

# Transport of Intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes

Zhong Jingshan,<sup>1,2,\*</sup> Rene A. Claus,<sup>3</sup> Justin Dauwels,<sup>1</sup> Lei Tian,<sup>2</sup> and Laura Waller<sup>2</sup>

<sup>1</sup> School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

<sup>2</sup> Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA

<sup>3</sup> Applied Science and Technology, University of California, Berkeley, USA

\*[zhongjingshan@hotmail.com](mailto:zhongjingshan@hotmail.com)

**Abstract:** We propose an alternative method for solving the Transport of Intensity equation (TIE) from a stack of through-focus intensity images taken by a microscope or lensless imager. Our method enables quantitative phase and amplitude imaging with improved accuracy and reduced data capture, while also being computationally efficient and robust to noise. We use prior knowledge of how intensity varies with propagation in the spatial frequency domain in order to constrain a fitting algorithm [Gaussian process (GP) regression] for estimating the axial intensity derivative. Solving the problem in the frequency domain inspires an efficient measurement scheme which captures images at exponentially spaced focal steps, significantly reducing the number of images required. Low-frequency artifacts that plague traditional TIE methods can be suppressed without an excessive number of captured images. We validate our technique experimentally by recovering the phase of human cheek cells in a brightfield microscope.

© 2014 Optical Society of America

**OCIS codes:** (100.5070) Phase retrieval; (100.3010) Image reconstruction techniques.

---

## References and links

1. G. Popescu, *Quantitative phase imaging of cells and tissues* (McGraw-Hill New York, 2011).
2. K. Nugent, D. Paganin, and T. Gureyev, "A phase odyssey," *Physics Today* **54**, 27–32 (2001).
3. M. R. Teague, "Deterministic phase retrieval: a Greens function solution," *J. Opt. Soc. Am.* **73**, 1434–1441 (1983).
4. N. Streibl, "Phase imaging by the transport equation of intensity," *Opt. Commun.* **49**, 6–10 (1984).
5. L. J. Allen and M. P. Oxley, "Phase retrieval from series of images obtained by defocus variation," *Opt. Commun.* **199**, 65–75 (2001).
6. M. Soto and E. Acosta, "Improved phase imaging from intensity measurements in multiple planes," *Appl. Opt.* **46**, 7978–7981 (2007).
7. L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Opt. Express* **18**, 12552–12561 (2010).
8. L. Waller, M. Tsang, S. Ponda, S. Yang, and G. Barbastathis, "Phase and amplitude imaging from noisy images by Kalman filtering," *Opt. Express* **19**, 2805–2814 (2011).
9. Z. Jingshan, J. Dauwels, M. A. Vázquez, and L. Waller, "Sparse ACEKF for phase reconstruction," *Opt. Express* **21**, 18125–18137 (2013).
10. L. Waller, S. S. Kou, C. J. R. Sheppard, and G. Barbastathis, "Phase from chromatic aberrations," *Opt. Express* **18**, 22817–22825 (2010).

11. L. Waller, Y. Luo, S. Y. Yang, and G. Barbastathis, "Transport of intensity phase imaging in a volume holographic microscope," *Opt. Lett.* **35**, 2961–2963 (2010).
12. P. M. Blanchard and A. H. Greenaway, "Simultaneous multiplane imaging with a distorted diffraction grating," *Appl. Opt.* **38**, 6692–6699 (1999).
13. S. S. Gorthi and E. Schonbrun, "Phase imaging flow cytometry using a focus-stack collecting microscope," *Opt. Lett.* **37**, 707–709 (2012).
14. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Opt. Lett.* **3**, 27–29 (1978).
15. J. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
16. D. Paganin and K. A. Nugent, "Noninterferometric phase imaging with partially coherent light," *Phys. Rev. Lett.* **80**, 2586–2589 (1998).
17. T. Gureyev and K. Nugent, "Rapid quantitative phase imaging using the transport of intensity equation," *Opt. Commun.* **133**, 339–346 (1997).
18. T. Gureyev, A. Pogany, D. Paganin, and S. Wilkins, "Linear algorithms for phase retrieval in the Fresnel region," *Opt. Commun.* **231**, 53–70 (2004).
19. M. Soto, E. Acosta, and S. Ríos, "Performance analysis of curvature sensors: optimum positioning of the measurement planes," *Opt. Express* **11**, 2577–2588 (2003).
20. C. Zuo, Q. Chen, Y. Yu, and A. Asundi, "Transport-of-intensity phase imaging using Savitzky-Golay differentiation filter-theory and applications," *Opt. Express* **21**, 5346–5362 (2013).
21. R. W. Schafer, "What is a Savitzky-Golay filter?[lecture notes]," *Signal Processing Magazine, IEEE* **28**, 111–117 (2011).
22. S. Zabner, P. Cloetens, J. Guigay, J. Baruchel, and M. Schlenker, "Optimization of phase contrast imaging using hard x rays," *Review of Scientific Instruments* **76**, 073705–073705 (2005).
23. J. Guigay, M. Langer, R. Boistel, and P. Cloetens, "Mixed transfer function and transport of intensity approach for phase retrieval in the Fresnel region," *Opt. Lett.* **32**, 1617–1619 (2007).
24. D. Paganin, T. E. Gureyev, K. M. Pavlov, R. A. Lewis, and M. Kitchen, "Phase retrieval using coherent imaging techniques based on the contrast transfer function," *Opt. Commun.* **234**, 87–105 (2004).
25. C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning* (the MIT Press, 2006).
26. B. Xue, S. Zheng, L. Cui, X. Bai, and F. Zhou, "Transport of intensity phase imaging from multiple intensities measured in unequally-spaced planes," *Opt. Express* **19**, 20244–20250 (2011).
27. K. Falaggis, T. Kozacki, and M. Kujawinska, "Optimum plane selection criteria for single beam phase retrieval techniques based on the contrast transfer function," *Opt. Lett.* **39**, 30–33 (2014).
28. N. Loomis, L. Waller, and G. Barbastathis, "High-speed phase recovery using chromatic transport of intensity computation in graphics processing units," in "Biomedical Optics and 3-D Imaging," (Optical Society of America, 2010), p. JMA7.
29. D. J. Lee, M. C. Roggemann, and B. M. Welsh, "Cramér-Rao analysis of phase-diverse wave-front sensing," *J. Opt. Soc. Am. A* **16**, 1005–1015 (1999).
30. L. Tian, J. C. Petrucci, and G. Barbastathis, "Nonlinear diffusion regularization for transport of intensity phase imaging," *Opt. Lett.* **37**, 4131–4133 (2012).
31. A. Kostenko, K. J. Batenburg, A. King, S. E. Offerman, and L. J. van Vliet, "Total variation minimization approach in in-line X-ray phase-contrast tomography," *Opt. Express* **21**, 12185–12196 (2013).
32. L. Tian, J. C. Petrucci, Q. Miao, H. Kudrolli, V. Nagarkar, and G. Barbastathis, "Compressive X-ray phase tomography based on the transport of intensity equation," *Opt. Lett.* **38**, 3418–3421 (2013).
33. P. Sollich and C. K. I. Williams, "Using the equivalent kernel to understand Gaussian process regression," in "Advances in Neural Information Processing Systems 17," (the MIT Press, 2005), pp. 1313–1320.

