



Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network

EMRAH BOSTAN,^{1,*} REINHARD HECKEL,² MICHAEL CHEN,¹  MICHAEL KELLMAN,¹  AND LAURA WALLER¹ 

¹Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California 94720, USA

²Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany

*Corresponding author: emrah.bostan@gmail.com

Received 28 January 2020; revised 28 March 2020; accepted 21 April 2020 (Doc. ID 389314); published 21 May 2020

Deep neural networks have emerged as effective tools for computational imaging, including quantitative phase microscopy of transparent samples. To reconstruct phase from intensity, current approaches rely on supervised learning with training examples; consequently, their performance is sensitive to a match of training and imaging settings. Here we propose a new approach to phase microscopy by using an untrained deep neural network for measurement formation, encapsulating the image prior and the system physics. Our approach does not require any training data and simultaneously reconstructs the phase and pupil-plane aberrations by fitting the weights of the network to the captured images. To demonstrate experimentally, we reconstruct quantitative phase from through-focus intensity images without knowledge of the aberrations. © 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

<https://doi.org/10.1364/OPTICA.389314>

Quantitative phase microscopy (QPM) enables label-free imaging of transparent samples such as unstained cells and tissues [1,2] and non-absorbing micro-elements [3]. QPM with partially coherent illumination provides improved spatial resolution and light throughput with reduced speckle. Examples include through-focus [4,5], interferometric [6,7], and angle-scanning [8,9] microscopes. All of these methods capture nonlinear (intensity) measurements and recover quantitative phase computationally. Generally, the performance and image quality is intrinsically governed by the phase reconstruction step [10].

Traditionally, the phase reconstruction inverse problem is solved by minimizing a least-squares loss that is based on the physics of the problem. This physics-based optimization approach is fundamental to phase imaging [11] and has the immediate advantage that prior assumptions on the images can be directly integrated through regularization. For example, one can constrain the phase image to admit a sparse representation in the wavelet domain [12]. Such regularizers work well and improve the reconstruction quality [7,13]. Another major advantage of the physics-based formulation is the possibility to incorporate algorithmic self-calibration [14,15]. It involves—in alternation with the phase retrieval step—minimizing the least-squares loss over unknown

or partially known system parameters such as pupil aberrations [10]. The concept hence accounts for the model-mismatch in the imaging pipeline. This provides us with great flexibility and allows phase reconstruction from measurements that are not fully characterized. In designing self-calibrating algorithms, the need for regularization (i.e., prior models for phase) is emphasized [16], since one simultaneously decouples the individual contributions of phase and aberrations to the measured images. However, typical regularization techniques are hand-crafted and require manual tuning of parameters.

More recently, deep neural networks (DNNs) trained in an end-to-end fashion on large datasets have been used for phase retrieval, directly mapping measured intensities back to phase. Trained DNNs give state-of-the-art performance in holographic [17], lensless [18], ptychographic [19], and through-scattering-media [20,21] phase retrieval configurations, among others [22]. The results validate the efficiency of properly trained DNNs to solve nonlinear inverse problems and shift the computational paradigm in QPM towards data-driven frameworks. However, for deep networks to work well, the proximity of training and experiment settings is critical, as the performance is susceptible to variations in sample features, instrumentation, and acquisition parameters [22]. Although improved DNN architectures have been proposed [23–26], training-based approaches are sensitive to misfits, in that they fundamentally rely on the phase distribution being close to one of the training images.

