∂H: Differentiable Holography

Ni Chen[1,*] *Congli Wang*[2] *Wolfgang Heidrich*[3]

Dr. N. Chen, Wyant College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA

Dr. C. Wang, Department of Electrical Engineering & Computer Sciences, University of California, Berkeley, CA 94720, USA

Prof. W. Heidrich, Visual Computing Center, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia

Email Address: nichen@arizona.edu

Keywords: *Holography, differentiable imaging, inverse problem*

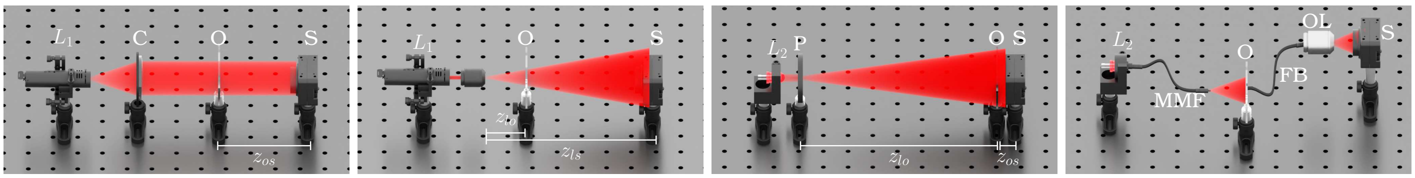
Over the past decade, the field of holography has gained significant ground due to advances in computational imaging. However, the utilization of computational tools is hampered by the mismatch between experimental setups and the conceptual model. We present differentiable holography (∂**H**), a novel framework for automatically self-calibrating experimental imperfections in inverse holographic imaging. The technique is demonstrated on auto-focused complex field imaging from a single intensity-only inline hologram.

1 Introduction

Imaging objects of various sizes ranging from nanometers to millimeters by investigating the object-light interaction is of great importance in numerous research fields, from physics [1], material science [2] to biology [3]. Direct quantification of light waves is challenging when the waves oscillate beyond the speed of electronic devices, while holography resolves this issue by using interferometry [4]. However, holographic imaging has been plagued by a number of challenges that have limited its adoption for the past century, including the unwanted terms (DC and twin-image) separation [5, 6], phase unwrapping, and auto-focusing [7, 8]. To mitigate these challenges, computational techniques surpass multiple variants of holography [9, 10, 11] in the possibility of maintaining a compact inline setup by employing optimization [5, 12] or deep learning [13, 14]. Suppose that imaging systems encode the target object x into holograms y with a forward model of $f(\cdot)$, computational techniques aim at decoding the holograms to obtain the target by inverting the forward model to obtain x with $f^{-1}(\mathbf{y}) \rightarrow \mathbf{x}'$.

Hand-crafted optimization methods rely on relatively simple and ideal imaging models $f(\cdot)$ [5, 12, 6], which may not accurately represent reality. The simplicity is mainly due to the need for hand-crafted optimization solvers. On the other hand, learning-based methods utilize general-purpose but opaque black-box models $f(\cdot)$, requiring a large amount of training data [15, 13] or time-consuming training [16, 17, 14], and often lacking generalization [15, 13]. Both approaches are sensitive to system-specific factors and unable to compensate for mismatches between numerical modeling and real-world conditions [18], such as experimental imperfections (e.g., light sources, optical alignments, sensor quantization) and physical interference. To solve the inverse problem, accurate modeling of these system imperfections is crucial; otherwise, the quality of reconstruction will be compromised due to the accumulation of modeling errors [19, 18]. However, as the complexity of modeling increases, the forward model may not have a simple form, making it difficult to solve using existing solvers. This challenge is further compounded when additional modeling parameters are introduced, further complicating the problem-solving process. Inspired by the principles of differentiable imaging [20], our proposed method, referred to as ∂H , incorporates system imperfections into the imaging model and employs a differentiable optimization technique to overcome the challenging inversion of the forward model. Specifically, we consider defocusing and illumination amplitude variation as additional variables to the model. This approach enables complex field imaging from a single-shot inline hologram without the need for additional hardware. The differentiable design of the imaging framework ensures flexibility and robustness, allowing it to be applied to various setups with minimal adjustments to the forward model.

2 Experimental results



(a) I: Plane illumination (b) II: Spherical illumination (c) III: Lensless holography (d) IV: with fiber bundle1 [21]

Figure 1. Various inline holography systems that are used for the verification: Inline holography with (a) plane wave illumination and (b) spherical wave illumination, (c) lensless holography, and (d) lensless holography with fiber bundle setup. L_1: laser light source; L_2: LED light source; C: collimator; O: object; P: Pinhole; S: camera sensor; MMF: Multi-mode fiber; FB: fiber bundle; OL: objective lens.

The ∂H has been verified on four different setups, as shown in **Figure 1**. Each of the setups induces inverse imaging with varying degrees of complexity. In type I of Figure 1(a), a plane wave laser beam is used as the illumination, and the primary factor that makes the forward model inaccurate is the object-camera distance and the perfection of the plane wave; in type II of Figure 1(b), there are two distances which define the spherical wave illumination, the light source to object distance z_{lo} and the light source to camera sensor distance z_{ls} in the forward model; in type III of Figure 1(c), the object-camera distance should be extremely small to maintain a coherence of the LED light source, which makes it more challenging to separate the twin-image and other terms in the hologram reconstructions; in type IV of Figure 1(d), apart from the difficulties in III, the fiber bundle also induces additional noise and errors in the holograms that hinder the imaging. The performance of the ∂H on these configurations is demonstrated in the following experiments.