## 1. Introduction

Quantitative phase imaging has found useful applications in biology, surface metrology and X-ray imaging [1, 2]. Methods that use a series of through-focus intensity images (e.g., [3–9]) are especially popular due to their experimental simplicity. In-focus intensity images contain no phase information; however, defocus introduces phase contrast. In fact, any imaging system with a complex transfer function will provide some phase contrast. These images can then be inverted to recover phase and amplitude quantitatively. In defocus-based methods, the experimental procedure involves simply moving the sample (or camera) axially while capturing multiple images through-focus (see Fig. 1), or using any of the recently proposed schemes for simultaneous multi-plane capture [10–13].

Recovering phase (and amplitude) from a series of defocused images is a nonlinear problem because intensity is bilinear with phase and amplitude. One of the most successful methods for this inversion is the iterative method, which takes a nonlinear convex optimization approach [14,

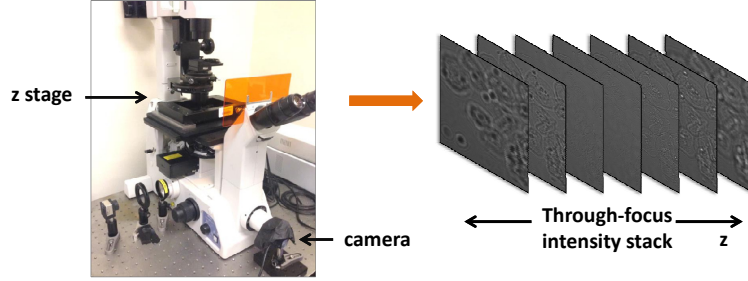


Fig. 1. Generalized experimental setup for an imaging system (e.g. a microscope) which captures intensity images at a range of axial defocus distances in order to recover phase.

15]. Though the problem is non-convex, results usually converge, especially when multiple images are used at various focal planes [5]. In another approach, where a direct solution is desired, the problem is linearized for small defocus distances using the Transport of Intensity Equation (TIE) [3, 4]. The TIE relates phase and amplitude to variations of intensity along the optical axis  $z$  [3]:

$$\left. \frac{\partial I(x, y, z)}{\partial z} \right|_{z=0} = -\frac{\lambda}{2\pi} \nabla_{\perp} \cdot [I(x, y, 0) \nabla_{\perp} \phi(x, y)], \quad (1)$$

where  $I(x, y, 0)$  is the intensity at focus,  $\phi(x, y)$  is the phase,  $\lambda$  is the spectrally-weighted mean wavelength of illumination and  $\nabla_{\perp}$  denotes the gradient operator in lateral directions  $(x, y)$  only. Besides being simple, the TIE method achieves phase accuracy equivalent to that of interferometric methods, but it is much less sensitive to coherence in the illumination [16]. Numerical solvers require no phase unwrapping and are computationally efficient, employing FFT-based inversion solvers [17].

Numerical solutions for the TIE can be understood by considering the case of a pure-phase object, where  $I(x, y, 0)$  is constant [18]. In this situation, the right hand side of Eq. (1) simplifies to a second derivative (Laplacian). Phase can then be recovered from a Laplacian inversion

$$\Phi(u, v) = F(u, v) / [-4\pi^2(u^2 + v^2)], \quad (2)$$

where  $\Phi(u, v)$  is the Fourier transform of  $\phi(x, y)$ ,  $F(u, v)$  is the Fourier transform of the measured first derivative  $\left. \frac{\partial I(x, y, z)}{\partial z} \right|_{z=0}$  scaled by  $2\pi/\lambda$ , and  $u, v$  are the spatial frequency variables. When  $I(x, y, 0)$  is not constant, two Laplacian inversions are required [3]. Note that the denominator of Eq. (2) goes towards zero as the spatial frequency goes towards zero, and a small regularization parameter must be added in order to avoid division-by-zero instability. This means that the DC phase term is lost and low frequency noise is amplified, resulting in phase errors that give cloudy phase results. We discuss here strategies for mitigating these problems.

Limitations of the TIE method mainly stem from noise and nonlinearity in the intensity derivative estimate, both of which we address here. Nonlinearity error comes from the estimation of the intensity derivative,  $\left. \frac{\partial I(x, y, z)}{\partial z} \right|_{z=0}$ , from through focus images. We cannot measure this linear (first) derivative directly and so it is usually estimated by finite difference methods. However, intensity is not linear through focus, due to diffraction, and so any nonlinearity in the through focus intensity corrupts the derivative estimate [6, 19]. Nonlinearity error can be removed by using higher order TIE [7], which performs polynomial fitting on each pixel's intensity vs.  $z$  plot, then extracts the first order derivative. This requires multiple images through-focus, which provides some inherent noise stability; however, there is a danger of over-fitting

(fitting to noise) and the optimal order to use will be phase dependent, making it a difficult parameter to optimize. The Savitzky-Golay differentiation filter (SGDF) TIE [20] was proposed to solve this trade-off between the order of fitting and noise in higher order TIE. It recovers phase images with different orders of polynomial fitting, then combines these phase images into a final reconstructed phase with band-pass filters. Limitations are that the fitting becomes unstable when the order is large [21] and the derivation assumes equally spaced focus steps, which we show here to be non-ideal.

We propose here a new TIE phase recovery method in which we also perform through-focus fitting of the data, except we do the fitting in Fourier space, rather than in real space. The advantage of this method can be understood from Fig. 2. On the left, we show a stack of through-focus intensity images, with a few example plots of the intensity vs.  $z$  behavior. These are the curves that would be fit to polynomials in higher order TIE [7]. It is easy to see that they do not conform to any particular functional form. On the right side of Fig. 2, we take 2D Fourier transforms of the intensity images at each of the defocus steps (showing here only the real part), then plot the intensity spectrum vs.  $z$  plots for a few example spatial frequencies. In contrast to the real space plots, these curves display a distinct sinusoidal behavior which can be predicted by the Contrast Transfer Function [22–24]. By using this a priori information of how light propagates in spatial frequency space, we can simultaneously deal with nonlinearity and noise in order to obtain a better derivative estimate. Here, we propose to fit the spectral data with Gaussian process (GP) regression [25]. GP regression does not require that the inputs are equally spaced and can suppress the unwanted high-frequency components in the fitted function by using the squared exponential covariance function (see details in Appendix A). The curves in Fig. 2 follow sinusoidal trajectories, but actual behavior may deviate from this model due to breaking of the small phase approximation or unmodeled coherence effects. GP regression is able to flexibly fit the data to variations in the model using the Contrast Transfer Function as a priori information to set a threshold for suppressing the unwanted high frequency components in the fitted function. In this way, the over-fitting to noise is avoided without knowing or fitting to the precise function of a sinusoid.