Here we propose a new QPM algorithm that is based on a deep network, but requires no ground-truth training data. Our approach is inspired by the idea of employing untrained generative DNNs as prior models for images, a concept pioneered by the so-called deep image prior [27]. Specifically, Ulyanov *et al.* [27] fitted a noisy image via optimizing over the weights of a randomly initialized, over-parameterized autoencoder (i.e., an autoencoder with more weights than the number of image pixels), and observed that early stopping of the regularization yielded good denoising performance, an effect theoretically explained in [28]. For denoising, regularization through early stopping is critical, since the network can in principle fit the noisy data perfectly. Subsequently, an under-parameterized image-generating network was proposed, named the deep decoder, that does not need early stopping or any other further regularization [29]. The framework acts as a

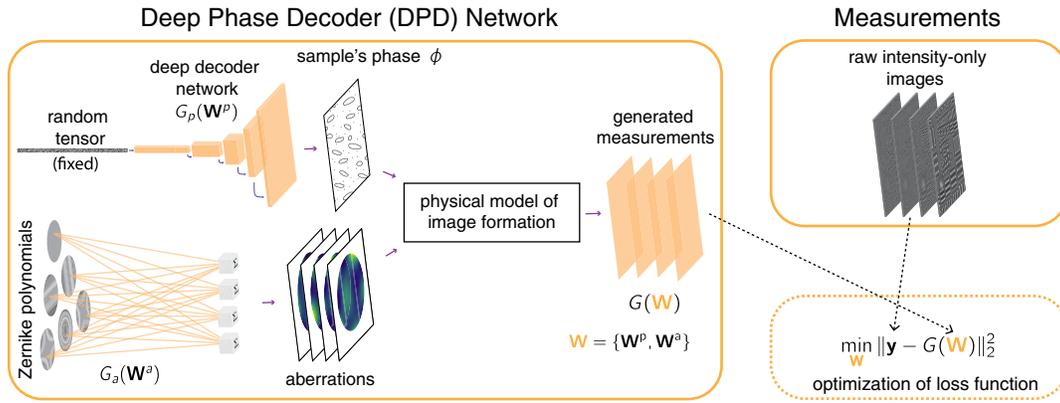


Fig. 1. Our deep phase decoder (DPD) algorithm aims to minimize the Euclidean distance between the measured intensity images and the hypothetical ones generated by our untrained deep network. The optimization problem, which is nonlinear and nonconvex, is stated in terms of the network's weights and is solved iteratively using a gradient-based procedure. Once the weights are optimized, the sought phase image is retrieved as the output of the deep decoder part of the network.

concise image model that provides a lower-dimensional description of images, akin to the sparse wavelet representations, and thus regularizes through its architecture. Unfortunately, a naive application of the method would not account for practical issues such as drift and sample-induced aberrations [16], which require properly incorporating our knowledge about optical physics in the self-calibration.

Our key contribution is a DNN-based self-calibrating reconstruction algorithm for QPM that is training-free and recovers quantitative phase from raw images recorded without the explicit knowledge of aberrations. We specify the measurement formation as an untrained DNN whose weights are fitted to the recorded images (Fig. 1). Leveraging the well-characterized system physics and nonlinear forward model, our network combines a fully connected layer that synthesizes aberrations from Zernike polynomials with the deep decoder that is used to generate phase. The proposed algorithm hence describes both the image and aberrations by a few weight coefficients, and as a consequence enables us to jointly retrieve the phase and individual aberration profile of each measurement without training data. We call our algorithm the deep phase decoder (DPD) and demonstrate it on a commercial widefield microscope.

We describe the image formation process for our optical setup (Fig. 2) and then describe our reconstruction approach in more detail. We consider an optically thin and transparent sample placed at the focal plane of the microscope's objective. The sample's complex-valued image (i.e., its transmission function) is characterized as

$$o(\mathbf{r}) = \exp(j\phi(\mathbf{r})), \quad (1)$$

where ϕ represents the spatial distribution of phase over 2D coordinates \mathbf{r} . The LED illumination is placed sufficiently far away enough that the illumination can be modeled as a monochromatic plane wave at the sample plane. Thus, the irradiance of the beam impinging on the sensor is given by

$$y(\mathbf{r}) = |c_{\text{psf}} * o|^2(\mathbf{r}), \quad (2)$$

where $*$ denotes spatial convolution and c_{psf} is the coherent point-spread function of the microscope. The sensor then measures the

sampled irradiance, $\mathbf{y} \in \mathbb{R}^p$, where p is the total number of pixels on it. In matrix form,

$$\mathbf{y} = |\mathbf{F}^{-1} \mathbf{P}_{\text{circ}} \mathbf{F} \mathbf{o}|^2, \quad (3)$$

where \mathbf{F} is the discrete Fourier transform matrix and \mathbf{P}_{circ} is the ideal and space-invariant exit pupil function, which is a circle with its radius determined by the numerical aperture (NA) of the objective and wavelength λ .

