PHOTONICS Research

Neural-field-assisted transport-of-intensity phase microscopy: partially coherent quantitative phase imaging under unknown defocus distance

Yanbo Jin,[1,2,3,†] Linpeng Lu,[1,2,3,†] Shun Zhou,[1,2,3] Jie Zhou,[1,2,3] Yao Fan,[1,2,3] and Chao Zuo[1,2,3,*]

[1]Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
[2]Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing 210019, China
[3]Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing 210094, China
[†]These authors contributed equally to this work.
*Corresponding author: zuochao@njust.edu.cn

Received 19 February 2024; revised 28 April 2024; accepted 9 May 2024; posted 9 May 2024 (Doc. ID 521056); published 1 July 2024

The transport-of-intensity equation (TIE) enables quantitative phase imaging (QPI) under partially coherent illumination by measuring the through-focus intensities combined with a linearized inverse reconstruction algorithm. However, overcoming its sensitivity to imaging settings remains a challenging problem because of the difficulty in tuning the optical parameters of the imaging system accurately and because of the instability to long-time measurements. To address these limitations, we propose and experimentally validate a solution called neural-field-assisted transport-of-intensity phase microscopy (NFTPM) by introducing a tunable defocus parameter into neural field. Without weak object approximation, NFTPM incorporates the physical prior of partially coherent image formation to constrain the neural field and learns the continuous representation of phase object without the need for training. Simulation and experimental results of HeLa cells demonstrate that NFTPM can achieve accurate, partially coherent QPI under unknown defocus distances, providing new possibilities for extending applications in live cell biology.

https://doi.org/10.1364/PRJ.521056

1. INTRODUCTION

Quantitative phase imaging (QPI) has gained increased interest in optical microscopy research for its capability to quantify optical thickness and morphologies of unlabeled samples [1–3]. The QPI approach can be categorized into iterative and deterministic methods [4–6], where the deterministic method requires the establishment of an analytical expression for the object phase with respect to the measured intensity images. Given that the image formation process in QPI is inherently non-linear, linearization approaches are commonly invoked to facilitate solving for phase as a function of intensity measurements. For example, as a well-established deterministic phase retrieval approach, the transport-of-intensity equation (TIE) applies paraxial approximation and slowly-varied approximation to linearize the phase retrieval problem and can recover the quantitative phase by utilizing intensity images at multiple axially defocused planes [6,7]. Under partially coherent illumination, TIE is expected to achieve improved spatial resolution beyond the coherent diffraction limit [8]. Nevertheless, in a conventional microscope with circular illumination, partial coherence tends to diminish the phase contrast, resulting in compromised imaging resolution [9].

To achieve high-resolution and high-contrast QPI, the annular illumination (AI) matching objective numerical aperture (NA) has been employed in deconvolution-based TIE, referred to as AI-TIE [10]. AI-TIE strongly boosts the phase contrast and significantly improves the practical imaging resolution to a 2-fold objective NA. The strong phase contrast is ultimately transformed to the quantitative phase images by WOTF (weak object transfer function) inversion, yielding high-quality results with enhanced resolution. However, AI-TIE is usually limited to weak scattering samples since it linearizes the image formation model by invoking weak object approximation with ignoring higher-order terms in the complex transmittance of the sample. In addition, WOTF is a function directly related to the light source distribution, objective pupil function, and defocus distance. Once WOTF is determined, AI-TIE is not capable of adaptively adjusting optical parameters such as the defocus distance during the imaging process. Therefore, such TIE-based methods may result in degraded quality of phase retrieval due to the inaccurate inverse reconstruction for nonweak objects or cases where optical parameters are incorrect.

In contrast to the aforementioned physics-based approaches [6,10], data-driven deep learning methods can establish the nonlinear pseudo-inverse mapping relation between the defocused intensity and the object phase [11–14], bypassing the obstacle of "solving nonlinear ill-posed inverse problems." Essentially, the major reason for the success of deep learning is the abundance of training data and the explicit agnosticism from *a priori* knowledge of how such data are generated [15]. However, high-quality paired data acquisition in experiments requires professional supervision and extensive labor. Furthermore, the lack of data diversity will restrict its generalization to out-of-domain cases with dissimilar optical parameters. Thus, the data-driven deep learning methods tend to fail in situations where it is difficult to obtain a large amount of high-quality paired data from a variety of different imaging systems.

To overcome the above limitations, researchers have developed untrained network approaches by incorporating physical priors into deep neural networks, such as the deep phase decoder [16] and PhysenNet [17]. These methods aim to achieve nonlinear optimization by minimizing the error between the prior model-generated image and the actual measurement. Their superiority lies in introducing neural networks as advanced regularization for automatic tuning. For instance, deep image prior (DIP) method can use a randomly initialized neural network as a prior to solve inverse problems such as pixel super-resolution [18]. Especially, the BlindNet method takes distance uncertainty into account and further addresses the phase retrieval problem with unknown defocus distance [19]. Additionally, we have witnessed the rise of neural field (NF), which has become a prominent self-supervised learning method [20]. NF can represent a three-dimensional (3D) scene as a continuous field, which is parameterized by a lightweight multilayer perceptron (MLP, i.e., fully connected network) and trained without ground truth data. In conjunction with computational imaging techniques, NF typically dispenses with training on a dataset and iterates the MLP network directly on the test data until the desired physical quantities are recovered, similar to physics-driven untrained network approaches. For example, NF can be incorporated into 3D diffraction tomography [21] or two-dimensional (2D) microscopy such as lensless microscopy [22] and Fourier ptychographic imaging [23]. However, these physics-driven deep learning methods only involve coherent imaging and are unsuitable for partially coherent imaging scenarios. In fact, considering partial coherence in phase retrieval helps to yield accurate results thanks to its better alignment with the actual situation [24]. Nevertheless, it needs to introduce additional parameters (such as coherence parameter) to establish a more complete forward model. Consequently, it remains a challenge to achieve stable partially coherent QPI under varying optical parameters.