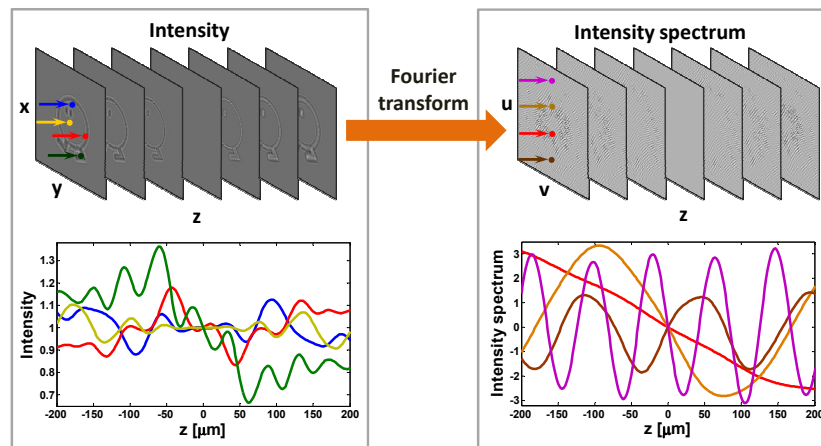


Fig. 2. Intensity evolution in frequency space is more predictable than in real space, demonstrating approximately sinusoidal behavior. (Top) Intensity images through-focus and their corresponding Fourier spectra's real part, from simulated data with  $dz = 1\mu m$ , size  $200\mu m \times 200\mu m$  using the object in Fig. 6. (Bottom left) Intensity variations over  $z$  for 4 sample pixels  $(x, y)$  denoted as dots above. (Bottom right) Real part of intensity spectrum over  $z$  for 4 spatial frequency  $(u, v)$  values denoted as dots above.

In studying the frequency behavior for through-focus images, we find that equally spaced defocus steps in  $z$  are not ideal. Previous methods generally use images equally spaced in  $z$ , though higher order TIE can accommodate arbitrary  $z$  steps [26]. In a recent study [27], an exponentially-growing measurement scheme is proposed to deal with the regularization problem. Here, we show that exponentially spaced defocus steps can provide a large reduction in the number of images required for a high quality phase result. We suggest an efficient scheme for choosing the  $z$  plane distances (within the limits of the motion stage), from the perspective of efficiently transferring phase information to intensity images, and show how our method can achieve high quality phase results with far fewer images than equally spaced schemes.

The rest of the paper is structured as follows. In Section 2, we apply GP regression to TIE phase recovery and propose an exponential spacing measurement scheme. In Section 3, we compare the error performance of the proposed algorithm with traditional TIE methods. In Section 4, we discuss the experimental results. We offer concluding remarks in Section 5.

## 2. Theory

### 2.1. Intensity spectrum fitting

We consider the intensity derivative estimation problem from the spatial frequency domain. When the phase  $\phi(x, y)$  is small, the variations of intensity over  $z$  can be approximated in frequency space by the Contrast Transfer Function [22–24]:

$$\mathcal{J}(u, v, z) = \delta(u, v) - 2U(u, v) \cos[\pi\lambda(u^2 + v^2)z] - 2\Phi(u, v) \sin[\pi\lambda(u^2 + v^2)z], \quad (3)$$

where  $\delta(u, v)$  denotes Dirac delta function, and  $\mathcal{J}(u, v, z)$ ,  $U(u, v)$ , and  $\Phi(u, v)$  are Fourier transforms of  $I(x, y, z)$ ,  $-\frac{1}{2} \ln I(x, y, 0)$ , and  $\phi(x, y)$ , respectively. For each spatial frequency  $(u_m, v_n)$ ,  $\mathcal{J}(u_m, v_n, z)$  follows a sinusoidal pattern that oscillates with frequency  $\pi\lambda(u_m^2 + v_n^2)$ . Thus, instead of fitting the intensity curves in real space to polynomials (as in higher order TIE [7]), we use the sinusoids in the intensity spectra through-focus as prior in the fitting.

### 2.2. TIE using Gaussian process regression

To fit the data to our model, we employ Gaussian process regression. We estimate the derivative of the intensity spectrum with respect to defocus  $\left. \frac{\partial \mathcal{J}(u, v, z)}{\partial z} \right|_{z=0}$  from  $\mathcal{J}(u, v, z_1), \dots, \mathcal{J}(u, v, z_N)$ , which are the 2D Fourier transforms of the measured intensity images. For each spatial frequency  $(u_m, v_n)$ ,  $\left. \frac{\partial \mathcal{J}(u_m, v_n, z)}{\partial z} \right|_{z=0}$  can be obtained from the continuous function fitted from the discrete data points  $\mathcal{J}(u_m, v_n, z_1), \dots, \mathcal{J}(u_m, v_n, z_N)$ .

GP regression is a suitable fitting method that allows us to fit the discrete points  $\mathcal{J}(u_m, v_n, z_1), \dots, \mathcal{J}(u_m, v_n, z_N)$  to a continuous sinusoidal function (see details in Appendix A). In the fitted function, frequency components of the axial oscillation that are higher than a threshold  $s_c$  can be suppressed by initializing appropriate hyper-parameters in the regression (see details in Appendix B). From the Contrast Transfer Function (Eq. (3)),  $\mathcal{J}(u_m, v_n, z)$ , as a function of  $z$ , does not contain frequency components higher than  $\pi\lambda(u_m^2 + v_n^2)$ . Therefore, the threshold can be set just above  $\pi\lambda(u_m^2 + v_n^2)$  in order to eliminate high-frequency noise while fitting over the inputs  $\mathcal{J}(u_m, v_n, z_1), \dots, \mathcal{J}(u_m, v_n, z_N)$ . Although the Contrast Transfer Function is derived under the weak phase assumption, it only serves as a guideline to pick a frequency threshold, above which variations are treated as noise. Thus, large phase variations can still be recovered, beyond the small phase approximation. Furthermore, by using GP regression, other prior knowledge about the sinusoidal pattern could be incorporated into the regression - for example, coherence effects of the illumination could be modeled in the Contrast Transfer

Table 1. Algorithm of GP TIE.

<p>For each frequency <math>v_m, v_n</math>:</p> <p>(1) Inputs are <math>N</math> data points <math>\mathcal{J}(v_m, v_n, z_1), \mathcal{J}(v_m, v_n, z_2), \dots, \mathcal{J}(v_m, v_n, z_N)</math>.</p> <p>(2) Set <math>s_c = \gamma\pi\lambda(u_m^2 + v_n^2)</math>, with <math>\gamma = 1.1 \sim 1.2</math>.</p> <p>(3) Initialize the hyper-parameters in GP regression, in order to suppress the frequency components higher than <math>s_c</math>.</p> <p>(4) Output is obtained by calculating <math>\left. \frac{\partial \mathcal{J}(v_m, v_n, z)}{\partial z} \right _{z=0}</math> through Eq. (17).</p>
--

Function as a product term, since reduced spatial coherence blurs defocused images and thus reduces the amplitude of the oscillations away from the focus plane.

The resulting algorithm of GP TIE is summarized in Table 1. The scaling factor  $\gamma$  in Step (2) of the table allows a trade-off between accuracy and noise filtering. Since the computational complexity of GP regression is proportional to the cubic of the number of measurements [25], the algorithm has a computational complexity of  $\mathcal{O}(N(N_z)^3)$ , where  $N$  is the number of pixels in one image, and  $N_z$  is the number of measurements. This algorithm is very amenable to parallel processing and can be implemented with Graphics Processing Units (GPUs) for near real-time performance, in conjunction with GPU-based TIE solvers [28]. Further, GP regression does not require the measuring positions  $z_1, z_2, \dots, z_N$  to be equally spaced, so any set of measurement positions along the axial dimension can be used as input. We show in the next session why this is not only convenient, but also leads to better phase recovery.