Phase is recovered based on multiple images with some type of data diversity that translates phase information into intensity (e.g., illumination coding [30] and pupil coding [4, 31]). Here we adopt a pupil-coding scheme where the wavefront at the exit pupil [11] is aberrated differently for each measurement. The pupil aberration is modeled as a weighted sum of Zernike polynomials, parameterized by a small number of coefficients:

$$\mathbf{P} = \mathbf{P}_{\text{circ}} \exp(j\mathbf{Z}\mathbf{c}), \quad (4)$$

where the Zernike basis $\mathbf{Z} = [\mathbf{z}_1 \mathbf{z}_2 \dots \mathbf{z}_M]$ is composed of M orthogonal modes in vectorized form, and \mathbf{c} contains the corresponding coefficients of each mode.

The microscope is probed with a known (or pre-calibrated) set of aberrations $\{\mathbf{P}\}_{n=1}^N$, where N is the total number of intensity images. The inverse problem then aims to recover the sample's transmission function as

$$\mathbf{o}^* = \arg \min_{\mathbf{o}} \sum_{n=1}^N \|\sqrt{y_n} - |\mathbf{F}^{-1} \mathbf{P}_n \mathbf{F} \mathbf{o}|\|_2^2. \quad (5)$$

This can be solved by gradient descent (or an accelerated variation), which is closely related to the well-known Gerchberg–Saxton method [10]. After solving for \mathbf{o}^* , the phase image is its argument. The conventional phase recovery in Eq. (5) does not necessarily impose any regularization on the recovered phase, and the aberrations must be known *a priori*. To address these without needing any training data, we introduce a deep network in the derived formulation.

At the core of our approach, we use a DNN that generates N intensity images. The network, denoted by $G(\mathbf{W})$, reparameterizes the measurement formation in Eq. (3) in terms of a weight tensor

\mathbf{W} rather than pixels in complex-image space as in Eq. (5). The network is untrained, and the weights, which are randomly initialized, are optimized by solving the following problem:

$$\mathbf{W}^* = \arg \min_{\mathbf{W}} \|\mathbf{Y} - G(\mathbf{W})\|_2^2, \quad (6)$$

where $\mathbf{Y} = [\sqrt{y_1} \dots \sqrt{y_N}] \in \mathbb{R}^{p \times N}$ accommodates the measured data. Once the optimal weights \mathbf{W}^* are obtained, phase is retrieved as the output of an appropriate layer in the network. The remarkable aspect—in terms of data requirement—is that the process is driven solely by the acquired images, without any training data. The main reason is that the generative network's weights (and hence its output image) are adjusted on-the-fly in Eq. (6) rather than training it for a certain class of images.

We design the network G to encapsulate two sub-generators, G_p and G_a , that synthesize a phase image and the pupil aberration of each individual measurement, respectively (see Fig. 1). For the phase generating network G_p , we use a deep decoder [29], which transforms a randomly chosen and fixed tensor $\mathbf{B}_0 \in \mathbb{R}^{n_0 \times k}$ consisting of k many n_0 -dimensional channels to an $n_d \times 1$ dimensional (gray-scale) image. In transforming the random tensor to a phase image, G_p applies i) a pixel-wise linear combination of the channels, ii) upsampling, iii) rectified linear units (ReLU), and iv) channel normalization. Specifically, the update at the $(i + 1)$ -th layer is given by

$$\mathbf{B}_{i+1} = \text{cn}(\text{ReLU}(\mathbf{U}_i \mathbf{B}_i \mathbf{W}_i^p)), \quad i = 0, \dots, d - 1. \quad (7)$$