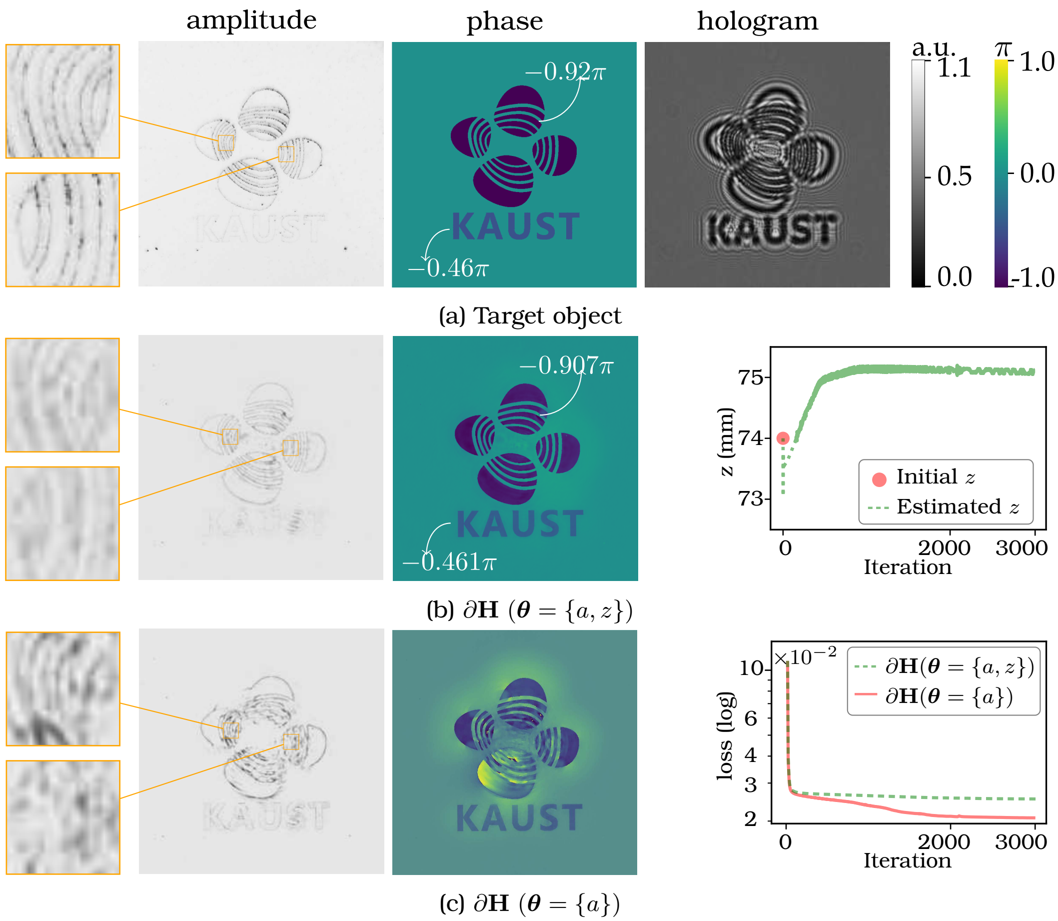
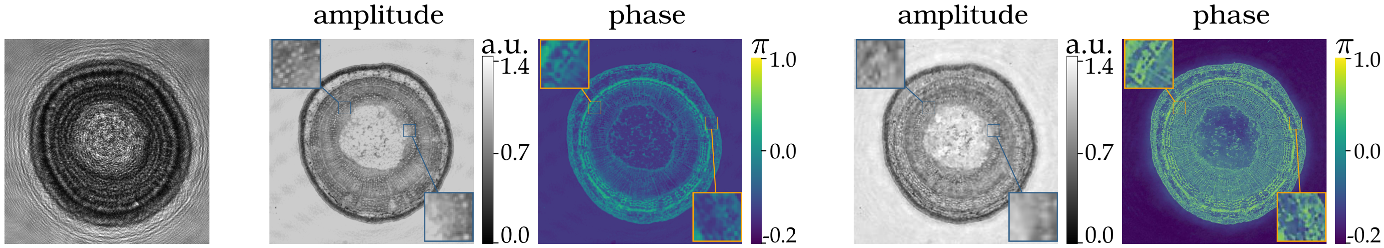


Figure 2. Experiment 1: verification of $\partial\mathbf{H}$ with the setup of Figure 1(a), on a phase-only sample (a) and the reconstructions with (b) and without system parameters (c). See section 4.3 for the definition of $\boldsymbol{\theta}$. z is the object to camera distance and a is the magnitude of the illumination light.

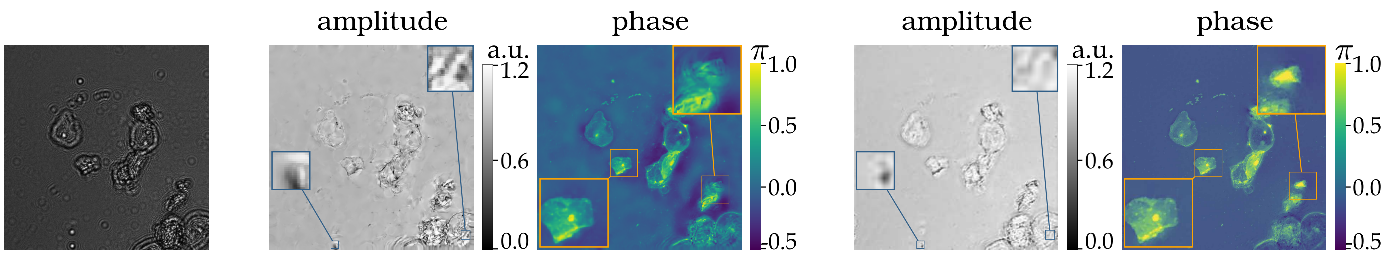
In the first experiment, we verified ∂H on an object with ground truth captured by setup I in Figure 1(a) with a wavelength of the laser centered at $532\,\mathrm{nm}$. The object was a highly transmitting silica wafer with an etched phase profile that is shown by the middle image of **Figure 2**(a). The amplitude photo was taken in the bright field imaging condition, revealing the target's amplitude profile, as the result of absorption at the edges of the etched parts. The ground truth phase image was calculated from the fabricated sample using the experimental parameters. An inline hologram (right image of Figure 2(a)) was captured by an image sensor of $3.45\,\mu\mathrm{m}$ pixel pitch. The reconstructions of ∂H with and without imperfection modeling parameters (here, defocus represented by $\boldsymbol{\theta}$ show a significant difference. With $\boldsymbol{\theta}$, the phase target was well reconstructed, as shown in Figure 2(b). The phase value also approximates the theory one. The right image of Figure 2(b) shows the convergence of the axial location z of the target, which is usually obtained in advance by auto-focusing in the conventional inverse imaging solver. On the contrary, without $\boldsymbol{\theta}$, the reconstructed images blur severely, as shown in Figure 2(c). The comparison of the loss in the right image of Figure 2(c), shows that with $\boldsymbol{\theta}$ included, the loss also converges much better than without it. This is because the forward model is essentially inaccurate without $\boldsymbol{\theta}$, which causes the convergence to stagnate. Because it has a high transmittance, the sample used in Figure 2 is typically treated as a phase-only target. However, in practice, other than the restricted options of phase-only substances, the misalignment of the system may cause light absorption in some portions of the target, making it crucial to consider the target as a more generalized complex field.



(a) hologram (b) MP (c) ∂**H**

Figure 3. Experiment 2: verification on a tilia root microscopy samples with the setup in Figure 1(a). The hologram (a), MP (b) and ∂**H** (c) reconstructions.