### 2.3. Exponential spacing measurement scheme

In looking at this problem from Fourier space, we see that equally spaced  $z$  steps are not the ideal measurement scheme. The choice of  $z$  distances is crucial, since it defines how much phase information from the object is transferred into intensity contrast in the defocused images [29]. Each  $z$  distance measurement can be thought of as having its own phase to intensity transfer function, with its own set of preferred spatial frequencies. The low-frequency information of phase is particularly poorly represented, as seen in the Laplacian inversion of Eq. (2). To better capture this low frequency phase information, images with large  $z$  are required, ideally out to the maximum range of the axial motion stage. However, the high-frequency components favor a small  $z$  and are important for recovering fine detail in the phase result. We would like to capture images *both* at the minimum  $z$  possible as well as at the maximum  $z$  possible. For linearly spaced schemes, this will require excessive data capture in between. To avoid this, we introduce an unequally spaced measurement scheme that balances these concerns. Since diffraction goes as  $(\Delta x)^2/\lambda z$  [18], with  $\Delta x$  relating to the object size, we expect the ideal spacing to follow a nonlinear trajectory.

To derive an ideal  $z$  sampling scheme, we start with Eq. (3) and extract the phase transfer function,  $g(u, v, z)$ :

$$g(u, v, z) = \sin[\pi\lambda(u^2 + v^2)z]. \quad (4)$$

This equation tells us how much phase information for a given spatial frequency is converted into intensity contrast, and thus relates to the signal-to-noise ratio (SNR) of the measurement. For a given position  $z$ , the phase transfer function is a sine function of  $\pi\lambda(u^2 + v^2)$  that will be maximized for particular spatial frequencies, as shown in Fig. 3. Since each object contains a different mix of spatial frequencies in its phase information, the theoretically optimal measurement planes are object-dependent [29], so it is not possible to determine them without knowing

the object itself. Thus, we aim not for the absolute optimal set of measurements, but rather for the optimal coverage of sensitivity across all spatial frequencies that pass through the imaging system, treating each equally.

Consider the goal of selecting a set of  $z$  planes such that the maximum phase measurement sensitivity (set by the value of the transfer function) is at least  $\alpha$  for the largest range of spatial frequencies possible. In the following, we demonstrate an exponential spacing measurement scheme that achieves this goal with the least number of measurement required. The highest frequency that can be recovered is set by the diffraction limit as  $NA/\lambda$ , where  $NA$  is the numerical aperture of the imaging system. This will define the minimum defocus distance from which we can capture relevant information. We define  $f = \pi\lambda(u^2 + v^2)$  and  $f \leq \pi\lambda(\frac{NA}{\lambda})^2$ . First, we select the minimum defocus distance  $z_1$  such that the sensitivity  $g(u, v, z_1)$  is  $\alpha$  at frequencies corresponding to the maximum frequency allowed  $f_1 = \pi\lambda(\frac{NA}{\lambda})^2$ , the solution of which is  $z_1 = \frac{\pi - \arcsin(\alpha)}{\pi(NA)^2} \lambda$ . Defocus steps smaller than  $z_1$  will not provide useful information and are thus unnecessary, though in practice the minimum  $z$  step size may be set by the axial motion stage. As shown by the blue curve in Fig. 3, the sensitivity remains above  $\alpha$  until  $f_2 = \frac{\arcsin(\alpha)\pi(NA)^2}{\lambda[\pi - \arcsin(\alpha)]}$ . Next, we select the second defocus distance  $z_2 = \beta z_1$ , where  $\beta = \frac{\pi - \arcsin(\alpha)}{\arcsin(\alpha)} > 1$ , such that the sensitivity  $g(u, v, z_2)$  is  $\alpha$  at  $f_2$ , and will remain at least  $\alpha$  for a range of frequencies, as illustrated by the green curve in Fig. 3. By induction, the optimal measurement scheme that satisfies a minimum phase measurement sensitivity  $\alpha$  should satisfy the following exponential relation for the defocus distances

$$z_{n+1} = \beta z_n. \quad (5)$$

The exponential spacing implies that a large  $z$  can be reached with far fewer measurements as compared to the equal-spacing measurement schemes. Large  $z$  images are crucial for transferring low-frequency phase information, and the exponential spacing scheme enables us to reach this without taking an excessive number of measurements or trading off high-frequency information. The larger the maximum  $z$ , the better the low-frequency result, and we desire to keep the minimum  $z$  as small as possible in order to maintain diffraction-limited phase resolution, both extremes being limited by the motion stage axial range. Accuracy can be traded off against number of images through the choice of  $\alpha$ . Larger values results in more accurate phase retrieval, at the cost of requiring more images.

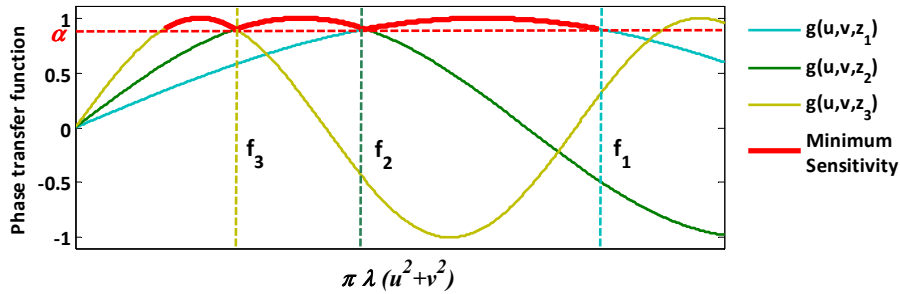


Fig. 3. Rationale for exponential spacing measurement scheme. Plot shows phase transfer functions for exponentially spaced  $z$  steps, which ensure a minimum sensitivity of  $\alpha$  across a range of frequencies.  $g(u, v, z_1)$ ,  $g(u, v, z_2)$ , and  $g(u, v, z_3)$  are the phase transfer functions at  $z_1$ ,  $z_2$ , and  $z_3$ , respectively. The minimum sensitivity plot shows the frequencies which are transferred at the sensitivity higher than  $\alpha$  by choosing  $z_1$ ,  $z_2$ , and  $z_3$ . Larger  $z$  brings more low-frequency sensitivity.

### 3. Simulations

We compare the performance of various phase recovery methods that use equal spacing and exponential spacing. In the simulation, the illumination wavelength is set as  $632.8\text{nm}$ , and each image has  $100 \times 100$  pixels (pixel size  $2\mu\text{m} \times 2\mu\text{m}$ ). **Equally Spaced Stack** has 9 intensity images simulated with a constant defocus step size of  $5\mu\text{m}$ . **Exponentially Spaced Stack** also has 9 images; however, they are exponentially spaced, at  $z$  positions around focus of  $\pm 5\mu\text{m}$ ,  $\pm 20\mu\text{m}$ ,  $\pm 80\mu\text{m}$  and  $\pm 320\mu\text{m}$  (see Fig. 4). Although both data sets use the same number of images, the exponentially spaced data set contains more information about the low-frequency phase information because it has a higher  $z$  range.