Here $\mathbf{W}_i^p \in \mathbb{R}^{k \times k}$ contains the coefficients for the linear combination of the channels, and the operator $\mathbf{U}_i \in \mathbb{R}^{n_{i+1} \times n_i}$ performs bi-linear upsampling. This is followed by a channel normalization operation, $\text{cn}(\cdot)$, which is equivalent to normalizing each channel individually to zero mean and unit variance, plus a bias term. A phase image, which is the output of the d -layer network, is then formed, with $\mathbf{W}_d^p \in \mathbb{R}^k$, as

$$\phi = 2\pi \text{sigmoid}(\mathbf{B}_d \mathbf{W}_d^p). \quad (8)$$

The aberration-generating network, $G_a(\mathbf{W}^a)$, relies on the parameterization in Eq. (4) represented as a fully connected layer (i.e., linear combination of Zernike modes), and the matrix \mathbf{W}^a contains the Zernike coefficients for all measurements. In combining the outputs of G_p and G_a , we reproduce the physical image formation using Eqs. (1), (3), and (4) in the network's architecture. The framework is implemented in *PyTorch*, allowing us to solve Eq. (6) using gradient-based algorithms thanks to auto-differentiation with respect to $\mathbf{W} = \{\mathbf{W}^p, \mathbf{W}^a\}$. Once the optimal weights \mathbf{W}^* are obtained, the reconstructed phase is given by $G_p(\mathbf{W}^{p*})$, where $\mathbf{W}^p = \{\mathbf{W}_0^p, \dots, \mathbf{W}_d^p\}$.

We now explain some implicit aspects of our method. First, we see from Eq. (6) that $G(\mathbf{W}^*)$ replicates the recorded intensities as closely as possible in the least-squares sense. Specifically, both G_p and G_a under-parameterize their corresponding outputs (fewer weights than the number of pixels in generated images and aberrations), so DPD imposes regularization on both phase and aberrations. Regularization is governed by the architecture of G_p , for phase has to lie in its range. Once G_p is constructed, the strength of regularization is not hand-tuned, as is typically done (e.g., adjusting the sparsity level for wavelet-based methods). As for G_a , it generates aberrations from randomly

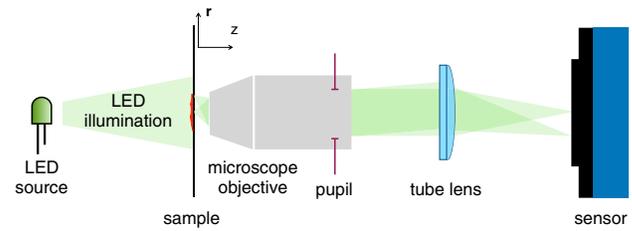


Fig. 2. We experimentally validate our method in a microscope by capturing intensity images with varying defocus distances. The images are fed into the DPD algorithm to computationally reconstruct the sample's phase and wavefront aberrations without knowing the pupil functions (defocus distances) that were used during acquisition.

initialized Zernike coefficients in contrast to other self-calibrating schemes using theoretical pupils as initialization [10,16]. Finally, since the retrieved aberrations are spanned by the Zernike polynomials, DPD is applicable when this physical model is valid [16].