In this work, we present a partially coherent QPI approach by using a neural field and taking the Abbe imaging model [25] as the physical prior. The proposed method, termed neural-field-assisted transport-of-intensity phase microscopy (NFTPM), is actually a gradient-based iterative algorithm. It drives a coordinate-based MLP through the physical prior to represent the phase distribution as a neural field and optimizes the MLP using the gradient computed by backpropagation in both the physical model and the MLP model. This framework empowers NFTPM to concurrently adjust the defocus distance of the physical model by introducing a tunable defocus parameter, enabling stable QPI under unknown defocus distance. Moreover, NFTPM is applicable to non-weak phase objects, since the weak object approximation is not applied to the forward image formation. Instead of an image-to-image 2D CNN, NFTPM forms a point-to-point mapping function from spatial coordinates to phase values, which effectively constrains the solution space and renders single-shot QPI possible. Unlike untrained networks based on coherent imaging systems, NFTPM can adapt to various partially coherent illuminations, which is validated by simulations under circular illumination and annular illumination. Furthermore, based on a bright-field microscope equipped with annular NA-matched illumination [26,27] formed by sparsely distributed light-emitting diode (LED) elements, we realize stable QPI of unstained Henrietta Lacks (HeLa) cells, demonstrating that NFTPM is a valid approach for adaptive correction of defocus aberration during the long-term phase microscopy. Given the simplicity and effectiveness of the NFTPM method, it promises to advance the integration of partially coherent imaging with physics-driven deep learning and open new possibilities for robust non-interferometric QPI in dynamic optical environments.

2. METHODS

A. Reconstruction Algorithm of NFTPM

The schematic diagram of NFTPM is outlined in Fig. 1(a), and Fig. 1(b) illustrates the image formation process in a partially coherent microscope, which corresponds to the physics prior used to drive NFTPM to perform phase recovery. The pipeline of NFTPM comprises a radial encoding module [21] and a 5-layer MLP ($\mathbf{W}$ is the weights) that maps 2D spatial coordinate $\mathbf{r} = (x, y)$ to phase value $\phi(\mathbf{r})$, which can finally represent the phase as a neural field $\Phi_{\mathbf{W}}(\mathbf{r})$. We adopt $M \times N$ grid coordinates $\mathbf{R} = \{(x_i, y_i)\}_{i=0}^{M\times N-1}$ for the field of view (FOV) of interest (generally $-1 \le x \le 1$, $-1 \le y \le 1$), and the coordinates correspond to pixels on the image sensor. Initially, we utilize the radial encoding module to map densely distributed two-dimensional coordinate points to sparsely distributed high-dimensional space, thus allowing the MLP to better discriminate between different coordinate positions in order to characterize high-frequency information. For $\mathbf{r} = (x, y) \in \mathbb{R}^{1\times 2}$, radial encoding can be expressed as

$$\mathbf{r}_{\text{rad}} = T_f\{\cos(\mathbf{T}_L\mathbf{r}\mathbf{T}_R), \sin(\mathbf{T}_L\mathbf{r}\mathbf{T}_R)\}, \tag{1}$$

where $\mathbf{r}_{\text{rad}}$ is the encoded feature, and $\mathbf{T}_L$ is the transformation matrix used for frequency expansion, which can be defined as

$$\mathbf{T}_L = [2^0\pi, 2^1\pi, \cdots, 2^{L-1}\pi]^{\mathrm{T}}, \tag{2}$$

where L is the number of the expanded frequencies. The purpose of introducing L frequencies is to characterize features at various scales in the radial positions. $\mathbf{T}_R$ contains multiple rotation matrices, and it can be specified as

$$\mathbf{T}_R = \begin{bmatrix} 1 & 0 & \cdots & \cos\theta_i & \sin\theta_i & \cdots \\ 0 & 1 & \cdots & -\sin\theta_i & \cos\theta_i & \cdots \end{bmatrix}, \tag{3}$$

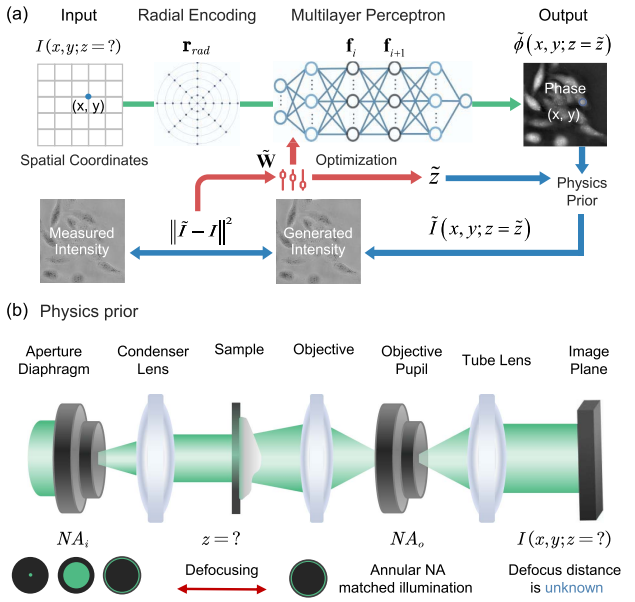

Fig. 1. (a) Schematic diagram of our proposed NFTPM method. (b) The physics prior (forward image formation model) of NFTPM.