∂H also works well on microscopy samples of realistic phase and amplitude distributions. **Figure 3** shows the imaging of a tilia root. The multi-plane (MP) phase retrieval result serves as a reference due to its relatively consistent nature and widespread usage [15]. Both the amplitude and the phase reconstructed by ∂H in Figure 3(c) match well with the reference reconstruction of the MP phase retrieval from five holograms in Figure 3(b). Except for the clear background of the reconstructed phase by ∂H, the magnified portions demonstrate that the ∂H achieves an even better resolution. It is crucial to recognize that the performance of the MP technique heavily depends on the alignment of the system. Implementing the MP approach requires capturing multiple images with axis translation, which introduces potential errors during the process. These errors may include inaccuracies in the translation distance and direction, as well as camera rotation.



(a) hologram (b) DCOD [14] (c) ∂**H**

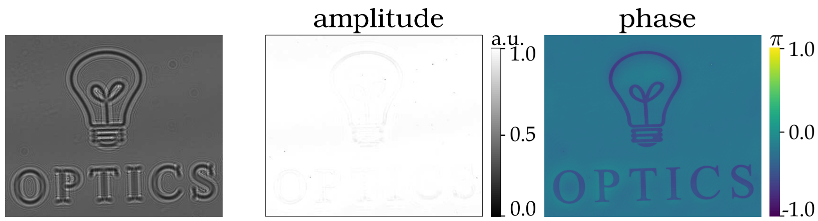
Figure 4. Experiment 3: verification on a cheek cell sample with the setup in Figure 1(c) (data from [14]). The hologram (a), DCOD (b) and ∂H (c) reconstructions.

Figure 4 shows a comparison to state-of-the-art learning-based complex field imaging with a single-shot inline hologram based on the physical model: Deep Compressed Object Decoder (DCOD) [14]. ∂H reconstructed the phase image with a clear background so that higher frequency with more detail can be observed after 25000 iterations in 46 seconds on an Nvidia GTX 1080 GPU, whereas it took 40 minutes to run 30000 iterations for DCOD on an Nvidia Tesla k80 GPU as reported in [14], shown in Table 1.

Table 1. Computational efficiency compared to DCOD [14].

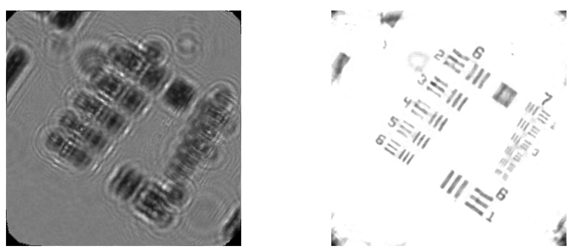
	image size (in pixel)	iteration	time cost	GPU
DCOD	512×512	30000	∼40 minutes	Nvidia Tesla k80
∂**H**	512×512	2500	∼36 seconds	Nvidia GTX 1080
	1024×1024	2500	∼113 seconds	

Besides, DCOD fails in reconstructing some parts of the target, especially around the image border, whereas ∂H reconstructed the image clearly everywhere. People may claim that other neural networks can achieve much faster imaging speed than DCOD [14]. In fact, there are numerous neural networks that could handle complex field imaging [15], phase imaging [17], or auto-focusing phase imaging with holograms [8]. It is true that neural networks outperform optimization in certain aspects, for example, computational speed. However, the inexplicable models trained by neural networks lack generalization and can hardly be further manipulated. Additionally, research related to neural networks driven by data is not practical if a large amount of training data is not accessible. Although the most recent self-supervised learning only needs simulation data and can conduct complex field imaging with two or more holograms, it is still unable to handle the experiment-theory mismatch [18]. On the contrary, the ∂H solves the widespread mismatch issues, which neither optimization nor neural networks have yet sufficiently addressed. The reference [14] was chosen for comparison because it is the most recent single-shot complex field imaging technique that uses only one hologram based on the physical model.



(a) hologram (b) ∂H reconstruction.

Figure 5. Experiment 4: Complex field imaging with the spherical wave illumination setup of Figure 1(b). The hologram (a) and (b) ∂H reconstructions.



(a) hologram (b) ∂H reconstruction

Figure 6. Experiment 5: Intensity-only object imaging with fiber bundle setup of Figure 1(d) (data from [21]). The hologram (a) extracted from a fiber-bundle setup, and (b) ∂H reconstructions.

In the preceding experiments, we validated the viability of the proposed ∂H method by conducting comparisons with ground-truth measurements, as well as existing approaches such as MP phase retrieval and learning-based methods. Nonetheless, it is important to recognize that the potential of ∂H extends beyond these comparisons. The interpretable approach of the proposed ∂H also enables it to work well on a variety of datasets from various setups, as demonstrated in **Figure 5** and **Figure 6**. In Figure 5, a divergent spherical wave was used as illumination. This setup can achieve a large field of view and reduce interference of reflected beams from the object and camera sensor. In this particular experiment, the object under investigation was a wafer containing a highly transmissive pattern etched onto its surface. The height of these patterns measured $200\,\mathrm{nm}$, which corresponds to a phase difference of 0.3174π when illuminated with a wavelength centered at $630\,\mathrm{nm}$. The reconstructed patterns exhibited a phase value of approximately -0.316π, closely matching the intended design. Regarding Figure 6, the experimental setup employed a fiber bundle lensless imaging system. The target object used in this setup was a positive resolution chart, which exhibited no phase difference and could therefore be considered an amplitude-only object. The reconstructed images obtained through the ∂H technique exhibited clear patterns without any presence of twin images. In both Figure 6 and Figure 4, the lensless setup utilized LED light sources, requiring a very short object-camera distance to maintain the desired coherence of the light sources. However, this proximity poses a challenge when it comes to eliminating unwanted terms in the holograms, especially when the object-camera distance is inaccurately determined. Despite this difficulty, the ∂H approach is capable of convincingly reconstructing clear images even with a roughly estimated distance.

3 Discussions and Conclusion

In conclusion, we have successfully demonstrated the effectiveness of ∂H as a method for enabling complex field imaging using a single inline hologram without the need for additional optics. This innovative approach has the potential to inspire new solutions for compact and in vivo 3D imaging [21].

When compared to optimization-based compressive holography [5] and its variants [6, 12, 22], ∂H offers several distinct advantages: (i) It adheres to a rigorous formulation of holographic imaging, ensuring a reliable and accurate reconstruction process. (ii) System imperfections are treated as additional variables to be optimized, allowing for better compensation and enhanced imaging performance. (iii) The framework is designed to be differentiable, enabling efficient optimization and allowing for the incorporation of gradient-based algorithms. (iv) ∂H is capable of reconstructing complex fields, providing more comprehensive and detailed imaging results. In this work, we have made the complex field, target location, and illumination light magnitude differentiable, facilitating the optimization process. Additionally, the differentiable design of the framework, as described in Section 4.3, allows for straightforward extensions by introducing other system parameters or modifying the object scattering model [23, 24]. This flexibility enables the adaptation of ∂H for various types of imaging, broadening its potential applications.