To experimentally validate our method, we use a commercial brightfield microscope (Nikon TE300) with LED illumination ($\lambda = 0.514 \mu\text{m}$) [9]. A phase target (Benchmark Technologies) is imaged by a 40×0.65 NA objective lens, and intensity images are captured by a PCO.edge 5.5 sCMOS camera placed on the front port of the microscope (which adds $2 \times$ magnification). Pupil coding is achieved practically by defocus; we capture a through-focus stack of eight images that are exponentially spaced (at 0, 1, 2, 4, 8, 16, 32, and $64 \mu\text{m}$ defocus) [33]. For comparison, we also reconstruct reference phase images with the accelerated Wirtinger flow algorithm [32] using Eq. (5) for eight and four (defocus of 4, 8, 16, and $32 \mu\text{m}$) measurements. We then use the same four measurements with our DPD method, using the RMSProp algorithm with 5×10^4 iterations (GPU time is ~ 25 minutes). Based on the observations in [29], the network is constructed with the following parameters: $k = 32$, $n_0 = 16 \times 16$, and $n_d = 512 \times 512$. \mathbf{B}_0 is randomly drawn from a uniform distribution in the range [0, 0.1], and bi-linear upsampling is fixed to a factor of two, making G_p a six-layer network. We use the first nine Zernike polynomials after piston for G_a . The reconstructions from Wirtinger flow (with known defocus distances) and our DPD method (without knowing the defocus distances) both show good agreement with the known phase profile of the target (Fig. 3). DPD jointly recovers defocus-like pupil functions, as expected. This validates our algorithm's ability to blindly reconstruct a reliable phase image, as well as the pupil aberrations, from the measured intensities.

In summary, we derived a new phase imaging algorithm that uses an untrained neural network, and demonstrated it on a phase-from-defocus dataset. Our DPD method, unlike its deep learning counterparts that are supervised, is training-free and does not require closely matching training and experiment conditions. Moreover, our method is self-calibrating, allowing us to directly reconstruct high-quality phase without *a priori* knowledge of the system's aberrations.

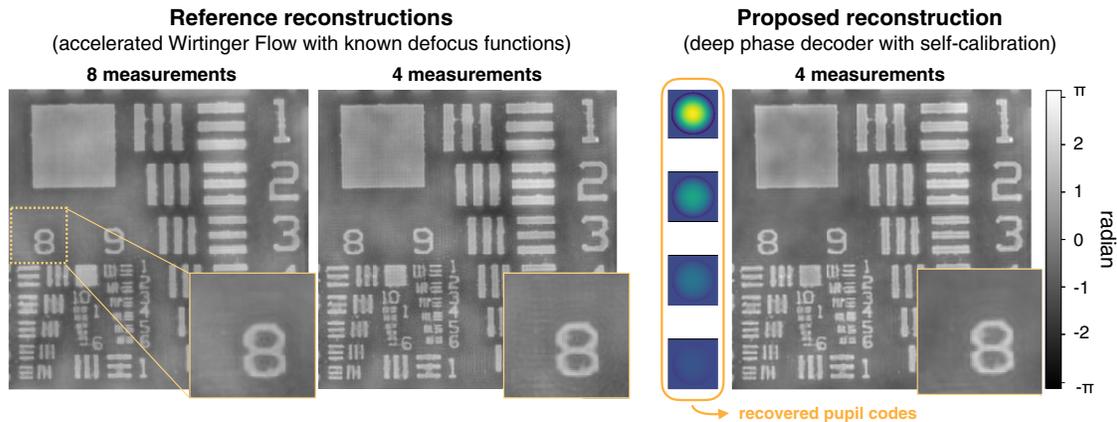


Fig. 3. Experimental validation of our DPD phase retrieval method from a stack of through-focus intensity images of a phase target with expected height of 150 nm (0.95 radians phase shift). (Left) Reconstructions by the accelerated Wirtinger flow algorithm [32] are shown for comparison, with different numbers of measurements in the focus stack and known defocus distances. (Right) Our proposed DPD reconstruction achieves a similar phase result without any explicit knowledge of the aberrations.

Funding. Swiss National Science Foundation (P2ELP2 17227); National Science Foundation (DMR 154892, DGE 1106400, ONR N00014-17-1-2401).