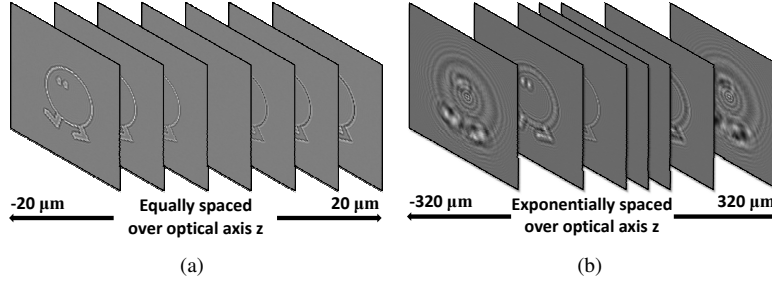


Fig. 4. Simulated data sets for equal and exponential spacing of  $z$  steps. (a) Equally Spaced Stack: images equally spaced by  $5\mu\text{m}$ . (b) Exponentially Spaced Stack: images exponentially spaced, with  $z$  of  $\pm 5\mu\text{m}$ ,  $\pm 20\mu\text{m}$ ,  $\pm 80\mu\text{m}$ , and  $\pm 320\mu\text{m}$ . The exponential spacing data contains more low-frequency phase information due to the larger range of  $z$  steps.

In order to assess the error performance, the intensity images of the equally and exponentially spaced stacks are corrupted by white noise with SNR ranging from  $18.5$  to  $8\text{ dB}$  (noise variance from  $0$  to  $0.02$ ). Fig. 5 shows the average mean square error (MSE) of the recovered phase over 50 trials as SNR decreases. For the Equally Spaced Stack, higher order TIE performs significantly worse than SGDF TIE, and GP TIE is slightly better than SGDF TIE. For the Exponentially Spaced Stack, we show the results for two possible choices of Higher order TIE:  $m = 9$ , which performs better in low noise, and  $m = 5$ , which performs better in high noise. GP TIE with exponential spacing clearly exhibits the lowest MSE. This can be explained by the fact that the exponentially spaced data contains more low-frequency phase content in the measurements than the equally spaced data, and there is no trade-off between noise and nonlinearity. Figure 6 gives an example of the recovered phase at SNR of  $11.1\text{dB}$ .

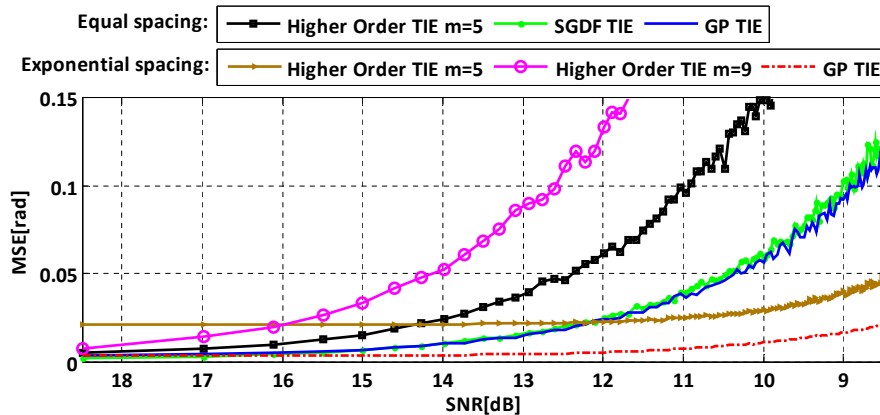


Fig. 5. Comparison of mean square error (MSE) in phase results for various methods as noise level increases. GP TIE with exponential spacing yields the best error performance.



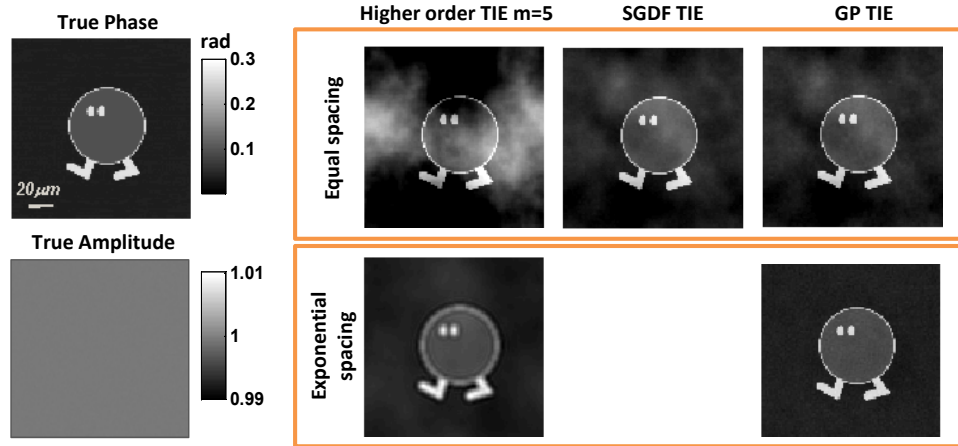


Fig. 6. Phase images recovered from simulated data with SNR of 11.1dB. (Top) Results of the Equally Spaced Stack, using higher order TIE ( $m=5$ , MSE 0.1194 in radian), SGDF TIE (0.0295), and the proposed GP TIE (0.0279). (Bottom) Results for Exponentially Spaced Stack, using Higher order TIE ( $m=5$ , MSE 0.0237) and GP TIE (0.0065). GP TIE using exponential spacing provides the best phase result.

#### 4. Experimental results

We tested our method experimentally in a microscope (magnification 20x, NA=0.5) with filtered white light illumination (center wavelength 650nm, 10nm bandwidth). **Data Set 1** comprises 129 images of human cheek cells, equally spaced by a constant small step size  $dz = 4\mu\text{m}$  over a large defocus range  $[-252\mu\text{m}$  to  $252\mu\text{m}]$  (Fig. 8(a)). In Fig. 7, the GP fitted intensity spatial frequency variations along the propagation direction  $z$  for 3 different  $(u, v)$  values are shown. Both the measured and fitted curves follow nearly sinusoidal patterns, as predicted by Eq. (3). Note that the delay of each sinusoid is dependent on the absorption of the object at the corresponding spatial frequency. When the frequency  $(u, v)$  is high, the intensity spectrum variations diminish for large defocus distance (see plot of  $(u_{60}, v_{60})$  in Fig. 7). This is due to the effect of partially coherent illumination, which will be the subject of future work.