where $\theta_i = 2\pi i/N_\theta\,(i = 0, 1, \ldots, N_\theta - 1)$, and N_θ is the number of rotation intervals. The rotation θ_i further enables the MLP to respond to features at diverse orientations, allowing for better feature representation and avoiding noise [21]. $T_f\{\cdot\}$ is applied to flatten concatenated matrices into a vector, which for matrices $\mathbf{A}$ and $\mathbf{B}$ can be defined as

$$T_f\{[\mathbf{A},\mathbf{B}]\} = T_f\left\{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} & b_{11} & b_{12} & \cdots & b_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} & b_{21} & b_{22} & \cdots & b_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \cdots & a_{\mathrm{MN}} & b_{M1} & b_{M2} & \cdots & b_{\mathrm{MN}} \end{bmatrix}\right\} = [a_{11}, a_{12}, \cdots, a_{1N}, b_{11}, b_{12}, \cdots, b_{1N}, \cdots, a_{M1}, a_{M2}, \cdots, a_{\mathrm{MN}}, b_{M1}, b_{M2}, \cdots, b_{\mathrm{MN}}]. \tag{4}$$

Here, we simplify the radial encoding module by replacing Eq. (3) with

$$\mathbf{T}_R = \begin{bmatrix} 1 & \cdots & \cos\theta_i & \cdots \\ 0 & & \sin\theta_i & \end{bmatrix}, \tag{5}$$

and including $\mathbf{r} = (x, y)$ into the encoded feature. Hence, $\mathbf{r} = (x, y)$ can be processed like the 2D Fourier series expansion

$$\mathbf{r}_{\mathrm{rad}} = [x, y, \cdots \cos(2\pi(\mathbf{u}_{l,i}x + \mathbf{v}_{l,i}y)), \sin(2\pi(\mathbf{u}_{l,i}x + \mathbf{v}_{l,i}y)), \cdots]_{l:0\le l\le L-1}, \tag{6}$$

where $\mathbf{r}_{\mathrm{rad}} \in \mathbb{R}^{1\times(2+2LN_\theta)}$, $\mathbf{u}_{l,i} = [2^{l-1}\cos(\theta_i)]_{i:0\le i\le N_\theta-1}$, and $\mathbf{v}_{l,i} = [2^{l-1}\sin(\theta_i)]_{i:0\le i\le N_\theta-1}$. To retrieve the phase distribution, we input $\mathbf{r}_{\mathrm{rad}}$ into the K-layer MLP of the neural field $\Phi_{\mathbf{W}}$ ($K = 5$), incorporating the following processing modules. (1) Linear module $\mathbf{W}_0 \in \mathbb{R}^{(2+2LN_\theta)\times C}$, converting $\mathbf{r}_{\mathrm{rad}}$ into hidden features with C channels ($C = 128$). (2) Linear module for hidden features $\mathbf{W}_i \in \mathbb{R}^{C\times C}$ ($i = 1, 2, \ldots, K - 2$). (3) The last linear transformation $\mathbf{W}_{K-1} \in \mathbb{R}^{C\times 1}$. (4) Leaky rectified linear unit (LeakyReLU) σ_i ($i = 0, \ldots, K - 2$). (5) Sigmoid activation function σ_{K-1}. Specifically, let the ith feature be denoted as $\mathbf{f}_i$, and then the $(i + 1)$th feature is given by

$$\mathbf{f}_{i+1} = \sigma_i(\mathbf{f}_i\mathbf{W}_i), \tag{7}$$

where $i = 0, \ldots, K - 1$, and $\mathbf{f}_0 = \mathbf{r}_{\mathrm{rad}}$. The phase value can be represented as $\phi(\mathbf{r}) = \Phi_{\mathbf{W}}(\mathbf{r}) = 2\pi\mathbf{f}_K$ ($\mathbf{f}_K$ is the output of the K-layer MLP), and $\phi(\mathbf{R})$ can be reshaped as an image of phase distribution. For most biological samples, the complex transmittance can be expressed as $t(\mathbf{r}) = e^{j\phi(\mathbf{r})}$. In a typical $6f$ optical imaging system, the source with distribution $S_{\mathrm{pc}}(\mathbf{u})$ at the aperture diaphragm plane ($\mathbf{u}$ corresponds to the 2D coordinates in Fourier space) provides partially coherent illumination, resulting in an image captured at the image plane,

$$I(\mathbf{r}) = \iint T(\mathbf{u}_1)T^*(\mathbf{u}_2)\mathrm{TCC}(\mathbf{u}_1,\mathbf{u}_2)e^{j2\pi\mathbf{r}(\mathbf{u}_1-\mathbf{u}_2)}\mathrm{d}\mathbf{u}_1\mathrm{d}\mathbf{u}_2, \tag{8}$$

where $T(\mathbf{u})$ is the Fourier transform of $t(\mathbf{r})$, and TCC (transmission cross-coefficient) [10,28] satisfies the following relation:

$$\mathrm{TCC}(\mathbf{u}_1,\mathbf{u}_2) = \int S_{\mathrm{pc}}(\mathbf{u})P(\mathbf{u}+\mathbf{u}_1)P^*(\mathbf{u}+\mathbf{u}_2)\mathrm{d}\mathbf{u}, \tag{9}$$

where $P(\mathbf{u}) = |P(\mathbf{u})|e^{jkz\sqrt{1-\lambda^2|\mathbf{u}|^2}}$ represents the complex pupil function of the imaging system, z is the defocus distance along the optical axis, k is the wavenumber, and $|P(\mathbf{u})|$ is a circular function determined by the objective NA and wavelength λ. The TCC formula is an abstraction for the spectral coupling of a light source to an objective pupil, intrinsically characterizing the imaging system compatible with partially coherent illumination. When the illumination distribution $S_{\mathrm{pc}}(\mathbf{u})$ of the imaging system is specified, the captured image is determined by the sample's inherent property (phase delay ϕ) and the defocus distance z. Therefore, we can use a function $H\{\phi, z\}$ to represent the image formation model of I.