The ∂H is capable of handling multiple specific variables as neural networks. Unlike conventional neural networks [15, 13, 16, 17, 14] that typically learn as black-box models, ∂H addresses the common challenge of experiment-theory mismatch. It provides a mechanism for incorporating system parameters into the imaging model, making the method adaptable to various data types acquired from different experimental setups. Furthermore, ∂H does not require an extensive amount of training data. To the best of our knowledge, this is the first learning-free approach for complex field imaging using a single-shot inline hologram. In contrast to traditional holography techniques, ∂H offers numerous advantages, such as effectively bridging the gap between theory and experiment, enabling generalization across different setups, and not relying on extensive training datasets.

Beyond the demonstrated success in holography, we firmly believe that the underlying philosophy of differentiable imaging will lead to significant advancements in imaging with complex systems or in scenarios that require multiple captures. By either addressing the theory-experiment gap or developing novel imaging modalities, differentiable imaging has the potential to revolutionize the field and unlock new possibilities for imaging applications.

4 Method

To showcase the versatility of ∂H and its ability to apply to various scenarios, we initially present the method in a generalized form in Sections of 4.1 to 4.2. Following that, we offer an example that incorporates domain-specific parameters, effectively tackling the complexities associated with field imaging from one single-shot inline hologram in Section 4.3.

4.1 General problem formulation

We approach holographic imaging as an inverse imaging problem. Unlike traditional computational imaging methods that employ an imaging forward model as $\mathbf{y} = f(\mathbf{x})$ with the objective of finding $f^{-1}(\mathbf{y}) \rightarrow \mathbf{x}$, we are considering a different approach here. Our approach involves a forward model $f(\mathbf{x}, \boldsymbol{\theta})$ that establishes a relationship between the intensity-only inline hologram y and the target object $\mathbf{x}$, while accounting for various unknown parametric conditions represented by $\boldsymbol{\theta}$. These conditions may include optical aberrations, the combined impact of the light source and the sensor on the illumination wave, and other similar factors.

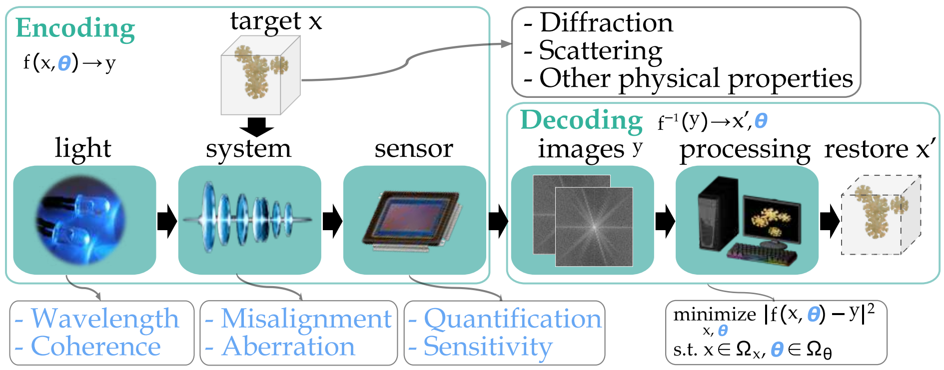


Figure 7. Realistic factors (blue color text represented by $\boldsymbol{\theta}$) in a typical holography system that may affect the modeling. ∂**H** models and retrieves some of these parameters.

The inversion of this forward model is achieved by minimizing an error metric with constraints, defined as (i) a least squares fitting to penalize the deviation of the model from measurement data (i.e., reality), (ii) regularization terms (soft constraints) to favor desirable properties, and (iii) hard constraints to guarantee specific physical constraints for the reconstruction to fulfill. These are collectively written as:

$$\begin{aligned} \underset{\mathbf{x},\,\boldsymbol{\theta}}{\text{minimize}} \quad & \|f(\mathbf{x}, \boldsymbol{\theta}) - \mathbf{y}\|^2 + \sum_{n=1}^{N} \beta_n \mathcal{R}_n(\mathbf{x}, \boldsymbol{\theta}), \\ \text{subject to} \quad & \mathbf{x} \in \Omega_{\mathbf{x}}, \;\; \boldsymbol{\theta} \in \Omega_{\boldsymbol{\theta}}, \end{aligned} \tag{1}$$

where $\|f(\mathbf{x}, \boldsymbol{\theta}) - \mathbf{y}\|^2$ ensures the fidelity of the data, $\mathcal{R}_n(\cdot)$ is a collection of soft regularizers, β_n are weights, and $\Omega_{\mathbf{x},\boldsymbol{\theta}}$ are respective physical constraints.

Given a large number of unknowns, the non-linearity of the model, and the non-convexity of the constraints, solving Eq. (1) is non-trivial and a naïve solver cannot converge in practice. Based on the modeling, the proposed differentiable holography mitigates these challenges, by adopting differentiable programming with reverse automatic differentiation [25]. Specifically, Eq. (1) is solved by projected gradient descent, as shown in Algorithm 1.

Algorithm 1 Differentiable holography solver for Eq. (1).

function RECONSTRUCT OBJECT AND SYSTEM PARAMETERS(y)
  Initialize $\{\mathbf{x}^0, \boldsymbol{\theta}^0\}$; ▷ Initialization
  $f^0 \leftarrow f(\cdot, \boldsymbol{\theta}^0)$; ▷ Initialization of the forward imaging model
  while not converged **do** ▷ Iteration
    $\mathbf{x}^{k+1}, \boldsymbol{\theta}^{k+1} \leftarrow \arg\min_{\mathbf{x},\,\boldsymbol{\theta}} \; \|f^k(\mathbf{x}, \boldsymbol{\theta}) - \mathbf{y}\|^2 + \sum_{n=1}^{N} \beta_n \mathcal{R}_n(\mathbf{x}, \boldsymbol{\theta})$; ▷ update with automatic differentiation (section 4.2)
    $\mathbf{x}^{k+1} \leftarrow \Omega_{\mathbf{x}}(\mathbf{x}^{k+1}),\; \boldsymbol{\theta}^{k+1} \leftarrow \Omega_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{k+1})$; ▷ Projection on physical constraints
    $f^{k+1} \leftarrow f(\cdot,\; \boldsymbol{\theta}^{k+1})$; ▷ Imaging model update
  return $\{\mathbf{x}^K, \boldsymbol{\theta}^K\}$; ▷ At final iteration K