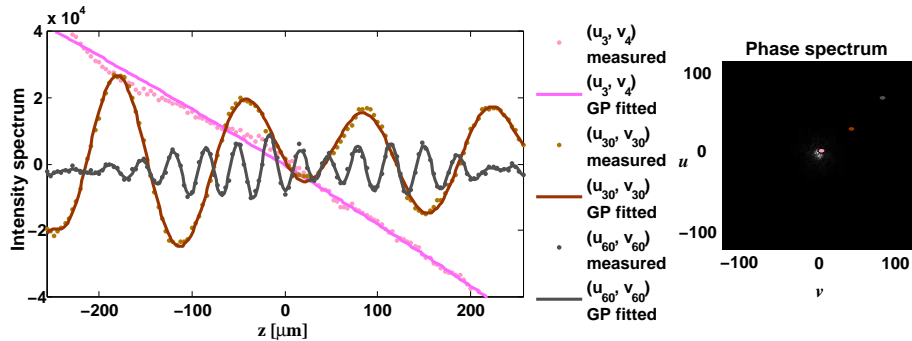


Fig. 7. (Left) GP regression of the intensity spectrum's real part over  $z$  for three sample frequency points  $(u, v)$  (Data Set 1). (Right) The frequency  $(u, v)$  of the three components depicted on real part of the recovered spectrum of phase (the image size is  $945 \times 888$  but only the central part of the spectrum is shown for clarity). According to Eq. (3), the values for  $\pi\lambda(u^2 + v^2)$  are  $0.029 \times 10^4 m^{-1}$ ,  $4.145 \times 10^4 m^{-1}$ , and  $17.299 \times 10^4 m^{-1}$ , respectively.

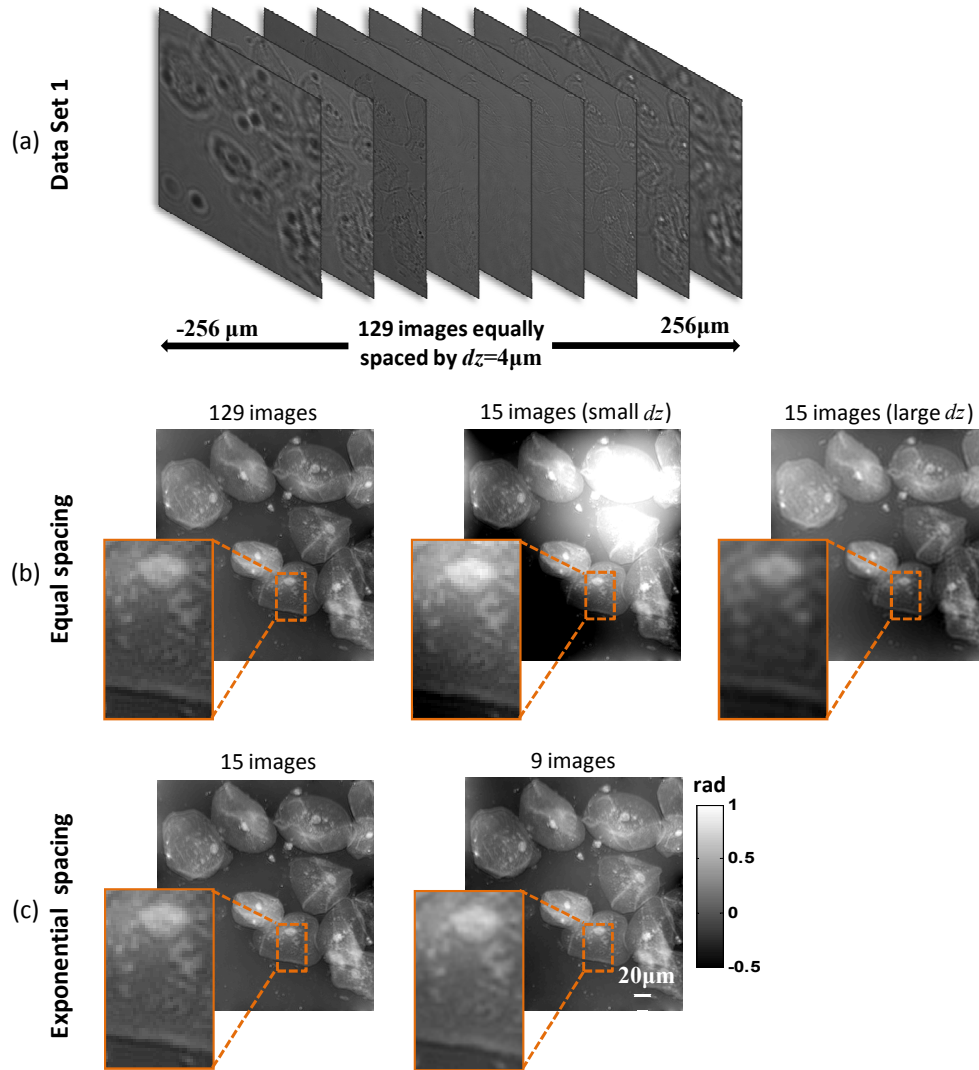


Fig. 8. Exponentially spaced defocus steps with GP TIE provide accurate phase results using less images. (a) Data Set 1: equally spaced  $dz = 4\mu\text{m}$ , from  $-256\mu\text{m}$  to  $256\mu\text{m}$ . Each image has  $945 \times 888$  pixels with effective size  $0.31\mu\text{m} \times 0.31\mu\text{m}$ . (b) Phase recovered with equally spaced  $z$  steps: (left) with all 129 images [ $-256\mu\text{m}$  to  $256\mu\text{m}$ ,  $dz = 4\mu\text{m}$ ]; (middle) subset of 15 images using minimum  $z$  step size [ $-28\mu\text{m}$  to  $28\mu\text{m}$ ,  $dz = 4\mu\text{m}$ ]; and (right) subset of 15 images using maximum  $z$  range [ $-252\mu\text{m}$  to  $252\mu\text{m}$ ,  $dz = 36\mu\text{m}$ ]. With equal spacing, there is a forced trade-off between low-frequency noise and high-frequency blurring, such that many images are required for good quality phase results. (c) Phase recovered with exponentially spaced  $z$  steps: (left) subset of 15 images ( $\beta = 2$ ), and (right) subset of 9 images ( $\beta = 4$ ). The minimum and maximum defocus distances are fixed at  $\pm 4\mu\text{m}$  and  $\pm 256\mu\text{m}$ , respectively.

With our exponentially spaced measurement scheme, GP TIE requires fewer images to be captured in order to obtain a high quality phase result. To demonstrate this, we compare the recovered phase by GP TIE from different subsets of Data Set 1 (Fig. 8) by sampling the image stack along  $z$  using various strategies. Figure 8(b) uses data subsets with equally spaced

z-planes. First, we show the best possible phase result, when all 129 images are used. With equally spaced planes, we have two possibilities for reducing the number of images: we can either reduce the defocus range (keeping the step size small) or increase the step size (reducing the defocus range). If we reduce the defocus range, the recovered phase becomes susceptible to low-frequency noise (see the middle of Fig. 8(b)). This is due to the fact that the low-frequency information of phase is not well captured at small defocus distances. If we instead increase the step size (keeping the defocus range large), high-frequency components are lost due to nonlinearity (see the right of Fig. 8(b)). In order to accurately capture both high and low-frequency information with the same reduced number of images, we need nonlinearly spaced measurements. Near the focus, the step size should be small, yet a large focus range can still be covered with only a few measurements in our exponentially spaced scheme. In Fig. 8(c), we extract images from Data Set 1 according to the exponential spacing scheme described in Section 2.3. As can be seen from Fig. 8(c), the phase results are free of low-frequency noise and also have high resolution. Even after further reducing the dataset to only 9 exponentially spaced images, we obtain a similar result to that with all 129 images. Thus, we have reduced the data capture requirement by more than an order of magnitude, without sacrificing quality.