In order to achieve phase retrieval with defocus distance prediction, the uncertain defocus distance can be incorporated into the computational graph as a tunable parameter z to be optimized along with the MLP. The trade-off in determining the optimal solution of NFTPM is to ensure the accuracy of the

predicted defocus distance while minimizing the error between the generated intensity image and the measurement. Given a captured intensity I, the spatial coordinates $\mathbf{R} = \{\mathbf{r}_i\}_{i=0}^{M\times N-1}$ are fed into $\Phi_{\mathbf{W}}$ to obtain the phase, which is then processed through the physical model $H\{\phi, z\}$ to generate intensity $\tilde{I}$ for comparison with I using the mean square error (MSE) loss function. The above operations can be abstracted into an optimization problem

$$\mathbf{W}^{\dagger}, z^{\dagger} = \arg\min_{\mathbf{W},z} \sum_{\mathbf{r}\in\mathbf{R}} \left\| H\{\Phi_{\mathbf{W}}(\mathbf{R}), z\} - I \right\|_2^2, \tag{10}$$

where $\Phi_{\mathbf{W}^{\dagger}}(\mathbf{R})$ is the retrieved phase, and $z^{\dagger}$ is the predicted defocus distance. The optimization is executed based on back-propagation and the gradient descent algorithm [29], and the specific optimization process is described in Section 7 of Ref. [30]. It is worth mentioning that the samples are assumed as pure phase objects in NFTPM, so the phase contrast provided by a single-shot defocused intensity is sufficient for precise phase recovery based on the principle of deep image prior [18,31].

B. Experimental Setup

Neural-field-assisted transport-of-intensity phase microscopy can be easily implemented on a commercial inverted bright-field microscope (IX83, Olympus, Japan) assisted by programmable LED array illumination due to the advantage of non-interferometric measurements. The LED array provides quasi-monochromatic illumination with a center wavelength of 525 nm and spectral bandwidth of 20 nm. These LED elements can be controlled to turn on to form point, circle, or annulus patterns by a field-programmable gate array (FPGA) unit (EP4CE10E22C8N, Intel, US). Twelve annularly distributed LED elements were selected in the array, with the center of the circle coinciding with the optical axis to provide matched annular illumination with a maximum illumination NA of 0.4. We utilized a Bertrand lens, positioned in an eyepiece observation tube in place of the normal eyepiece, to examine the rear focal plane of the objective lens. This examination is crucial for confirming that the circular illumination is centered precisely in the field of view or that the annular illumination is accurately inscribed within the objective lens's pupil. A CMOS camera (Hamamatsu ORCA-Flash 4.0 C13440) with a resolution of 2048 × 2048 and a pixel size of 6.5 μm was used to record the intensity information under a detection objective (10×/0.4 UPLSAPO, Olympus). This study was conducted on a workstation equipped with an Intel i9-10900K 3.70 GHz CPU and an NVIDIA GeForce RTX 3090 GPU. The proposed algorithm was operated by Python 3.7.16 and PyTorch 1.12.1.

3. RESULTS

A. Comparison with TIE, AI-TIE, BlindNet, and GS Algorithms

To validate the effectiveness of NFTPM in partially coherent QPI, we conducted simulations to compare the proposed NFTPM with TIE, AI-TIE, BlindNet [19], and GS algorithms [4,5] under both coherent illumination and circular illumination with a coherence parameter (denoted by S, illumination NA/objective NA) of 0.85. It is noteworthy that AI-TIE here refers to all deconvolution-based TIE methods, adapted to coherent (point), partially coherent illuminations and not limited to annular illumination case. As shown in Figs. 2(a1) and 2(b1), the phase distribution of a HeLa cell (0–1 rad) was used to simulate an intensity image at a defocus distance of 5 μm, defined within a grid of 256×256 pixels (pixel size is 6.5 μm). The objective NA is 0.4 (20× magnification), and the wavelength of monochromatic illumination is 550 nm. The comprehensive comparison under coherent illumination is detailed in Figs. 2(a2)–2(a6). Except for BlindNet and NFTPM (randomly given an initial value as the defocus distance, e.g., $z = 2$ μm), other methods were provided with the correct defocus value ($z = 5$ μm). It can be observed that all methods achieve accurate phase recovery due to the exact match of the physical model and the optical parameters. Specifically, NFTPM and BlindNet show the ability to correctly predict the defocus distance, enabling robust QPI even with incorrect initialization of the defocus parameter. However, since BlindNet, TIE, and GS methods ignore the effect of partial coherence on the forward image formation process, the physical priors used by these methods do not accurately apply to the circular illumination situation, leading to a significant loss of high-frequency information in the partially coherent QPI [Figs. 2(b2)–2(b4)].

In contrast, AI-TIE and NFTPM demonstrate better performance at $S = 0.85$ [Figs. 2(b5) and 2(b6)], as they establish a nonlinear forward model that conforms to partially coherent illumination by considering illumination distribution in modeling. Besides, we extended the simulated phase range to 0–6 rad (non-weak object) to validate that NFTPM is beyond weak object approximation. The results in Fig. 2(c) show that AI-TIE suffers from low-frequency underestimation, and Fig. 2(e) quantitatively reflects the inaccuracy of the phase image recovered by AI-TIE, while the result of NFTPM is consistent with the ground truth. It is worth noting that NFTPM, with a parameter count of 6×10^4 (MLP) and the inference time of 3×10^{-4} s, outperforms BlindNet in terms of speed, which requires a much larger parameter count of 3.0063×10^7 (UNet) and a comparatively longer inference time of 5.8×10^{-3} s. Furthermore, when the network size is reduced (fewer channels C per layer), the representational capacity of 2D UNet (1.786×10^6) is significantly weaker, resulting in the deterioration of phase retrieval in BlindNet [Fig. 2(d)] and the mismatch profile [Fig. 2(f)], while NFTPM (6×10^3) remains robust. In addition, as shown in Fig. S1 in Ref. [30], we also discuss the impact of hyperparameter tuning (K, L, N_θ) on NFTPM, elucidating the stability of NFTPM against the layer changes in the MLP as well as the significance of radial encoding for the high-frequency characterization (see Section 1 in Ref. [30] for detailed analysis).