The target $\mathbf{x}$ and the system parameters $\boldsymbol{\theta}$ are updated by gradient descent, while the derivatives are calculated using backward-mode automatic differentiation, which computes accurate derivatives of a computer program by operating directly on the parameters of interest. Complex numbers are treated as mono variables with the conjugate Wirtinger derivative. For non-differentiable functions, approximate gradients are adapted. The detail is presented in section 4.2. The physical constraints are employed by projecting the updated parameters to the physical domain. Regularizers for complex numbers are applied to amplitude and phase, respectively, as described in appendix B. We implemented the method in PyTorch, which provides a dynamic computational graph and allows the forward imaging model to be altered at run-time, thus allowing for optimizing more than one target. It is important to note that PyTorch is not the only programming language available for implementation. Other languages such as TensorFlow [26], JAX [27], or Julia [28] can also be utilized. The selection of a programming language largely relies on the user's preferences, familiarity, and the specific needs of the problem. A projection adaptive momentum estimation [29] was used as the optimizer. Compared to compressive holography [5, 6], a more accurate imaging model with system parameters is modeled instead of an approximated linear model, and the model differentiability allows versatility in formulating the inverse problem, as well as incorporating plug-and-play priors and physical constraints, thus enable single-shot complex field imaging.

4.2 Automatic differentiation optimization

Optimization of the unconstrained part in the algorithm of Eq. (1) is achieved by using gradient descent of the loss function that relies on iterative-refined optimization. This concerns solving the following sub-optimization problem:

$$\min_{\mathbf{x},\,\boldsymbol{\theta}} \quad \mathcal{L}(\mathbf{x},\boldsymbol{\theta}) = \|f(\mathbf{x},\boldsymbol{\theta}) - \mathbf{y}\|^2 + \sum_{n=1}^{N} \beta_n \mathcal{R}_n(\mathbf{x},\boldsymbol{\theta}), \tag{2}$$

where at iteration n, given a step size τ, we update $\mathbf{x}^{n+1}$ and $\boldsymbol{\theta}^{n+1}$ by

$$\begin{cases} \mathbf{x}^{n+1} \leftarrow \mathbf{x}^{n} - \tau \left.\dfrac{\partial \mathcal{L}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}^n}, & \text{(3a)} \\ \boldsymbol{\theta}^{n+1} \leftarrow \boldsymbol{\theta}^{n} - \tau \left.\dfrac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^n}. & \text{(3b)} \end{cases}$$

To achieve optimization of Eq. (3), gradient descent methods like mirror descent or its variants can be utilized. In our scenario, we found that the default gradient implementation in PyTorch, specifically the conjugate Wirtinger derivative, is sufficient for convergence purposes and offers simplicity. Generally, the analytic expression of $\left(\frac{\partial \mathcal{L}}{\partial \mathbf{x}}, \frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}}\right)$ is derived by writing an explicit expression for the error metric $\mathcal{L}$ and symbolically differentiating with respect to each of the input parameters. However, it is tedious and even impossible in our multivariate case, and the partial derivatives indicate that both $\mathbf{x}$ and $\boldsymbol{\theta}$ affect the error metric locally. Here we apply the reverse-mode automatic differentiation [30]. Automatic differentiation takes advantage of the fact that any computer program, regardless of how complex it may be, executes a series of basic arithmetic operations (addition, subtraction, multiplication, division, etc.) and basic functions (exp, log, sin, cos, etc.). Applying the chain rule repeatedly to these operations allows one to automatically compute derivatives of arbitrary order precisely while using at most a small constant factor more arithmetic operations than the original program. By defining the loss function as a composition of elementary operators with derivatives, we can compute the loss function's derivative by chain rule [31].

Eq. (3a) needs to optimize real-valued loss functions with complex variables, i.e., $\mathcal{L}(\mathbf{x}) : \mathbb{C} \to \mathbb{R}$. Since a non-constant real-valued function of a complex variable is not complex analytic and, therefore, is not differentiable, a complex-variable function is usually viewed as a function of the real and imaginary components of the complex variable, which may not satisfy the Cauchy–Riemann equations and cannot be addressed by complex differentiation [32, 33]. Besides, the real parametrization leads to a multivariate optimization problem that is two times larger than the actual size of the medium and ill-posed. We adopt the Wirtinger derivative [34, 35] and perform a monovarietal optimization in the complex domain. By rewriting a real differentiable function $\mathcal{L}(\mathbf{x})$ as two variable holomorphic function $\mathcal{L}(\mathbf{x}, \mathbf{x}^*)$, where $\mathbf{x} = a + jb$ and $\mathbf{x}^* = a - jb$, we can simplify the complex variable update formula of Eq. (3a) to only refer to the conjugate Wirtinger derivative $\frac{\partial \mathcal{L}}{\partial \mathbf{x}^*}$ as $\mathbf{x}^{n+1} = \mathbf{x}^n - \tau \frac{\partial \mathcal{L}}{\partial \mathbf{x}^*}$, where $\frac{\partial \mathcal{L}}{\partial \mathbf{x}^*} = \frac{1}{2}\left(\frac{\partial}{\partial a} + j\frac{\partial}{\partial b}\right)$ is given by the classic definition of Wirtinger calculus [36].

Automatic calculation of partial derivatives allows us to modify Eq. (2) in a number of ways without alternating the optimization framework. Instead of having a complex field, $\mathbf{x}$ could alternatively be a composite function of physical properties such as the refractive index; The misalignment of the illumination angle in optical diffraction tomography [37] or Fourier ptychography microscopy [38], or any other imaging system with multiple captures, could also be considered as system parameters of $\boldsymbol{\theta}$; We may also simply incorporate the forward imaging model of $f(\cdot)$, which is susceptible to multiple scattering and could complicate the inverse imaging problem [23, 24, 39]; Additionally, the least square data term and regularization can be tailored domain-specifically without taking optimization into account. In summary, Eq. (2) can be constructed as a mix-and-match approach, allowing us to concentrate on problem-solving rather than optimization techniques.