Having shown that exponential spacing is advantageous, we now compare GP TIE with Higher order TIE, both of which are capable of using unequally spaced data. **Data Set 2** comprises 9 images of human cheek cells, exponentially spaced about focus at  $\pm 5.7\mu\text{m}$ ,  $\pm 11.4\mu\text{m}$ ,  $\pm 22.8\mu\text{m}$ , and  $\pm 45.6\mu\text{m}$  (Fig. 9(a)). The imaging system has 10x magnification and NA of 0.5. Each image has  $350 \times 360$  pixels of size  $0.62\mu\text{m} \times 0.62\mu\text{m}$ . Figure 9(b) shows the recovered phase of Data Set 2 obtained by Higher order TIE and GP TIE. We show the results for higher order TIE with the order of polynomial fitting  $m$  equal to 2, 3, and 4. The phase images of  $m = 2$  and 3 have small low-frequency noise but appear blurred. The phase of  $m = 4$  has strong contrast in some regions, but contains low-frequency noise. In contrast, GP TIE does not suffer from this tradeoff, so the phase recovered has less low-frequency noise and also exhibits high contrast. We can see details inside cells in the phase image recovered by GP TIE.

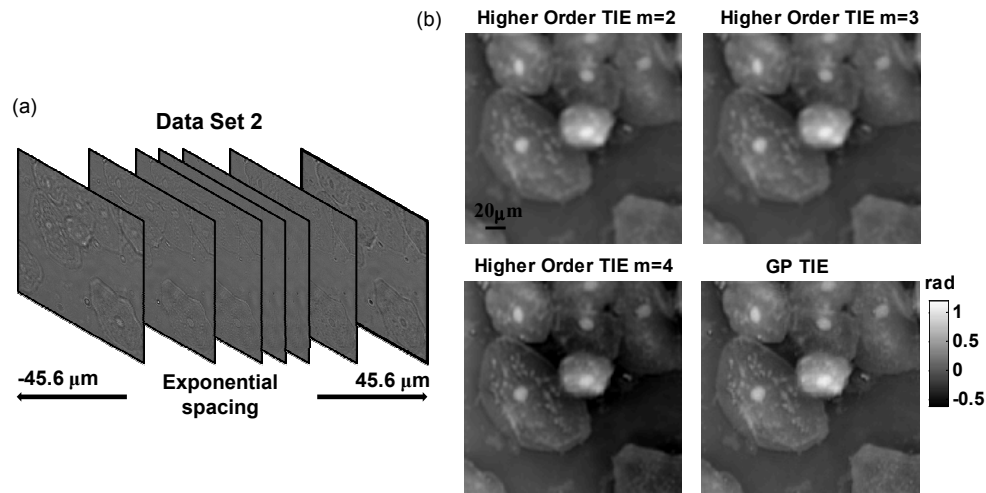


Fig. 9. (a) Data Set 2: exponentially spaced, with  $z$  of  $\pm 5.7\mu\text{m}$ ,  $\pm 11.4\mu\text{m}$ ,  $\pm 22.8\mu\text{m}$ , and  $\pm 45.6\mu\text{m}$ . Each image has  $350 \times 360$  pixels of size  $0.62\mu\text{m} \times 0.62\mu\text{m}$ . (b) Phase images of Data Set 2 by Higher order TIE ( $m = 2, 3$ , and 4) and GP TIE.

## 5. Conclusions

In this paper, we proposed a TIE phase recovery method that exploits prior knowledge of spatial frequency variations of the intensity in order to yield more accurate phase images from exponentially spaced images through-focus. Our proposed method estimates the intensity derivative by means of GP regression. It is robust and stable with noise, while incorporating the a priori knowledge of the sinusoidal patterns predicted when fitting spectrum variations over  $z$ . We derived an exponentially spaced measurement scheme which guarantees a minimum phase measurement sensitivity across the full range of available spatial frequencies with the least possible number of images required. In future work, we will extend the current technique to add more a priori information, such as partial coherence of the illumination and large phase effects. With the freedom to choose nonlinearly spaced measurement positions, we can develop new strategies to optimize measurements if any prior knowledge of the phase spectrum is known.

It should be noted that improving the TIE result by incorporating prior knowledge has been previously considered in [30–32]. The difference is that the previous approach considers priors over the phase [30, 31] or refractive index distribution [32], whereas here we use a prior on intensity spectral evolution to improve the intensity derivative estimate. We expect the approach considered here should have wide application as it not only can be directly apply to the special cases in [30–32], but also to more general situations without constraints on the phase distribution. If any reader is interested in this algorithm, open source code can be obtained by emailing the authors or visiting the website either [www.laurawaller.com](http://www.laurawaller.com) or [www.dauwels.com](http://www.dauwels.com).

## Appendix A: review of Gaussian process regression

We review the basics of Gaussian process regression [25]. Consider the problem of 2-D regression: given input/output pairs  $(z_n, f_n)$ , where  $n = 1, \dots, N$ , we would like to estimate  $f(z)$  at arbitrary position  $z$ . Under the Gaussian process assumption, the outputs  $f_n$  are drawn from the zero-mean Gaussian distribution with the covariance as a function of  $z_n$ :

$$(f_1, f_2, \dots, f_N | z_1, z_2, \dots, z_N) \sim \mathcal{N}(\mathbf{0}, \mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma_n \mathbf{I}), \quad (6)$$

where  $\mathbf{K}(\mathbf{Z}, \mathbf{Z})$  is the covariance matrix of the outputs given the input set  $\mathbf{Z}$ , and  $\sigma_n$  is the variance of additive Gaussian noise in outputs. Generally, the squared exponential covariance function is used to model the covariance matrix:

$$\mathbf{K}_{ij} = \sigma_f^2 \exp\left[-\frac{1}{2\ell^2}(z_i - z_j)^2\right], \quad (7)$$

where  $\mathbf{K}_{ij}$  is the element at coordinate  $(i, j)$  of the matrix  $\mathbf{K}(\mathbf{Z}, \mathbf{Z})$ . The parameters  $\sigma_f$ ,  $\ell$ , and  $\sigma_n$  in Eq. (6) are defined as the hyper-parameters of the GP model. We can write the joint distribution of the observed input/output pairs with the unknown value of  $f(z)$  at  $z$  as:

$$\begin{bmatrix} \mathbf{f} \\ f(z) \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma_n \mathbf{I} & \mathbf{K}(\mathbf{Z}, z) \\ \mathbf{K}(z, \mathbf{Z}) & \mathbf{K}(z, z) \end{bmatrix}\right), \quad (8)$$

where  $\mathbf{f} = [f_1, f_2, \dots, f_N]^T$ . The conditional distribution of the unknown output  $f(z)$  at  $z$  is calculated as:

$$(f(z) | f_1, f_2, \dots, f_N, z_1, z_2, \dots, z_N, z) \sim \mathcal{N}(\tilde{f}(z), \tilde{\mathbf{K}}), \quad (9)$$

where

$$\tilde{f}(z) = \mathbf{K}(z, \mathbf{Z})(\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma_n \mathbf{I})^{-1} \mathbf{f}, \quad (10)$$

$$\tilde{\mathbf{K}} = \mathbf{K}(z, z) - \mathbf{K}(z, \mathbf{Z})(\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma_n \mathbf{I})^{-1} \mathbf{K}(\mathbf{Z}, z). \quad (11)$$