B. Verification of QPI at Unknown Defocus Distances for Different Illuminations

Further simulations are shown to verify that NFTPM can accurately recover the phase without prior knowledge of the defocus distance in different partially coherent illuminations. In Figs. 3(a1)–3(a4), we simulated intensity images at $z = 7$ μm under circular illuminations ($S = 0.10$, 0.40, and 0.75) and annular NA-matched illumination. We randomly provided

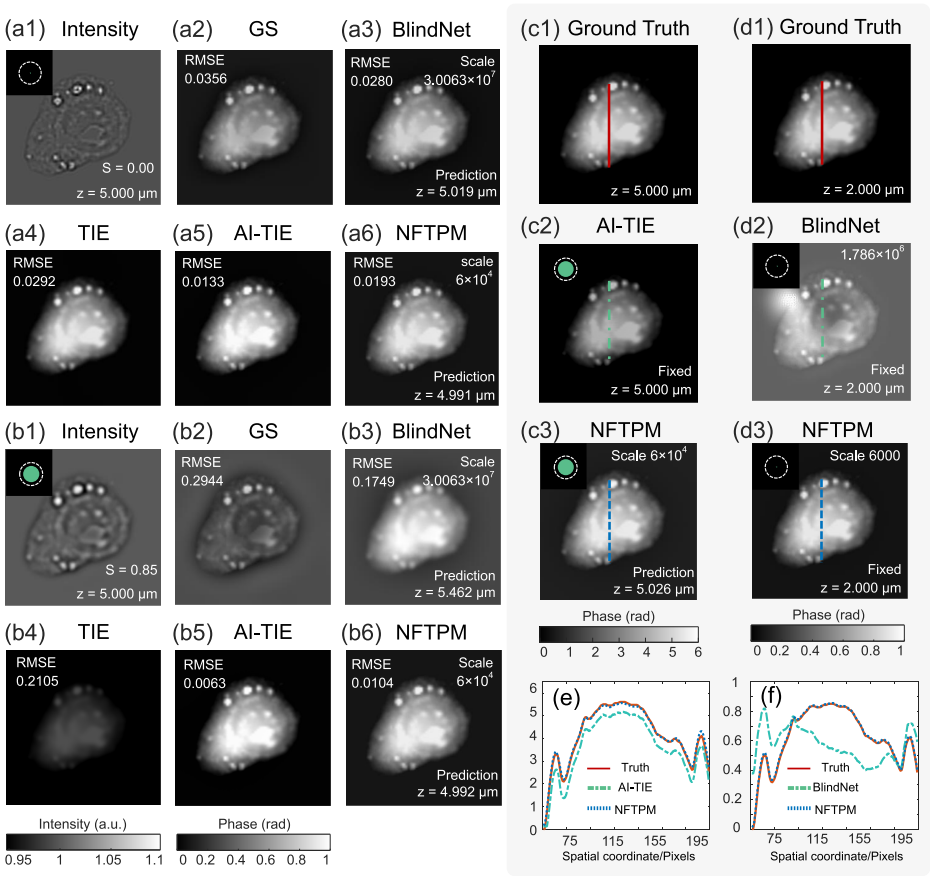

Fig. 2. Comparison of phase (0–1 rad) reconstruction results of a simulated cell sample using NFTPM, BlindNet, TIE, AI-TIE, and GS methods under different illumination settings and a defocus distance of 5 μm. (a) Results under coherent illumination. (b) Results under partially coherent illumination ($S = 0.85$). (c) Comparison between AI-TIE and NFTPM for the large phase (0–6 rad) under partially coherent illumination ($S = 0.85$). (d) Comparison between BlindNet and NFTPM after downsizing the network under coherent illumination. (e), (f) Phase line profiles corresponding to (c), (d).

an incorrect initial value of the defocus distance (10 μm) for the AI-TIE and NFTPM (the robustness of NFTPM to defocus distance initialization is verified in Fig. S2 and Section 2 of Ref. [30]) and evaluated the quality of the retrieved phase using the root mean square error (RMSE). In Figs. 3(b1)–3(b4), severe artifacts appear in the results of AI-TIE, as the phase transfer function (PTF) is mis-estimated due to the uncorrected z value. In contrast, NFTPM achieves high-precision phase retrieval (RMSE < 0.06) in diverse illuminations and accurately predicts the defocus value based on the tunable defocus parameter z. We also supplement simulations under other special illuminations (e.g., asymmetric semicircular illumination) to display the adaptability of NFTPM to arbitrary source distribution (see Fig. S3 and Section 3 of Ref. [30] for details). As depicted in Figs. 3(c1)–3(c4), the RMSE of the phase recovered by NFTPM progressively decreases with the increasing maximum illumination angle, demonstrating that NFTPM has higher imaging accuracy at large illumination angles in the absence of noise. However, under noisy conditions (Gaussian noise with a standard deviation of 0.005), as the increase in the illumination angle reduces the response amplitude of the PTF [see Figs. S4(a)–S4(c) in Ref. [30]], the sensitivity to noise instead leads to an escalation in the RMSE, as illustrated in Figs. 3(d1)–3(d3). Although the annular NA-matched illumination has a larger illumination angle compared to the circular illumination with S of 0.4 and 0.75, it has a relatively smaller RMSE in the presence of noise [Fig. 3(d4)], owing to its improved spatial frequency response that allows for higher robustness to noise [see Figs. S4(e) and S4(f) in Ref. [30]]. Essentially, NFTPM can be regarded as an iterative process that simultaneously seeks the optimal solutions for defocus distance prediction and phase retrieval. As shown in Figs. 3(f1)–3(f4), the loss function exhibits a steady decline along with a converging trend of defocus distance z (towards 7 μm in all cases), indicating the parallel optimization of the model parameters and the defocus parameter. Additionally, simulations under annular illumination for various z (5 μm, 8 μm, 12 μm, and 17 μm) in Fig. 3(g) validate the stability of NFTPM for defocus distance prediction.