4.3 The ∂H formulation of complex field imaging problem

Solving Eq. (1) can be challenging in general. Instead, it is more desirable to focus on solving specific scenarios tailored to match the experimental conditions. In inline holography, the forward imaging model is greatly affected by factors of the distance the wave travels from the object to the sensor and the intensity of the illumination. Traditionally, autofocusing techniques [40, 41] have been used before applying inverse optimization solvers, and intensity compensation has been achieved by utilizing background images. However, it is crucial to acknowledge that autofocusing and obtaining appropriate background images may not always be practical or readily accessible, which presents challenges for conventional hand-crafted optimization and learning-based methods. In the complex field imaging case where $\mathbf{x}$ is a complex transmittance function, denoted as $\mathbf{t}$, the major factors that affect the forward model are the illumination amplitude a and the object-camera distance z (auto-focusing). Let $\boldsymbol{\theta} = \{a, z\}$, the problem rephrases as follows:

$$\begin{aligned} \underset{\mathbf{t},\, a,\, z}{\text{minimize}} \quad & \|f(a \cdot \mathbf{t},\, z) - \mathbf{y}\|^2 + \beta_1 \mathcal{R}_{\ell_1}(\mathbf{t}) + \beta_2 \mathcal{R}_{\mathrm{TV}}(\mathbf{t}), \\ \text{subject to} \quad & |\mathbf{t}| \leq 1,\; a > 0,\; z > 0, \end{aligned} \tag{4}$$

where $f(\cdot)$ is essentially the free-space wave propagator with a subsequent interference operator, as defined in Appendix A. The constraint $|\mathbf{t}| \leq 1$ enforces sparse energy conservation in the imaging process [42]. The regularizations of ℓ_1 norm $\mathcal{R}_{\ell_1}(\mathbf{t})$ and total variation norm $\mathcal{R}_{\mathrm{TV}}(\mathbf{t})$ are used to favor a sparse and spatially smooth transmittance function. In this specific case, we devised intricate regularization techniques, detailed in Appendix B.

By implementing the technique mentioned in the previously discussed sections, specifically from Section 4.1 to Section 4.2, we can successfully tackle the challenging field imaging problem. However, it is crucial to acknowledge that in addition to the factors depicted in Figure 7, $\boldsymbol{\theta}$ might encompass additional elements to achieve a more realistic imaging formation model for various imaging objectives.

Appendices

Appendix A Forward model of inline holography

Let $\mathbf{r} = (x, y, z)$ be a three-dimensional vector in space. An illumination reference beam $t_i(\mathbf{r})$ propagates from the light source to the detector camera plane located at the original coordinates. In general, the penetration of a wave through an object is described by a complex transmission function that consists of amplitude and phase $t_o(\mathbf{r}) = e^{-a_o(\mathbf{r})}e^{j\phi_o(\mathbf{r})}$, where $a_o(\mathbf{r})$ describes object's absorption property, and $\phi_o(\mathbf{r})$ is the phase delay introduced by the object into the incident reference beam. Suppose $t_s(\mathbf{r}) = e^{-a_s(\mathbf{r})}e^{j\phi_s(\mathbf{r})}$ is the complex transmission function of the surrounding medium (when there is no object that exists), the transmission function with the object in the surrounding medium is $t(\mathbf{r}) = e^{-(a_s(\mathbf{r})+a_o(\mathbf{r}))}e^{j(\phi_s(\mathbf{r})+\phi_o(\mathbf{r}))} = t_s(\mathbf{r})t_o(\mathbf{r})$, and $t(\mathbf{r}) = t_i(\mathbf{r})t_o(\mathbf{r})$ if the surrounding medium is a vacuum, i.e., $t_s(\mathbf{r}) = t_i(\mathbf{r})$. We assume the change in irradiance caused by the object is significantly smaller than the illumination of the beam, and the transmission function of the object can be expressed as $1 + t_o(\mathbf{r})$, where 1 corresponds to the transmittance in the absence of the object and $t_o(\mathbf{r})$ is a complex function caused by the presence of the object, so we have $t(\mathbf{r}) = t_i(\mathbf{r})(1+t_o(\mathbf{r}))$. The captured hologram, i.e., the output of the forward model $f(\cdot)$ in Eq. (1), thus is $f(\cdot) = |t(\mathbf{r}) \otimes h(\mathbf{r})|^2$, which is a self-interference of the propagated two-dimensional scalar field $t(\mathbf{r})$ at $z = 0$, $\otimes$ is the convolution operator, and $h(\mathbf{r}) = \frac{1}{2\pi}\frac{z}{\mathbf{r}}\frac{1-jk\mathbf{r}}{\mathbf{r}^2}\exp(jk\mathbf{r})$ is the point spread function of the free-space wave propagation [43]. In the case of $t_i(\mathbf{r}) = A_i(\mathbf{r})e^{-jkz}$ being a plane wave, the free-space wave propagation is implemented by the hybrid Taylor-Rayleigh-Sommerfeld diffraction [44], while for the spherical illumination of $t_i(\mathbf{r}) = A_i(\mathbf{r})\frac{e^{-jk\mathbf{r}}}{\mathbf{r}}$, there is a magnification of the hologram defined by $M = \frac{z_{ls}}{z_{lo}}$, where z_{ls} is the light source to camera sensor distance and z_{lo} is the light source to object distance [45]. Notably, in contrast to compressive holography [5, 6, 22], despite the fact that there is no linear approximation in the forward imaging model $f(\cdot)$, we could also adopt inline holography systems with plane and spherical wave illuminations.

Appendix B Regularization for complex numbers

In section 4.2, we introduced a real-valued loss function with complex variables that we aimed to optimize. Typically, complex numbers are split into real and imaginary components to fit into the traditional regularization techniques used for real-only numbers [46]. However, because the amplitude and phase of the complex field are internally related, we developed a regularization method that applies to both the amplitude and phase of the complex numbers. Specifically, we utilized the ℓ_1 norm $\mathcal{R}_{\ell_1}(t)$ and total variation norm $\mathcal{R}_{\mathrm{TV}}(t)$ for a complex number t, as follows:

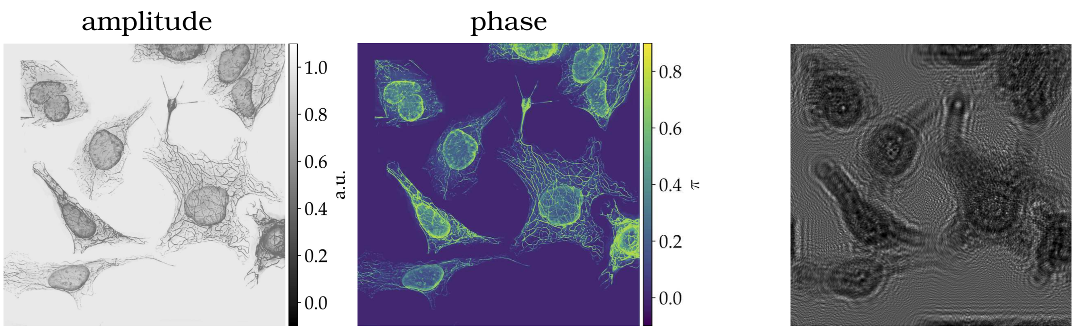
$$\begin{aligned}\mathcal{R}_{\ell_1}(t) &= \mathcal{R}_{\ell_1}(|t|) + \mathcal{R}_{\ell_1}(\angle t)\\ &= \|1 - |t|\|_1 + \|\sin t\|_1,\end{aligned}$$

