28] Latychevskaia T, Gehri F and Fink H-W 2010 Depth-resolved holographic reconstructions by three-dimensional deconvolution *Opt. Express* **18** 22527–44

[29] Shimobaba T, Takahashi T, Yamamoto Y, Endo Y, Shiraki A, Nishitsuji T, Hoshikawa N, Kakue T and Ito T 2019 Digital holographic particle volume reconstruction using a deep neural network *Appl. Opt.* **58** 1900–6

[30] Latychevskaia T 2019 Iterative phase retrieval for digital holography *J. Opt. Soc. Am.* A **36** 31–40

[31] Wolf E 1969 Three-dimensional structure determination of semi-transparent objects from holographic data *Opt. Commun.* **1** 153–6

[32] Lin Y, Chen H, Tu H, Liu C and Cheng C-J 2017 Optically driven full-angle sample rotation for tomographic imaging in digital holographic microscopy *Opt. Lett.* **42** 1321–4

[33] Choi W, Fang-Yen C, Badizadegan K, Oh S, Lue N, Dasari R R and Feld M S 2007 Tomographic phase microscopy *Nat. Methods* **4** 717–19

[34] Zhang W, Cao L, Brady D J, Zhang H, Cang J, Zhang H and Jin G 2018 Twin-image-free holography: a compressive sensing approach *Phys. Rev. Lett.* **121** 093902

[35] Gabor D 1956 Theory of electron interference experiments *Rev. Mod. Phys.* **28** 260

[36] Goodman J W 2005 *Introduction to Fourier Optics* (Roberts and Company Publishers)

[37] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica* D **60** 259–68

[38] Beck A and Teboulle M 2009 A fast iterative shrinkage-thresholding algorithm for linear inverse problems *SIAM J. Imag. Sci.* **2** 183–202

[39] Moreau J-J 1965 Proximité et dualité dans un espace hilbertien *Bull. Soc. Math. Fr.* **93** 273–99

[40] Combettes P L and Pesquet J-C 2011 Proximal Splitting Methods in Signal Processing *Fixed-Point Algorithms for Inverse Problems in Science and Engineering* (Berlin: Springer) pp 185–212

[41] Beck A and Teboulle M 2009 Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems *IEEE Trans. Image Proc.* **18** 2419–34

[42] Carter W H 1970 Computational reconstruction of scattering objects from holograms *J. Opt. Soc. Am.* **60** 306–14

[43] Charriére F, Rappaz B, Kühn J, Colomb T, Marquet P and Depeursinge C 2007 Influence of shot noise on phase measurement accuracy in digital holographic microscopy *Opt. Express* **15** 8818–31

[44] Morozov V A 1966 On the solution of functional equations by the method of regularization *Dokl. Akad. Nauk.* **167** 510–12