For convenience, we can define a vector of functions  $\mathbf{h}(z)$ :

$$\mathbf{h}(z)^T = \mathbf{K}(z, \mathbf{Z})(\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma_n \mathbf{I})^{-1}. \quad (12)$$

Thus we have

$$\bar{f}(z) = \mathbf{h}(z)^T \mathbf{f}. \quad (13)$$

The predicted function  $\bar{f}(z)$  can be understood as a weighted combination of the shifted equivalent kernel  $h(z)$  [25, 33]. In [33], the Fourier transform of the equivalent kernel  $h(z)$  for the squared exponential covariance function is given as:

$$\tilde{h}_{SE}(\mathbf{s}) = \frac{1}{1 + b \exp(2\pi^2 \ell^2 |\mathbf{s}|^2)}, \quad (14)$$

where  $b = \sigma_n^2 / \rho(2\pi\ell^2)^{1/2}$  and  $\rho$  is the average number of observations per unit (for example length). When  $b$  is small,  $\tilde{h}_{SE}(\mathbf{s})$  is approximated by a step function. The rapid change from 1 to 0 happens at the point when

$$s_c^2 = \log(1/b) / (2\pi^2 \ell^2). \quad (15)$$

Therefore, the frequency components above the threshold  $s_c$  in the fitted function  $\bar{f}(z)$  are suppressed. From Eq. (15), we can set the desired threshold  $s_c$  by changing the hyper-parameters. The property of suppressing the unwanted high frequency components in the fitted function is applied in the regression of the intensity images in TIE.

## Appendix B: derivation of GP TIE

We have a 3D stack  $\mathcal{J}(u, v, z_1), \dots, \mathcal{J}(u, v, z_N)$ , which are the 2D Fourier transforms of the measured intensity images. Our goal is to use regression to estimate the first derivative of the intensity spectrum at  $z = 0$ ,  $\left. \frac{\partial \mathcal{J}(u, v, z)}{\partial z} \right|_{z=0}$ . Instead of doing 3D regression, we perform GP regression for each lateral spatial frequency  $(u_m, v_n)$  on  $N$  data points  $\mathcal{J}(u_m, v_n, z_1), \mathcal{J}(u_m, v_n, z_2), \dots, \mathcal{J}(u_m, v_n, z_N)$ . It is easy to observe that  $\mathcal{J}(u_m, v_n, z)$  is taken from  $\mathcal{J}(u, v, z)$  at the same frequency coordinates  $(u_m, v_n)$ . From Eq. (3),  $\mathcal{J}(u_m, v_n, z)$  follows a sinusoidal pattern with frequency  $\pi\lambda(u_m^2 + v_n^2)$ . The sinusoids prior is incorporated into the GP regression by setting the frequency threshold  $s_c$  as  $\pi\lambda(u_m^2 + v_n^2)$ . This is realized by setting appropriate hyper-parameters  $\sigma_f$ ,  $\sigma_n$ , and  $\ell$  in the regression. The hyper-parameters  $\sigma_f$  and  $\sigma_n$  are initialized to keep  $b$  in Eq. (14) small. Next, the parameter  $\ell$  is solved from Eq. (15) with  $s_c$  and  $\sigma_f$ ,  $\sigma_n$  already known. The frequency threshold  $s_c$  can be larger than  $\pi\lambda(u_m^2 + v_n^2)$  to allow trade-offs between accuracy and noise filtering.

Define  $\left. \frac{\partial \mathcal{J}(u_m, v_n, z)}{\partial z} \right|_{z=0}$  as the element of  $\left. \frac{\partial \mathcal{J}(u, v, z)}{\partial z} \right|_{z=0}$  at the coordinate  $(u_m, v_n)$ . From Eqs. (10)(12), the function  $\bar{\mathcal{J}}(u_m, v_n, z)$  fitted by GP regression is expressed as:

$$\bar{\mathcal{J}}(u_m, v_n, z) = \mathbf{h}(u_m, v_n, z)^T \mathbf{I}_{mn}, \quad (16)$$

where  $\mathbf{I}_{mn} = [\mathcal{J}(u_m, v_n, z_1), \mathcal{J}(u_m, v_n, z_2), \dots, \mathcal{J}(u_m, v_n, z_N)]^T$  and  $\mathbf{h}(u_m, v_n, z)$  can be obtained from Eq. (12). The fitted function  $\bar{\mathcal{J}}(u_m, v_n, z)$  is a function of the single variable  $z$ . Therefore, its first derivative at  $z = 0$ ,  $\left. \frac{\partial \bar{\mathcal{J}}(u_m, v_n, z)}{\partial z} \right|_{z=0}$ , is approximated by:

$$\left. \frac{\partial \bar{\mathcal{J}}(u_m, v_n, z)}{\partial z} \right|_{z=0} = \left. \frac{\partial \mathbf{h}(u_m, v_n, z)^T}{\partial z} \right|_{z=0} \mathbf{I}_{mn}, \quad (17)$$

where

$$\begin{aligned} \frac{\partial \mathbf{h}(u_m, v_n, z)^T}{\partial z} &= \left[ -\frac{\sigma_f^2}{\ell^2} (z - z_1) \exp\left[-\frac{1}{2\ell^2} (z - z_1)^2\right], \dots, -\frac{\sigma_f^2}{\ell^2} (z - z_N) \exp\left[-\frac{1}{2\ell^2} (z - z_N)^2\right] \right] \\ &\cdot (\mathbf{K}(\mathbf{Z}, \mathbf{Z}) + \sigma_n \mathbf{I})^{-1}. \end{aligned} \quad (18)$$

In Eq. (17), the vector  $\mathbf{I}_{mn}$  is known from the measurements, and the vector  $\frac{\partial \mathbf{h}(u_m, v_n, z)^T}{\partial z} \Big|_{z=0}$  can be derived from the GP regression with hyper-parameters already known. Therefore,  $\frac{\partial \mathcal{J}(u_m, v_n, z)}{\partial z} \Big|_{z=0}$  is estimated by performing GP regression over data points  $\mathcal{J}(u_m, v_n, z_1), \mathcal{J}(u_m, v_n, z_2), \dots, \mathcal{J}(u_m, v_n, z_N)$ .

By repeating the same process, we can obtain all of the frequency components for the first derivative of intensity spectrum at focus  $\frac{\partial \mathcal{J}(u, v, z)}{\partial z} \Big|_{z=0}$ , then the phase is recovered with the Laplacian inversion in Eq. (1)(2). Because the computational complexity of GP regression is proportional to the cubic of the number of measurements [25], the algorithm has a computational complexity of  $\mathcal{O}(N(N_z)^3)$ , where  $N$  is the number of pixels in one image, and  $N_z$  is the number of measurements. In order to save computational time, the frequency components which have similar  $u_m^2 + v_n^2$  values can share the same hyper-parameters and hence the same  $\frac{\partial \mathbf{h}(u_m, v_n, z)^T}{\partial z} \Big|_{z=0}$ .

## Acknowledgments

The authors would like to thank Sijia Liu and Jingyan Wang for helping with experiments, and Aamod Shanker, Daniel Shuldman, and Andrew R. Neureuther for helpful discussions.