where $\mathrm{Im}(\cdot)$ is the imaginary operator and $\angle(\cdot)$ is the argument of a complex number. For the amplitude component, we apply the ℓ_1 norm to $1 - |t|$ rather than $|t|$ due to the fact that the illumination beam is significantly larger than the object area in inline holography systems. While for the phase component, we apply it to $|\sin t|$, since $\sin t = \mathrm{Im}(\exp(j\angle t))$ contains the phase component $\angle t$. Similarly, the isotropic total variant is defined as

$$\begin{aligned}\mathcal{R}_{\mathrm{TV}}(t) &= \mathcal{R}_{\mathrm{TV}}(|t|) + \mathcal{R}_{\mathrm{TV}}(\angle t)\\ &= \sum_k \sqrt{\nabla_x^k|t|^2 + \nabla_y^k|t|^2 + \epsilon^2} + \sum_k \sqrt{\nabla_x^k|\sin t|^2 + \nabla_y^k|\sin t|^2 + \epsilon^2},\end{aligned} \tag{5}$$

where $\epsilon > 0$ is a small number that is used to avoid a staircase effect, $\nabla^k_{x,y}(\cdot)$ is a linear operator that performs the finite difference operator along the x, y directions at the k^{th} pixel location.

Appendix C Numerical verification



(a) Ground truth complex field

(b) The hologram

Figure 8. The complex object (a) and the hologram (b) used in the simulations.

To assess the performance of our approach, we conducted numerical simulations. In these simulations, we created an image resembling a cell with a specific refractive index. Figure 8(a) illustrates the amplitude and phase resulting from the object immediately after it. We then simulated the inline hologram of the cell-like image with a plane wave illumination of a wavelength centered at $532\,\text{nm}$, and the corresponding holograms located at $z = 100\,\mu\text{m}$ are shown in Figure 8(b). The camera sensor pixel size was $8\,\mu\text{m}$.

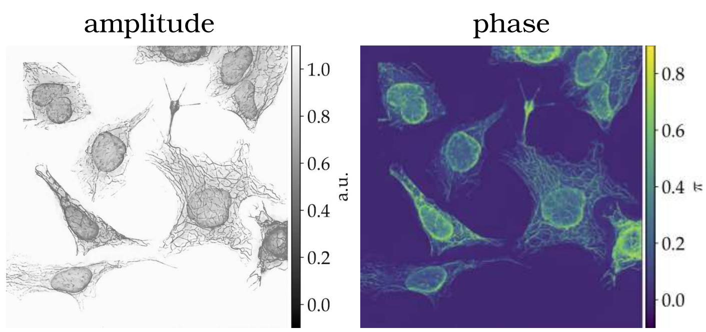


Figure 9. $\partial\mathbf{H}$ reconstruction of the hologram in Figure 8(b).

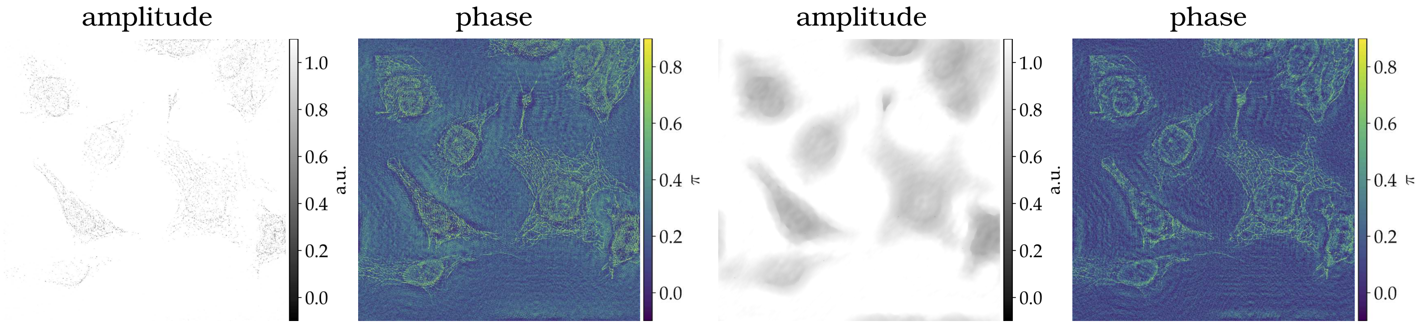
	SSIM	NRMSE	PSNR
Amplitude	0.9514	0.0330	30.2607
Phase	0.8290	0.1504	29.6130

Table 2. Measurement of the reconstructed images in Fig. 9.

The reconstructed image in Figure 9 faithfully represents the complex field, including both the amplitude and phase components. To evaluate the quality of this reconstruction, we analyze the structural similarity index (SSIM), normalized root mean square error (NRMSE), and peak signal-to-noise ratio (PSNR) values, which are provided in Table 2. These metrics, computed using the scikit-image library [47], enable a comprehensive comparison between the reconstructed image and the ground truth.

To showcase the effectiveness of complex regularizations, we perform a reconstruction of the hologram depicted in Figure 8(b) using real-valued regularizations while maintaining the same weights. The reconstructions obtained using $\mathcal{R}_{\ell_1}$ and $\mathcal{R}_{\mathrm{TV}}$ in the real domain are shown in Figure 10, and the assessment metrics are presented in Table 3.

Comparing Figure 9 and Table 2 with Figure 10 and Table 3, it is evident that the conventional real-valued regularizations exhibit lower performance when contrasted with the complex regularizations.



(a) Real $\mathcal{R}_{\ell_1}$ (b) Real $\mathcal{R}_{\mathrm{TV}}$

Figure 10. Simulation with real-valued regularizations.

		SSIM	NRMSE	PSNR
Real $\mathcal{R}_{\ell_1}$	Amplitude	0.6873	0.1719	15.9293
	Phase	0.0703	1.2183	11.4466
Real $\mathcal{R}_{\mathrm{TV}}$	Amplitude	0.6667	0.1148	19.4327
	Phase	0.1070	0.9347	13.7488

Table 3. Measurement of the image reconstructed with real regulations in Fig. 10.

Funding Information

This work was partially supported by the KAUST individual baseline funding.