Acknowledgment. The authors thank Gautam Gunjala for helpful discussions.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

- P. Marquet, B. Rappaz, P. J. Magistretti, É. Cuche, Y. Emery, T. Colomb, and C. Depeursinge, *Opt. Lett.* **30**, 468 (2005).
- G. Popescu, *Quantitative Phase Imaging of Cells and Tissues* (McGraw-Hill, 2011).
- A. Barty, K. A. Nugent, D. Paganin, and A. Roberts, *Opt. Lett.* **23**, 817 (1998).
- L. Waller, L. Tian, and G. Barbastathis, *Opt. Express* **18**, 12552 (2010).
- A. Descloux, K. Grubmayer, E. Bostan, T. Lukes, A. Bouwens, A. Sharipov, S. Geissbuehler, A.-L. Mahul-Mellier, H. Lashuel, M. Leutenegger, and T. Lasser, *Nat. Photonics* **12**, 165 (2018).
- Z. Wang, L. Millet, M. Mir, H. Ding, S. Unarunotai, J. Rogers, M. U. Gillette, and G. Popescu, *Opt. Express* **19**, 1016 (2011).
- T. Kim, R. Zhou, M. Mir, S. D. Babacan, P. S. Carney, L. L. Goddard, and G. Popescu, *Nat. Photonics* **8**, 256 (2014).
- G. Zheng, R. Horstmeyer, and C. Yang, *Nat. Photonics* **7**, 739 (2013).
- L. Tian and L. Waller, *Optica* **2**, 104 (2015).
- L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, *Opt. Express* **23**, 33214 (2015).
- J. R. Fienup, *Appl. Opt.* **21**, 2758 (1982).
- A. Pein, S. Loock, G. Plonka, and T. Salditt, *Opt. Express* **24**, 8332 (2016).
- E. Bostan, E. Froustey, M. Nilchian, D. Sage, and M. Unser, *IEEE Trans. Image Process.* **25**, 807 (2016).
- X. Ou, G. Zheng, and C. Yang, *Opt. Express* **22**, 4960 (2014).
- Z. Jingshan, L. Tian, J. Dauwels, and L. Waller, *Biomed. Opt. Express* **6**, 257 (2015).
- M. Chen, Z. F. Phillips, and L. Waller, *Opt. Express* **26**, 32888 (2018).
- Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, *Light Sci. Appl.* **7**, 17141 (2018).
- A. Sinha, J. Lee, S. Li, and G. Barbastathis, *Optica* **4**, 1117 (2017).
- T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, *Opt. Express* **26**, 26470 (2018).
- S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, *Optica* **5**, 803 (2018).
- N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, *Optica* **5**, 960 (2018).
- Y. Jo, H. Cho, S. Y. Lee, G. Choi, G. Kim, H. Min, and Y. Park, *IEEE J. Sel. Top. Quantum Electron.* **25**, 1 (2019).
- Y. Li, Y. Xue, and L. Tian, *Optica* **5**, 1181 (2018).
- Y. Xue, S. Cheng, Y. Li, and L. Tian, *Optica* **6**, 618 (2019).
- M. Kellman, E. Bostan, N. Repina, and L. Waller, *IEEE Trans. Comput. Imaging* (IEEE, 2019).
- M. Kellman, E. Bostan, M. Chen, and L. Waller, "Data-driven design for Fourier ptychographic microscopy," in *Proceedings of the IEEE International Conference for Computational Photography* (2019). In press
- D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2018), pp. 9446–9454.
- R. Heckel and M. Soltanolkotabi, "Denoising and regularization via exploiting the structural bias of convolutional generators," in *International Conference on Learning Representations (ICLR)* (2020).
- R. Heckel and P. Hand, "Deep decoder: concise image representations from untrained non-convolutional networks," in *International Conference on Learning Representations (ICLR)* (2019).
- L. Tian, X. Li, K. Ramchandran, and L. Waller, *Biomed. Opt. Express* **5**, 2376 (2014).
- R. Horisaki, Y. Ogura, M. Aino, and J. Tanida, *Opt. Lett.* **39**, 6466 (2014).
- E. J. Candès, X. Li, and M. Soltanolkotabi, *IEEE Trans. Inf. Theory* **61**, 1985 (2015).
- Z. Jingshan, R. A. Claus, J. Dauwels, L. Tian, and L. Waller, *Opt. Express* **22**, 10661 (2014).