C. QPI Experiments for Live HeLa Cells

In the actual experiment, the long-time imaging of living HeLa cells [Fig. 4(a)] was performed using the inverted microscope (IX83) without motor drive adjustment or manual correction. Under the influence of temperature fluctuation and other

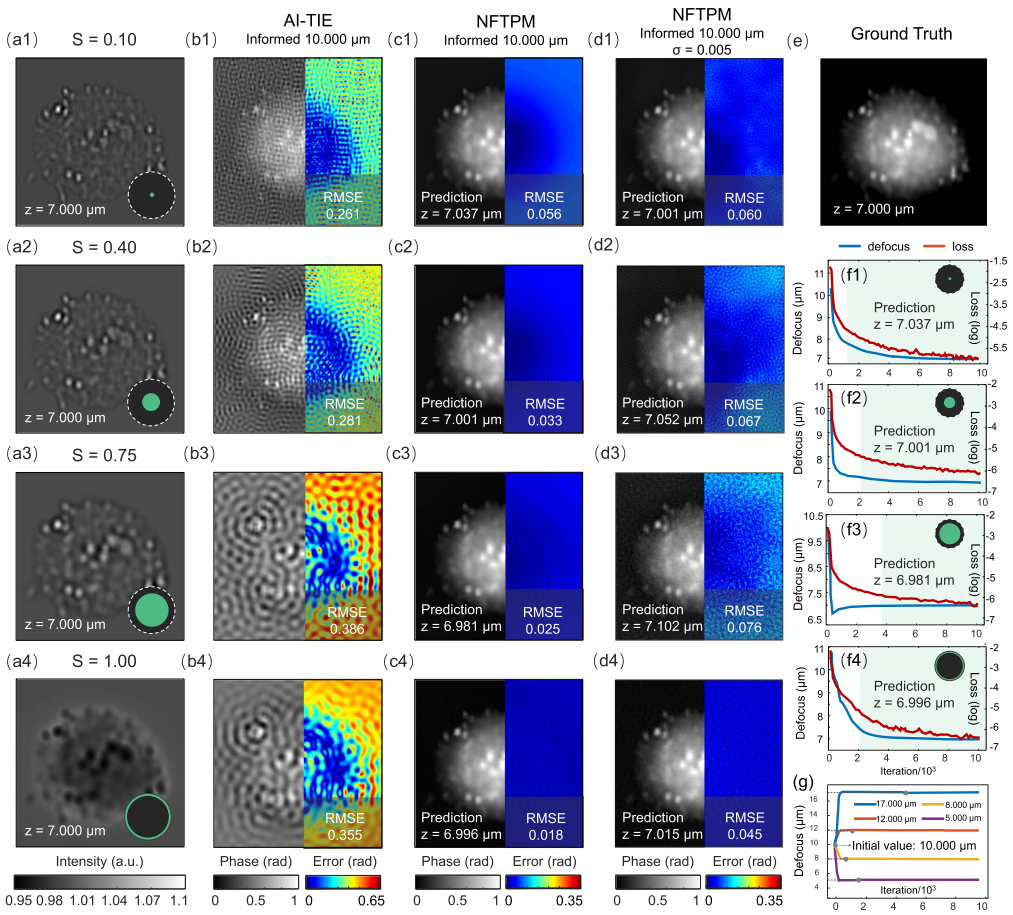

Fig. 3. Comparison between AI-TIE and NFTPM with an incorrect initial value of the defocus distance in various illuminations. (a) Intensity images under circular illuminations ($S = 0.10$, 0.40, and 0.75) and annular NA-matched illumination. (b) The results of AI-TIE. (c) The results of NFTPM. (d) The results of NFTPM in the presence of noise. (e) Ground truth. (f) Convergence curves of the defocus value and loss value (in logarithmic form). (g) The defocus distance prediction process of NFTPM for intensity images simulated at other defocus distances under annular illumination.

factors that lead to focal drift, NFTPM shows its stability of phase recovery by correctly predicting the unknown defocus distance. Figure 4(b) demonstrates the iterative process of predicting the defocus distance using the NFTPM in Area 1 and Area 2 at different moments. Taking 1 μm as the initial value, we recovered the phase of the first frame measurement by NFTPM with the predicted defocus value of 6.522 μm and adopted 6.5 μm as the defocus distance for AI-TIE. Since the defocus distance was fixed at all moments, the retrieved results of AI-TIE, as depicted in Figs. 4(c1) and 4(c2), gradually deteriorated over time due to the model mismatch induced by the time-varying defocus distance. On the contrary, NFTPM dynamically reconstructed the phase information at different moments through adaptive defocus parameter correction, revealing distinct subcellular details such as nuclei and lipid droplets [Figs. 4(d1) and 4(d2)]. Once NFTPM reconstructs the phase for the measurement at a given moment using random initialization, the QPI of subsequent frames can be accelerated by initializing NFTPM with the MLP model and the predicted z corresponding to the present frame, which utilizes the correlation between frames in the same FOV. Additionally, the efficiency of full FOV phase retrieval can be improved fivefold by utilizing the pre-iterated model of the subregion to initialize the neural field.

4. DISCUSSION AND CONCLUSION

In summary, we have proposed a new partially coherent QPI method called NFTPM using the neural field. NFTPM, a single-shot non-interferometric iterative method, employs a straightforward MLP model for continuous phase representation and can accurately predict the defocus distance without prior knowledge, which eliminates the necessity for precise motor drive adjustment or manual correction for focus drift.

The rough defocus distance provided by the focusing device can be manually adjusted to reduce the WOTF error and thus improve the reconstruction results of AI-TIE. But with the limited time and manpower required, AI-TIE still cannot be applied as an effective method for long-term live cell imaging. In contrast, NFTPM replaces the costly manual operation with gradient-based tuning, which is based on the backpropagation algorithm in the prior model and the MLP. Therefore, NFTPM can adaptively obtain reconstruction results without defocus artifacts.