Acknowledgments

N.C. conceived and developed the idea, conducted the validation, and wrote the manuscript. C.W. contributed to result discussions, and edited the draft of the manuscript. W.H. provided partial financial support for the project. The authors would like to express their gratitude to Hadi Amata for fabricating the sample of the KAUST logo.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

[1] S. Kadlecek, J. Sebby, R. Newell, T. G. Walker, *Optics Letters* **2001**, *26*, 3 137.

[2] T. P. Almeida, A. Palomino, S. Lequeux, V. Boureau, O. Fruchart, I. L. Prejbeanu, B. Dieny, D. Cooper, *APL Materials* **2022**, *10*, 6 061104.

[3] Y. Sung, W. Choi, C. Fang-Yen, K. Badizadegan, R. R. Dasari, M. S. Feld, *Optics Express* **2009**, *17*, 1 266.

[4] D. Gabor, *Nature* **1948**, *161*, 4098 777.

[5] D. J. Brady, K. Choi, D. L. Marks, R. Horisaki, S. Lim, *Optics Express* **2009**, *17*, 15 13040.

[6] W. Zhang, L. Cao, D. J. Brady, H. Zhang, J. Cang, H. Zhang, G. Jin, *Physical Review Letters* **2018**, *121*, 9 093902.

[7] T.-C. Poon, *Digital Holography and Three-Dimensional Display*, Springer, **2006**.

[8] L. Huang, T. Liu, X. Yang, Y. Luo, Y. Rivenson, A. Ozcan, *ACS Photonics* **2021**, *8*, 6 1763.

[9] E. N. Leith, J. Upatnieks, *Journal of the Optical Society of America A* **1962**, *52*, 10 1123.

[10] I. Yamaguchi, T. Zhang, *Optics Letters* **1997**, *22*, 16 1268.

[11] T.-C. Poon, M. H. Wu, K. Shinoda, Y. Suzuki, *Proceedings of the IEEE* **1996**, *84*, 5 753.

[12] F. Momey, L. Denis, T. Olivier, C. Fournier, *Journal of the Optical Society of America A* **2019**, *36*, 12 D62.

[13] Z. Ren, Z. Xu, E. Y. Lam, *Optica* **2018**, *5*, 4 337.

[14] F. Niknam, H. Qazvini, H. Latifi, *Scientific Reports* **2021**, *11*, 1.

[15] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, A. Ozcan, *Light: Science and Applications* **2017**, *7*, 2 17141.

[16] E. Bostan, R. Heckel, M. Chen, M. Kellman, L. Waller, *Optica* **2020**, *7*, 6 559.

[17] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, G. Situ, *Light: Science and Applications* **2020**, *9*, 1.

[18] H. Chen, L. Huang, T. Liu, A. Ozcan, Gedankennet: Self-supervised learning of hologram reconstruction using physics consistency, **2023**.

[19] T. Zeng, E. Y. Lam, *IEEE Transactions on Computational Imaging* **2021**, *7* 1080.

[20] N. Chen, L. Cao, T. Poon, B. Lee, E. Y. Lam, *Advanced Physics Research* **2023**, *2*, 6.

[21] M. R. Hughes, *Applied Optics* **2020**, *60*, 4 A1.

[22] N. Chen, C. Wang, W. Heidrich, *Laser and Photonics Reviews* **2021**, *15*, 8 2100008.

[23] M. Lee, H. Hugonnet, Y. Park, *Optica* **2022**, *9*, 2 177.

[24] M. Chen, D. Ren, H.-Y. Liu, S. Chowdhury, L. Waller, *Optica* **2020**, *7*, 5 394.

[25] D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Nature* **1986**, *323*, 6088 533.

[26] Y. Yu, M. Abadi, P. Barham, E. Brevdo, M. Burrows, A. Davis, J. Dean, S. Ghemawat, T. Harley, P. Hawkins, M. Isard, M. Kudlur, R. Monga, D. Murray, X. Zheng, Dynamic control flow in large-scale machine learning, https://www.tensorflow.org, **2015**, Software available from tensorflow.org.

[27] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, Q. Zhang, JAX: composable transformations of Python+NumPy programs, http://github.com/google/jax, **2018**.

[28] J. Bezanson, A. Edelman, S. Karpinski, V. B. Shah, *SIAM Review* **2017**, *59*, 1 65.

[29] B. Heo, S. Chun, S. J. Oh, D. Han, S. Yun, G. Kim, Y. Uh, J.-W. Ha, In *International Conference on Learning Representations*. arXiv, **2020** .

[30] A. Griewank, *Evaluating derivatives*, Society for Industrial and Applied Mathematics, **2008**.

[31] M. Blennow, *Mathematical Methods for Physics and Engineering*, CRC Press, **2018**.

[32] Y. Peng, S. Choi, N. Padmanaban, G. Wetzstein, *ACM Transactions on Graphics* **2020**, *39*, 6 1.

[33] C. Chen, B. Lee, N.-N. Li, M. Chae, D. Wang, Q.-H. Wang, B. Lee, *Optics Express* **2021**.

[34] R. Remmert, In *Theory of Complex Functions*, 9–44. Springer New York, New York, NY, ISBN 9781461209393, **1991**.

[35] J. A. Palmer, K. Kreutz-Delgado, S. Makeig, The complex gradient operator and the cr-calculus, **2009**.

[36] N. Chen, C. Wang, W. Heidrich, *Frontiers in Photonics* **2022**, *3*.

[37] Y. Baek, Y. Park, *Nature Photonics* **2021**, *15*, 5 354.

[38] A. Zhou, W. Wang, N. Chen, E. Y. Lam, B. Lee, G. Situ, *Optics Express* **2018**, *26*, 18 23661.

[39] H. Wang, W. Tahir, J. Zhu, L. Tian, *Optics Express* **2021**, *29*, 11 17159.

[40] Z. Ren, N. Chen, E. Y. Lam, *Optics Letters* **2017**, *42*, 9 1720.

[41] Y. Zhang, Z. Huang, S. Jin, L. Cao, *Applied Optics* **2023**, *62*, 10 D23.

[42] T. Latychevskaia, H.-W. Fink, *Physical Review Letters* **2007**, *98*, 23.

[43] J. W. Goodman, *Introduction to Fourier optics*, W. H. Freeman, Englewood, Colo, **2017**.

[44] N. Chen, C. Wang, W. Heidrich, *Optics Express* **2022**, *30*, 21 37727.

[45] T. Latychevskaia, H.-W. Fink, *Applied Optics* **2015**, *54*, 9 2424.

[46] Y. Gao, L. Cao, *Light: Advanced Manufacturing* **2023**, *4*, LAM2022050014 1.

[47] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, *PeerJ* **2014**, *2* e453.