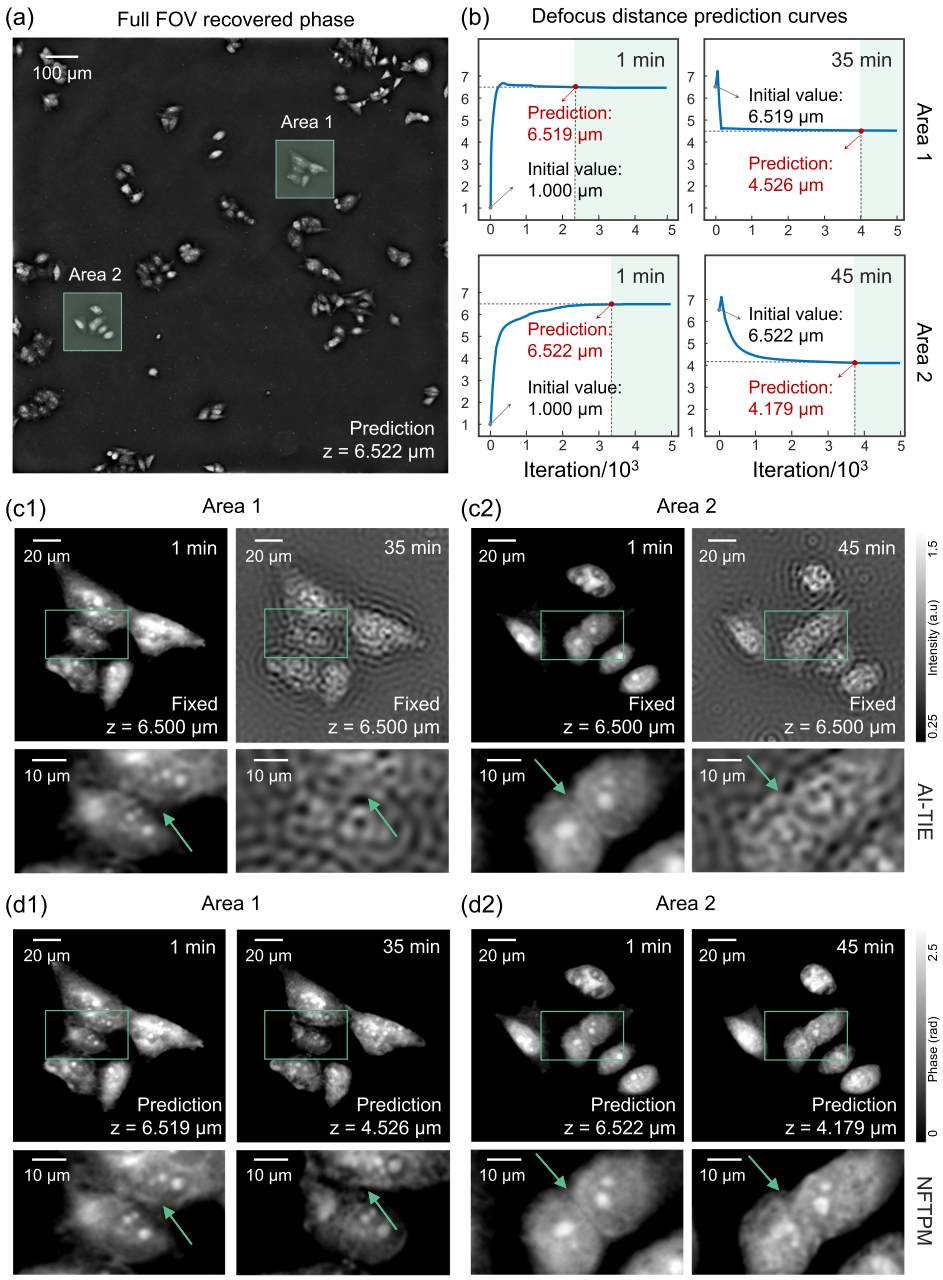

Fig. 4. Experimental observation of HeLa cells via NFTPM under annular NA-matched illumination (see Visualization 1). (a) The full FOV of the reconstruction result of NFTPM. (b) The defocus distance prediction process. (c1), (c2) The phase retrieved by AI-TIE. (d1), (d2) The phase retrieved by NFTPM.

For circular illumination, high-resolution reconstruction results from large-angle illumination, providing more high-frequency information. However, this comes at the expense of a steadily diminishing PTF response with increasing illumination angle, making noise more detrimental to the phase reconstruction of NFTPM. In contrast, annular illumination exhibits strong noise immunity due to a uniformly high response in its pass-band over a large illumination angle. To obtain high-quality QPI results, the defocus distance also needs to be selected appropriately. The low-frequency response of the PTF becomes weak when the defocus distance is too small, which is not conducive to the recovery of low-frequency information. Besides, the PTF obtained at excessive defocus distance has a low response in its pass-band and contains multiple deep dips and zero-crossings, rendering this part of information susceptible to noise (see Figs. S5 and S6 in Section 4 of Ref. [30]). Remarkably, NFTPM can also be applied in pixel-aliasing conditions by additionally introducing pixel binning as a prior. Its capability of pixel super-resolution QPI is validated in Fig. S7 (Section 5 of Ref. [30]) by simulating a pixel-aliased defocused intensity of a USAF resolution test target.

Although our method incorporates partially coherent illumination into the forward image formation model, other optical parameters that are beneficial for improving reconstruction quality are still overlooked. Therefore, in the future, more optical parameters will be considered in the physical model to further promote the quality of phase reconstruction. For instance, it is possible to achieve prediction of unknown illumination by a grid search in a preset series of coherence parameters.

Funding. National Natural Science Foundation of China (62227818, 62105151, 62175109, U21B2033); National Key Research and Development Program of China (2022YFA1205002); Leading Technology of Jiangsu Basic Research Plan (BK20192003); Youth Foundation of Jiangsu Province (BK20210338); Biomedical Competition Foundation of Jiangsu Province (BE2022847); Key National Industrial Technology Cooperation Foundation of Jiangsu Province (BZ2022039); Fundamental Research Funds for the Central Universities (30920032101, 30923010206); Fundamental Scientific Research Business Fee Funds for the Central Universities (2023102001); Open Research Fund of Jiangsu Key Laboratory of Spectral Imaging Intelligent Sense (JSGP202105, JSGP202201).

Disclosures. The authors declare no conflicts of interest.

Data Availability. The data that support the plots and maps within this paper and other findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. G. Popescu, *Quantitative Phase Imaging of Cells and Tissues* (McGraw-Hill Education, 2011).
2. X. Chang, L. Bian, and J. Zhang, "Large-scale phase retrieval," eLight **1**, 4 (2021).
3. Z. Huang, P. Memmolo, P. Ferraro, *et al.*, "Dual-plane coupled phase retrieval for non-prior holographic imaging," PhotoniX **3**, 3 (2022).
4. R. W. Gerchberg, "Phase determination from image and diffraction plane pictures in the electron microscope," Optik **34**, 275–284 (1971).
5. R. W. Gerchberg, "A practical algorithm for the determination of plane from image and diffraction pictures," Optik **35**, 237–246 (1972).
6. M. R. Teague, "Deterministic phase retrieval: a Green's function solution," J. Opt. Soc. Am. **73**, 1434–1441 (1983).
7. C. Zuo, J. Li, J. Sun, *et al.*, "Transport of intensity equation: a tutorial," Opt. Lasers Eng. **135**, 106187 (2020).
8. E. Barone-Nugent, A. Barty, and K. Nugent, "Quantitative phase-amplitude microscopy I: optical microscopy," J. Microsc. **206**, 194–203 (2002).
9. L. Lu, J. Li, Y. Shu, *et al.*, "Hybrid brightfield and darkfield transport of intensity approach for high-throughput quantitative phase microscopy," Adv. Photonics **4**, 056002 (2022).
10. C. Zuo, J. Sun, J. Li, *et al.*, "High-resolution transport-of-intensity quantitative phase microscopy with annular illumination," Sci. Rep. **7**, 7654 (2017).
11. A. Sinha, J. Lee, S. Li, *et al.*, "Lensless computational imaging through deep learning," Optica **4**, 1117–1125 (2017).
12. Y. Wu, Y. Rivenson, Y. Zhang, *et al.*, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," Optica **5**, 704–710 (2018).
13. Y. Rivenson, Y. Zhang, H. Günaydn, *et al.*, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light Sci. Appl. **7**, 17141 (2018).
14. K. Wang, J. Di, Y. Li, *et al.*, "Transport of intensity equation from a single intensity image via deep learning," Opt. Lasers Eng. **134**, 106233 (2020).
15. C. Zuo, J. Qian, S. Feng, *et al.*, "Deep learning in optical metrology: a review," Light Sci. Appl. **11**, 39 (2022).
16. E. Bostan, R. Heckel, M. Chen, *et al.*, "Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network," Optica **7**, 559–562 (2020).
17. F. Wang, Y. Bian, H. Wang, *et al.*, "Phase imaging with an untrained neural network," Light Sci. Appl. **9**, 77 (2020).
18. D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 9446–9454.
19. X. Zhang, F. Wang, and G. Situ, "Blindnet: an untrained learning approach toward computational imaging with model uncertainty," J. Phys. D **55**, 034001 (2021).
20. B. Mildenhall, P. P. Srinivasan, M. Tancik, *et al.*, "NeRF: representing scenes as neural radiance fields for view synthesis," Commun. ACM **65**, 99–106 (2021).
21. R. Liu, Y. Sun, J. Zhu, *et al.*, "Recovery of continuous 3D refractive index maps from discrete intensity-only measurements using neural fields," Nat. Mach. Intell. **4**, 781–791 (2022).
22. H. Zhu, Z. Liu, Y. Zhou, *et al.*, "Dnf: diffractive neural field for lensless microscopic imaging," Opt. Express **30**, 18168–18178 (2022).
23. H. Zhou, B. Y. Feng, H. Guo, *et al.*, "Fourier ptychographic microscopy image stack reconstruction using implicit neural representations," Optica **10**, 1679–1687 (2023).
24. L. Lu, Y. Fan, J. Sun, *et al.*, "Accurate quantitative phase imaging by the transport of intensity equation: a mixed-transfer-function approach," Opt. Lett. **46**, 1740–1743 (2021).
25. E. Abbe, "Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung," Arch. Mikroskopische Anatomie **9**, 413–468 (1873).
26. S. Zhou, J. Li, J. Sun, *et al.*, "Transport-of-intensity Fourier ptychographic diffraction tomography: defying the matched illumination condition," Optica **9**, 1362–1373 (2022).
27. Y. Shu, J. Sun, J. Lyu, *et al.*, "Adaptive optical quantitative phase imaging based on annular illumination Fourier ptychographic microscopy," PhotoniX **3**, 24 (2022).
28. C. J. Sheppard, "Three-dimensional phase imaging with the intensity transport equation," Appl. Opt. **41**, 5951–5955 (2002).
29. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv, arXiv:1412.6980 (2014).
30. "Supplemental document for 'Neural-field-assisted transport-of-intensity phase microscopy: partially coherent quantitative phase imaging under unknown defocus distance'," Figshare, 2024, https://figshare.com/articles/journal_contribution/Supplementary_document_for_Neural-field-assisted_transport-of-intensity_phase_microscopy_partially_coherent_quantitative_phase_imaging_under_unknown_defocus_distance/25835650.
31. D. Paganin, S. C. Mayo, T. E. Gureyev, *et al.*, "Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object," J. Microsc. **206**, 33–40 (2002).